



Complete Genome Sequences of Four Halophilic *Thermus thermophilus* Strains Isolated from Arima Hot Spring in Japan

 Kentaro Miyazaki,^{a,b} Natsuki Tomariguchi,^{a,c} Yuko Ueno^d

^aBioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Ibaraki, Japan

^bDepartment of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Chiba, Japan

^cFaculty of Life Sciences, Toyo University, Itakura, Gunma, Japan

^dFrontier Development Division, Leave a Nest Co., Ltd., Shinjuku, Tokyo, Japan

ABSTRACT We isolated four *Thermus thermophilus* strains from Arima Hot Spring in Japan. Complete genome sequencing revealed that they showed average nucleotide identities of $\geq 99.21\%$ to each other and to strains previously isolated from the same spot, but of $\leq 97.86\%$ to strains from geographically different spots in Japan, reflecting habitat-specific genomic conservation.

Thermus thermophilus, which grows optimally at around 70 to 75°C, was first isolated from Mine Hot Spring in Japan in 1968 (1, 2). So far, complete genome sequences have been reported for 13 strains (<https://www.ncbi.nlm.nih.gov/genome/browse#!/prokaryotes/461/>), of which six were isolated from Mine Hot Spring. Comparative genomic analyses (3) of these six strains (4–7) revealed high ($\geq 98.53\%$) average nucleotide identities (ANIs) to each other, supporting habitat-specific genomic conservation. In this study, to gain insight into the habitat-dependent evolutionary patterns of *T. thermophilus*, we investigated the genomes of four *T. thermophilus* strains that were newly isolated from Arima Hot Spring in Japan (34.7974 N, 135.2494 E). Arima Hot Spring lies in a mountain setting; however, the hot spring water, known as Arima-type brine (8), is unique in its high salinity and high solute concentrations. For bacterial isolation, a hot water sample (pH 6.6, 88°C, ca. 42,000 ppm Cl) was spread over marine agar (Sigma) plates. After incubation at 65°C for 24 h, four well-separated colonies were isolated; repetitive colony isolation by streaking ensured axenicity. These strains, designated AA1-1, AA2-2, AA3-7, and AK1, were subjected to genome analyses.

To prepare the genomic DNA, cells were grown in 5 ml of marine broth (Sigma) at 70°C for 24 h with vigorous shaking (200 rpm). The genomic DNA was purified using a blood and cell culture DNA Midi kit (Qiagen). For long-read sequencing, unshredded genomic DNA (1 μ g) was treated with a short-read eliminator kit (Circulomics) to remove fragments of < 10 kbp, and a library was constructed using a ligation sequencing kit (Oxford Nanopore Technologies [ONT]). Sequencing was performed using a GridION X5 system on a FLO-MIN106 R9.41 rev D flow cell (ONT). Base calling was conducted using Guppy v.4.0.11. The raw sequencing data (Table 1) were filtered ($Q < 10$; length, $< 1,000$ bases) using NanoFilt v.2.7.1 (9). For short-read sequencing, a library was constructed using an MGIEasy FS PCR-free DNA library prep set (MGI) with a ~ 400 to 500-bp insert. Paired-end sequencing (2×150 bases) was then performed on a DNBSEQ-400 instrument (MGI). The raw sequencing data (Table 1) were filtered ($Q < 30$; length, < 10 bases) using fastp v.0.20.1 (10). For all software, default parameters were used.

A hybrid assembly of the trimmed long- and short-read data was conducted using Unicycler v.0.4.8 (11), and the assembly was polished using Pilon v.1.23 (12). Each strain

Citation Miyazaki K, Tomariguchi N, Ueno Y. 2021. Complete genome sequences of four halophilic *Thermus thermophilus* strains isolated from Arima Hot Spring in Japan. *Microbiol Resour Announc* 10:e00874-21. <https://doi.org/10.1128/MRA.00874-21>.

Editor Julia A. Maresca, University of Delaware

Copyright © 2021 Miyazaki 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 Kentaro Miyazaki, miyazaki-kentaro@aist.go.jp.

Received 2 September 2021

Accepted 17 September 2021

Published 7 October 2021

TABLE 1 Characteristics and accession numbers of four halophilic *Thermus thermophilus* genomes

Strain	DNBSEQ (short read)			GridION (long read)			SRA accession no.	Length (bp)	GC content (%)	No. of coding sequences	Avg read depth (x)	GenBank accession no.
	No. of paired-end reads	Total length (Mb)	SRA accession no.	No. of reads	N_{50} (bp)	Total length (Mb)						
AA1-1 Chromosome pAA1-1b pAA1-1c	7,640,627	2,292	DRR308214	151,074	26,864	1,200	DRR308215	1,954,152 227,954 19,112	69.2 68.3 69.4	2,092 263 24	1,056.2 684.0 5,199.4	AP024926 AP024927 AP024928
AA2-2 Chromosome pAA2-2b pAA2-2c pAA2-2d	6,156,709	1,847	DRR308216	602,728	9,095	2,546	DRR308217	1,948,114 298,084 164,734 13,273	69.2 68.8 66.6 66.7	2,076 327 196 7	376.0 452.1 222.6 1,753.2	AP024929 AP024930 AP024931 AP024932
AA3-7 Chromosome pAA3-7b pAA3-7c pAA3-7d	4,675,366	1,402	DRR308218	426,341	7,807	1,781	DRR308219	1,980,678 213,325 143,809 1,4256	69.1 69.3 68.2 67.7	2,139 243 72 8	273.3 280.2 212.9 1,419.4	AP024933 AP024934 AP024935 AP024936
AK1 Chromosome pAK1b pAK1c	7,754,567	2,326	DRR308220	294,172	9,819	1,614	DRR308221	1,946,623 116,199 91,082	69.2 68.7 68.2	2,096 135 20	767.6 1,716.6 497.8	AP024937 AP024938 AP024939

