



# Complete Genome Sequence of the Hyperthermophilic and Acidophilic Archaeon *Saccharolobus caldissimus* Strain HS-3<sup>T</sup>

 Hiroyuki D. Sakai,<sup>a</sup>  Norio Kurosawa<sup>a</sup>

<sup>a</sup>Department of Science and Engineering for Sustainable Innovation, Faculty of Science and Engineering, Soka University, Hachioji, Tokyo, Japan

**ABSTRACT** The complete genome sequence of the hyperthermophilic archaeon *Saccharolobus caldissimus* strain HS-3<sup>T</sup> was determined. The genome is 3,075,795 bp with a GC content of 32.9%. Genes for the complete semiphosphorylative Entner-Doudoroff pathway, gluconeogenesis, tricarboxylic acid cycle, pentose phosphate pathway, and 3-hydroxypropionate 4-hydroxybutyrate cycle were present in the genome.

*Saccharolobus caldissimus* strain HS-3<sup>T</sup> is a facultatively anaerobic, hyperthermophilic, and acidophilic archaeon belonging to the order *Sulfolobales* (1). It grows at 65 to 93°C (optimal, 85°C) and pH 1.5 to 6.0 (optimal, 3.0) and is capable of utilizing a variety of organic substrates as a carbon source. Although physiological properties of *S. caldissimus* were examined in detail (1), the genome sequence of this species has not been determined. Here, we report the complete genome sequence of *S. caldissimus* strain HS-3<sup>T</sup>.

Strain HS-3<sup>T</sup> was isolated as described previously (1). The strain was cultivated in modified Brock's basal salt medium (2) supplemented with 1 g/L yeast extract (pH 3.0; 80°C). Microbial cells were collected as a pellet by centrifugation (12,000 × *g*, 4°C, and 15 min) from 200 mL of the culture. Genomic DNA was extracted from the cell pellet using Genomic-tips 100/G (Qiagen). Genome sequencing was performed using the same protocol as described previously (3). For short-read sequencing, a DNA library was prepared with the NEBNext Ultra II FS DNA library preparation kit for Illumina (New England BioLabs) and then subjected to an Illumina NovaSeq 6000 sequencer (2 × 150 bp). For long-read sequencing, a DNA library was prepared with kits SQK-LSK109 and EXP-NBD104 (Oxford Nanopore Technologies [ONT]), following the protocol described by ONT (NBE\_9065\_v109\_revZ\_14Aug2019). DNA fragments of 3 kb or longer were enriched by the protocol. Long-read sequencing was conducted on a MinION sequencer with an R9 flow cell. Base calling was carried out by MinKNOW v.4.2.8 software. As a result, a total of 8,375,640 short reads (1,256,346,000 bp) and 62,599 long reads (336,859,512 bp [*N*<sub>50</sub>, 15,506 bp]) were obtained. The short reads and long reads were quality filtered using fastp v.0.20.1 (4) and Filtlong v.0.2.0. (<https://github.com/rrwick/Filtlong>) (Table 1). After the quality-filtering step, a total of 8,322,466 short reads (1,211,007,369 bp) and 56,658 long reads (294,131,198 bp [*N*<sub>50</sub>, 8,652 bp]) were obtained. Using all the quality-filtered reads as the input, genome assembly was carried out with Unicycler v.0.4.8 (5) (Table 1). Genome annotation was conducted by DFAST v.1.4.0 (6). Default parameters were used for all software unless otherwise specified.

As a result of assembly, the complete genome sequence of strain HS-3<sup>T</sup> was obtained as a circular contig composed of 3,075,795 bp with a GC content of 32.9%. The genome encoded 3,307 coding DNA sequences (CDSs), a single copy of the rRNA operon (16S-23S), 47 tRNAs, and 7 CRISPR repeats. The genome contains genes for complete glycolysis (semiphosphorylative Entner-Doudoroff pathway), gluconeogenesis, tricarboxylic acid cycle, and pentose phosphate pathway. The archaeal carbon fixation pathway 3-hydroxypropionate 4-hydroxybutyrate cycle (7) was also found in the genome. Genes related to oligosaccharide transporters (e.g., *malK*, *msmXK*, *smoK*, *msiK*, and *msmX*) and monosaccharide transporters (e.g., *glcSUTV* and *malK*), were present, which was in good agreement with physiological characteristics of HS-3<sup>T</sup>

**Editor** Irene L. G. Newton, Indiana University, Bloomington

**Copyright** © 2022 Sakai and Kurosawa. 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 Hiroyuki D. Sakai, shiroyuki@soka.ac.jp.

