AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequences of *Thermus thermophilus* Strains AA2-20 and AA2-29, Isolated from Arima Onsen in Japan

Kentaro Miyazaki,[a,b] Natsuki Tomariguchi[a,c*]

[a]Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Ibaraki, Japan
[b]Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Chiba, Japan
[c]Faculty of Life Sciences, Toyo University, Itakura, Gunma, Japan

ABSTRACT We isolated halophilic and thermophilic *Thermus thermophilus* strains AA2-20 and AA2-29 from nonvolcanic, oceanic Arima Onsen (hot spring) in Japan. Here, we report the complete genome sequences of these organisms to gain insights into halophilicity.

Thermus thermophilus was first isolated by Oshima and Imahori from Mine Onsen (hot spring) in Japan (1). The bacterium grows optimally at 70 to 75°C, and its thermophilic nature aids the study of biomacromolecules, such as enzymes (2–7) and ribosomes (8, 9), whose functional investigations are often limited due to the lack of stability. We isolated *T. thermophilus* strains AA2-20 and AA2-29 from nonvolcanic, oceanic Arima Onsen in Kobe, Japan (10). The environmental sample was spread over TYSS (1% [wt/vol] tryptone, 0.5% [wt/vol] yeast extract, and 4% [wt/vol] sea salts [Sigma]) agar plates. After overnight incubation at 65°C, two colonies with red or yellow pigmentation (named AA2-20 and AA2-29, respectively) were isolated. DNA sequencing analysis of the rRNA gene suggested that the two strains belong to *T. thermophilus* (99% identity to the *T. thermophilus* HB8 gene). Both the strains displayed limited growth in low-salt media, such as Lennox LB (1% [wt/vol] tryptone, 0.5% [wt/vol] yeast extract, and 0.5% [wt/vol] NaCl) broth, thereby suggesting their halophilic nature.

To prepare genomic DNA, the two strains were grown in TYSS broth at 65°C until saturation. Genomic DNA was extracted using the Nexttec 1-step DNA isolation kit for bacteria (Nexttec Biotechnologie) according to the manufacturer's instructions.

For long-read sequencing, the extracted genomic DNA was passed through a Circulomics short-read eliminator kit to remove short fragments. Sequencing was performed using a GridION X5 system (Oxford Nanopore Technologies [ONT]); the library was constructed from unfragmented genomic DNA (1.0 μg) using a ligation sequencing kit (ONT) and applied to a FLO-MIN106 R9.41 flow cell (ONT). The long-read sequences, which were base called using Guppy v.3.0.3, further generated 278,671 reads (1,550 Mb) with an average length of 5,563 bases during a 24-h run time (the data are associated with quality-filtered reads with average phred quality values of >8.0, obtained using NanoFilt v.2.3.0 [11]). The longest read had 91,228 bases.

For short-read sequencing, the extracted genomic DNA was used for library preparation with the Nextera DNA library preparation kit (Illumina). Prepared libraries were subjected to 100-bp paired-end sequencing on the Illumina HiSeq 2500 platform. Adapters and low-quality sequencing data were trimmed using fastp v.0.14.1 (12), and 6.9 million paired-end reads (660 Mb) with an average length of 96 bases were obtained.

Citation 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.

Editor Frank J. Stewart, Georgia Institute of Technology

Address correspondence to Kentaro Miyazaki, miyazaki-kentaro@aist.go.jp.

* Present address: Natsuki Tomariguchi, Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Chiba, Japan.

Received 9 July 2019
Accepted 11 July 2019
Published 1 August 2019

TABLE 1 Genome statistics and genomic features of *T. thermophilus* isolates

Strain name	Chromosome or plasmid	Length (bp)	GC content (%)	No. of CDSs[a]	No. of rRNAs	No. of tRNAs	DDBJ accession no.
AA2-20	Chromosome	1,928,002	69.2	2,045	6	51	AP019792
	Plasmid	219,630	68.9	258	0	0	AP019793
AA2-29	Chromosome	19,290,112	69.2	2,042	6	51	AP019794
	Plasmid	215,278	69.1	252	0	0	AP019795

[a] CDSs, coding DNA sequences.

Hybrid assembly of long-read and short-read data was conducted using Unicycler v.0.4.7 (13), followed by a final polishing with Pilon v.1.23 (14), which resulted in the production of a single circular chromosome and a single circular plasmid. Automated annotation was performed using DFAST v.1.1.0 (15). Default parameters were used for all software unless otherwise noted.

