



# Complete Genome Sequences of *Rhodothermus marinus* Strains AA2-13 and AA3-38, Isolated from Arima Onsen Hot Spring in Japan

Natsuki Tomariguchi,<sup>a,b\*</sup>  Kentaro Miyazaki<sup>a,c</sup>

<sup>a</sup>Department of Life Science and Biotechnology, Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki, Japan

<sup>b</sup>Faculty of Life Sciences, Toyo University, Itakura, Gunma, Japan

<sup>c</sup>Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Chiba, Japan

**ABSTRACT** We isolated *Rhodothermus marinus* strains AA2-13 and AA3-38 from Arima Onsen, a hot spring in Japan, and sequenced their genomes. The average nucleotide identity between their genomes was 99.2%, and that with the genome of *R. marinus* strain DSM 4252<sup>T</sup> (isolated from Iceland) was ~95.2%, suggesting close relationships among these strains.

The *Rhodothermus* genus belongs to the phylum *Bacteroidetes* and has been found in geothermal habitats. *Rhodothermus* has gained much attention as a source of industrial enzymes (1) and as a target for protein engineering (2) because of its thermostability. Since the discovery of *Rhodothermus marinus* DSM 4252 (the type strain of *R. marinus*) from the Blue Lagoon geothermal spa in Iceland (3), several *Rhodothermus* species have been isolated from high-temperature environments (4, 5).

We isolated thermophilic and halophilic *R. marinus* strains from the nonvolcanic, oceanic Arima Onsen (a hot spring) in Kobe, Japan, on 20 March 2017 (6). The environmental sample was spread over marine agar plates (Difco) and incubated at 65°C. After 48 h, dark-pink colonies that appeared on the plates were examined by sequencing of the full length of the 16S rRNA gene, which was amplified using the primers Thermus\_1F and Thermus\_1521R (7). Many of the colonies were associated with *Thermus thermophilus* (8), one with *Rubrobacter xylanophilus* (9), and the others with *R. marinus* (~99% identity with the *R. marinus* DSM 4252<sup>T</sup> 16S rRNA gene [GenBank accession number [NR\\_029282.2](https://www.ncbi.nlm.nih.gov/nuccore/NR_029282.2)]). Among the *R. marinus* strains, two strains, AA2-13 and AA3-38, were subjected to genome analysis. The strains were grown in marine broth (Difco) at 65°C for 24 h. Genomic DNA was purified using the 1-Step DNA isolation kit for bacteria (nexttec Biotechnologie), according to the manufacturer's instructions.

Genome analysis was conducted by employing a combination of Illumina short-read and Oxford Nanopore Technologies (ONT) long-read sequencing. For short-read sequencing, genomic DNA was sheared to 100-bp fragments, adaptor ligated, and amplified to generate DNA libraries using the Nextera DNA library preparation kit (Illumina). The libraries were subjected to paired-end sequencing (2 × 100 bases) on the Illumina HiSeq 2500 platform. A total of 8,154,381 reads (815 Mb) and 7,894,203 reads (789 Mb) were produced for AA2-13 and AA3-38, respectively. Adapters and low-quality data were trimmed using fastp v.0.19.5 (10). Default parameter settings were applied for software analyses throughout this study.

For long-read sequencing, short genomic fragments were removed using a short-read eliminator kit (Circulomics). Sequencing was performed using a GridION X5 sequencer (ONT). Unfragmented genomic DNA (1 μg) was used for library construction with a ligation sequencing kit (ONT). The library was then analyzed in a FLO-MIN106

**Citation** Tomariguchi N, Miyazaki K. 2020. Complete genome sequences of *Rhodothermus marinus* strains AA2-13 and AA3-38, isolated from Arima Onsen hot spring in Japan. *Microbiol Resour Announc* 9:e01475-19. <https://doi.org/10.1128/MRA.01475-19>.

**Editor** Frank J. Stewart, Georgia Institute of Technology

**Copyright** © 2020 Tomariguchi and Miyazaki. 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](mailto: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** 28 November 2019

**Accepted** 13 December 2019

**Published** 9 January 2020

**TABLE 1** Genome statistics and features of *R. marinus* isolates

Strain	Chromosome or plasmid	Length (bp)	GC content (%)	No. of coding sequences	No. of rRNAs	No. of tRNAs	GenBank accession no.
AA2-13	Chromosome	3,440,105	64.1	2,981	3	45	AP019796
AA3-38	Chromosome	3,430,829	64.1	2,980	3	46	AP019797
DSM 4252 <sup>T</sup>	Chromosome	3,261,604	64.5	2,797	3	45	NC_013501.1
	Plasmid (pRMAR01)	125,133	58.2	107	0	0	NC_013502.1

R9.41 flow cell (ONT) for 12 h. Base calling was conducted using Guppy v.3.0.3 to generate 140,887 reads (2.09 Gb) and 467,442 reads (2.87 Gb), with average lengths of 14,800 bases (longest read, 269,105 bases;  $N_{50}$ , 37,152 bases) and 6,143 bases (longest read, 156,140 bases;  $N_{50}$ , 20,011 bases) for AA2-13 and AA3-38, respectively. The raw read data were filtered (average Phred quality values, >8.0) using NanoFilt v.2.3.0 (11).