The authors declare no conflict of interest.

**Received** 11 November 2021

**Accepted** 27 February 2022

**Published** 9 March 2022

**TABLE 1** Commands used in bioinformatics analyses

Method	Commands
Quality filtering of raw short reads (fastp v.0.20.1)	fastp -i raw-short-read1.fastq -l raw-short-read2.fastq -o quality-filtered-short-read1.fastq -O quality-filtered-short-read2.fastq
Quality filtering of raw long reads (Filtilong v.0.2.0)	filtilong -1 quality-filtered-short-read1.fastq -2 quality-filtered-short-read2.fastq --min_length 1000 --trim --split 100 --mean_q_weight 10 long-read.fastq > quality-filtered-long-read.fastq
Hybrid assembly (Unicycler v.0.4.8)	unicycler -1 quality-filtered-short-read1.fastq -2 quality-filtered-short-read2.fastq -l quality-filtered-long-read.fastq -o assembly-files

which utilizes a variety of sugars as a carbon source (1). Based on the full-length 16S rRNA gene, the closest species of HS-3<sup>T</sup> is *Saccharolobus shibatae* with sequence similarity of 96.4%, which was calculated using the EzBioCloud 16S-based identifier (ID) (8). Average nucleotide identity (ANI) between HS-3<sup>T</sup> and *Saccharolobus shibatae* DSM5389<sup>T</sup> was 74.7%, which was calculated using the EzBioCloud ANI calculator (9).

In this report, we determined the complete genome sequence of the strain HS-3<sup>T</sup> of *Saccharolobus caldissimus*. The data deposited in the DDBJ/ENA/GenBank database contribute to our understanding of the genomic diversity and evolution of members of the order *Sulfolobales*.

**Data availability.** The genome sequence of strain HS-3<sup>T</sup> and raw reads have been deposited in DDBJ/ENA/GenBank under the accession numbers [AP025226](#), [DRR326939](#), and [DRR326940](#).

## REFERENCES

- Sakai HD, Kurosawa N. 2018. *Saccharolobus caldissimus* gen. nov., sp. nov., a facultatively anaerobic iron-reducing hyperthermophilic archaeon isolated from an acidic terrestrial hot spring, and reclassification of *Sulfolobus solfataricus* as *Saccharolobus solfataricus* comb. nov. *Int J Syst Evol Microbiol* 68:1271–1278. <https://doi.org/10.1099/ijsem.0.002665>.
- Kurosawa N, Itoh YH, Iwai T, Sugai A, Uda I, Kimura N, Horiuchi T, Itoh T. 1998. *Sulfurisphaera ohwakuensis* gen. nov., sp. nov., a novel extremely thermophilic acidophile of the order *Sulfolobales*. *Int J Syst Bacteriol* 48:451–456. <https://doi.org/10.1099/00207713-48-2-451>.
- Nakamura K, Kurosawa N, Sakai HD. 2021. Complete genome sequence of a novel thermoacidophilic and facultatively anaerobic crenarchaeon, *Stygiolobus* sp. *Microbiol Resour Announc* 10:e00581-21. <https://doi.org/10.1128/MRA.00581-21>.
- 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>.
- 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>.
- Tanizawa Y, Fujisawa T, Nakamura Y. 2018. DFAST: a flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics* 34:1037–1039. <https://doi.org/10.1093/bioinformatics/btx713>.
- Berg IA, Kockelkorn D, Buckel W, Fuchs G. 2007. A 3-hydroxypropionate/4-hydroxybutyrate autotrophic carbon dioxide assimilation pathway in Archaea. *Science* 318:1782–1787. <https://doi.org/10.1126/science.1149976>.
- Yoon SH, Ha SM, Kwon S, Lim J, Kim Y, Seo H, Chun J. 2017. Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int J Syst Evol Microbiol* 67:1613–1617. <https://doi.org/10.1099/ijsem.0.001755>.
- Yoon SH, Ha S, 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>.