The genome statistics and genomic features are listed in Table 1. A JSpecies analysis (16) revealed that the genomic sequences of AA2-20 and AA2-29 were nearly identical (97.2% average nucleotide identity) with no large gaps or rearrangement. Additionally, these sequences showed high average nucleotide identity (89.6%) with the genomic sequence of *T. thermophilus* JL-18 (GenBank accession number NC_017587), which was isolated from freshwater hot springs in Great Basin National Park, USA. The strains AA2-20 and AA2-29 differed from JL-18 with respect to large rearrangements in their sequences, including inversions and indels.

Data availability. The accession numbers for the complete genome sequences of *T. thermophilus* strains AA2-20 and AA2-29 are listed in Table 1. Raw sequencing data have been deposited in the DDBJ SRA database under the accession number DRA008626 (BioProject number PRJDB7414, BioSample numbers SAMD00177813 [AA2-20] and SAMD00177814 [AA2-29]). The *T. thermophilus* strains AA2-20 and AA2-29 have been deposited in the RIKEN Bioresource Center, JCM, under the accession numbers JCM 33047 and JCM 33048, respectively.

ACKNOWLEDGMENT

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

REFERENCES

1. 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 Syst Evol Microbiol 24:102–112. https://doi.org/10.1099/00207713-24-1-102.
2. Yamada T, Akutsu N, Miyazaki K, Kakinuma K, Yoshida M, Oshima T. 1990. Purification, catalytic properties, and thermal stability of *threo*-Ds-3-isopropylmalate dehydrogenase coded by *leuB* gene from an extreme thermophile, *Thermus thermophilus* strain HB8. J Biochem 108:449–456. https://doi.org/10.1093/oxfordjournals.jbchem.a123220.
3. Miyazaki K, Kakinuma K, Terasawa H, Oshima T. 1993. Kinetic analysis on the substrate specificity of 3-isopropylmalate dehydrogenase. FEBS Lett 332:35–36. https://doi.org/10.1016/0014-5793(93)80477-C.
4. Miyazaki K, Kadono S, Sakurai M, Moriyama H, Tanaka N, Oshima T. 1994. Chemical modification and site-directed mutagenesis of Tyr36 of 3-isopropylmalate dehydrogenase from *Thermus thermophilus* HB8. Protein Eng Des Sel 7:99–102. https://doi.org/10.1093/protein/7.1.99.
5. Miyazaki K, Yaoi T, Oshima T. 1994. Expression, purification, and substrate specificity of isocitrate dehydrogenase from *Thermus thermophilus* HB8. Eur J Biochem 221:899–903. https://doi.org/10.1111/j.1432-1033.1994.tb18805.x.
6. Miyazaki K. 2005. A hyperthermophilic laccase from *Thermus thermophilus* HB27. Extremophiles 9:415–425. https://doi.org/10.1007/s00792-005-0458-z.
7. Miyazaki K, Oshima T. 1994. Co-enzyme specificity of 3-isopropylmalate dehydrogenase from *Thermits thermophilus* HB8. Protein Eng Des Sel 7:401–403. https://doi.org/10.1093/protein/7.3.401.
8. Tocilj A, Schlunzen F, Janell D, Gluhmann M, Hansen HA, Harms J, Bashan A, Bartels H, Agmon I, Franceschi F, Yonath A. 1999. The small ribosomal subunit from *Thermus thermophilus* at 4.5 Å resolution: pattern fittings and the identification of a functional site. Proc Natl Acad Sci U S A 96:14252–14257. https://doi.org/10.1073/pnas.96.25.14252.
9. Wimberly BT, Brodersen DE, Clemons WM, Jr, Morgan-Warren RJ, Carter AP, Vonrhein C, Hartsch T, Ramakrishnan V. 2000. Structure of the 30S ribosomal subunit. Nature 407:327–339. https://doi.org/10.1038/35030006.
10. Kusuda C, Iwamori H, Nakamura H, Kazahaya K, Morikawa N. 2014. Arima hot spring waters as a deep-seated brine from subducting slab. Earth Planet Space 66:119. https://doi.org/10.1186/1880-5981-66-119.
11. 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.
12. Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890. https://doi.org/10.1093/bioinformatics/bty560.
13. 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.
14. 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.
15. 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.
16. 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.