contained a single circular chromosome and two or three circular plasmids; circularity was confirmed using Unicycler. Automatic annotation was conducted using DFAST v.1.4.0 (13); the genomic features are summarized in Table 1. A JSpecies analysis (14) revealed that the genome sequences of bacterial strains from Arima Hot Spring (four strains presented in this study and two previously documented strains, AA2-20 [15] and AA2-29 [15]) showed ANIs of $\geq 99.21\%$ to each other but of $\leq 97.86\%$ to six strains originating from Mine Hot Spring (HB8, HB27 [4], HC11 [5], HB5002 [7], HB5008 [7], and HB5018 [6]).

Data availability. All four *T. thermophilus* strains reported in this paper are associated with BioProject accession number [PRJDB7414](https://www.ncbi.nlm.nih.gov/bioproject/PRJDB7414). The genome sequences and raw sequencing data are available under the accession numbers listed in Table 1.

ACKNOWLEDGMENTS

This work was partly supported by the following grants awarded to K.M. from the Japan Society for the Promotion of Science (JSPS): a Grant-in-Aid for Scientific Research (A) (19H00936) and a Grant-in-Aid for Challenging Research (Pioneering) (19H05538).

We thank the following elementary school students for preliminary characterization of AK1: Shoko Endo, Yuto Nakajima, Kento Makuuchi, Satoki Matsumoto, Kohaku Abe, Kokomi Uyama, Kaon Mimura, and Sahoyo Yogo (Discovery Lab, Leave a Nest Co., Ltd.).

REFERENCES

- Oshima T, Imahori K. 1971. Isolation of an extreme thermophile and thermostability of its transfer ribonucleic acid and ribosomes. *J Gen Appl Microbiol* 17:513–517. <https://doi.org/10.2323/jgam.17.513>.
- Oshima T, Imahori K. 1974. Description of *Thermus thermophilus* (Yoshida and Oshima) comb. nov., a nonsporulating thermophilic bacterium from a Japanese thermal spa. *Int J Sys Evol Microbiol* 24:102–112. <https://doi.org/10.1099/00207713-24-1-102>.
- Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 9:5114. <https://doi.org/10.1038/s41467-018-07641-9>.
- Henne A, Bruggemann H, Raasch C, Wiezer A, Hartsch T, Liesegang H, Johann A, Lienard T, Gohl O, Martinez-Arias R, Jacobi C, Starkuviene V, Schlenczcek S, Dencker S, Huber R, Klenk H-P, Kramer W, Merkl R, Gottschalk G, Fritz H-J. 2004. The genome sequence of the extreme thermophile *Thermus thermophilus*. *Nat Biotechnol* 22:547–553. <https://doi.org/10.1038/nbt956>.
- Miyazaki K. 2019. Complete genome sequencing of *Thermus thermophilus* strain HC11, isolated from Mine Geyser in Japan. *Microbiol Resour Announc* 8:e00873-19. <https://doi.org/10.1128/MRA.00873-19>.
- Miyazaki K, Moriya T, Nemoto N, Oshima T, Yura K, Bessho Y. 2021. Complete genome sequence of *Thermus thermophilus* strain HB5018, isolated from Mine Hot Spring in Japan. *Microbiol Resour Announc* 10:e00039-21. <https://doi.org/10.1128/MRA.00039-21>.
- Miyazaki K, Moriya T, Tokito N, Oshima T, Yura K, Bessho Y. 2021. Complete genome sequences of *Thermus thermophilus* strains HB5002 and HB5008, isolated from Mine Hot Spring in Japan. *Microbiol Resour Announc* 10:e00272-21. <https://doi.org/10.1128/MRA.00272-21>.
- Kusuda C, Iwamori H, Nakamura H, Kazahaya K, Morikawa N. 2014. Arima hot spring waters as a deep-seated brine from subducting slab. *Earth Planets Space* 66:119. <https://doi.org/10.1186/1880-5981-66-119>.
- De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 34:2666–2669. <https://doi.org/10.1093/bioinformatics/bty149>.
- Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ pre-processor. *Bioinformatics* 34:i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>.
- Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>.
- Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. <https://doi.org/10.1371/journal.pone.0112963>.
- Tanizawa Y, Fujisawa T, Kaminuma E, Nakamura Y, Arita M. 2016. DFAST and DAGA: Web-based integrated genome annotation tools and resources. *Biosci Microbiota Food Health* 35:173–184. <https://doi.org/10.12938/bmfh.16-003>.
- Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J. 2016. JSpeciesWS: a Web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics* 32:929–931. <https://doi.org/10.1093/bioinformatics/btv681>.
- Miyazaki K, Tomariguchi N. 2019. Complete genome sequences of *Thermus thermophilus* strains AA2-20 and AA2-29, isolated from Arima Onsen in Japan. *Microbiol Resour Announc* 8:e00820-19. <https://doi.org/10.1128/MRA.00820-19>.