The hybrid assembly of long- and short-read data was conducted using Unicycler v.0.4.4 (12), followed by final polishing with Pilon v.1.23 (13). For both strains, a single circular chromosome was generated. Unlike in DSM 4252<sup>T</sup>, no plasmids were found in AA2-13 or AA3-38. Automated annotation was performed using DFAST v.1.1.0 (14). The features of the genomes are presented in Table 1. Average nucleotide identity (ANI) analysis was carried out using the Web-based OrthoANu calculator (15), which revealed that the genome sequences of AA2-13 and AA3-38 were nearly identical (99.2% ANI), with no large gaps or rearrangements. As of 28 November 2019, two complete genome sequences had been determined for *R. marinus* (strains DSM 4252<sup>T</sup> [GenBank accession number NC\_013501.1] and SG0.5JP17-172 [GenBank accession number NC\_015966.1]). Despite the distant habitats, strains AA2-13 and AA3-38 showed gene arrangements similar to that of strain DSM 4252<sup>T</sup>, with ANIs of ~95.2%, suggesting close relationships among these strains.

**Data availability.** The GenBank accession numbers for the complete genome sequences of *R. marinus* strains AA2-13 and AA3-38 are AP019796 and AP019797, respectively (Table 1). Raw sequencing data were deposited in the DDBJ Sequence Read Archive under accession number DRA008620 (GenBank BioProject number PRJDB8496 and BioSample accession numbers SAMD00177810 [strain AA2-13] and SAMD00177811 [strain AA3-38]).

## ACKNOWLEDGMENTS

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

## REFERENCES

- Bjornsdottir SH, Blondal T, Hreggvidsson GO, Eggertsson G, Petursdottir S, Hjorleifsdottir S, Thorbjarnardottir SH, Kristjansson JK. 2006. *Rhodothermus marinus*: physiology and molecular biology. *Extremophiles* 10: 1–16. <https://doi.org/10.1007/s00792-005-0466-z>.
- Kan SB, Lewis RD, Chen K, Arnold FH. 2016. Directed evolution of cytochrome *c* for carbon-silicon bond formation: bringing silicon to life. *Science* 354:1048–1051. <https://doi.org/10.1126/science.aah6219>.
- Alfredsson GA, Kristjansson JK, Hjorleifsdottir S, Stetter KO. 1988. *Rhodothermus marinus*, gen. nov., sp. nov., a thermophilic, halophilic bacterium from submarine hot springs in Iceland. *Microbiology* 134:299–306. <https://doi.org/10.1099/00221287-134-2-299>.
- Sako Y, Takai K, Ishida Y, Uchida A, Katayama Y. 1996. *Rhodothermus obamensis* sp. nov., a modern lineage of extremely thermophilic marine bacteria. *Int J Syst Bacteriol* 46:1099–1104. <https://doi.org/10.1099/00207713-46-4-1099>.
- Marteinsson VT, Bjornsdottir SH, Bienvenu N, Kristjansson JK, Birrien JL. 2010. *Rhodothermus profundus* sp. nov., a thermophilic bacterium isolated from a deep-sea hydrothermal vent in the Pacific Ocean. *Int J Syst Evol Microbiol* 60:2729–2734. <https://doi.org/10.1099/ijs.0.012724-0>.
- 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>.
- Miyazaki K, Tomariguchi N. 2019. Occurrence of randomly recombined functional 16S rRNA genes in *Thermus thermophilus* suggests genetic interoperability and promiscuity of bacterial 16S rRNAs. *Sci Rep* 9:11233. <https://doi.org/10.1038/s41598-019-47807-z>.
- Miyazaki K, Tomariguchi N, 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>.
- Tomariguchi N, Miyazaki K, Tomariguchi N, Miyazaki K. 2019. Complete genome sequence of *Rubrobacter xylanophilus* strain AA3-22, isolated from Arima Onsen in Japan. *Microbiol Resour Announc* 8:e00818-19. <https://doi.org/10.1128/MRA.00818-19>.
- 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>.
- 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. 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>.
  13. 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>.
  14. 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>.
  15. Yoon SH, Ha SM, Lim J, Kwon S, Chun J. 2017. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek* 110:1281–1286. <https://doi.org/10.1007/s10482-017-0844-4>.