A

s extreme thermophiles belonging to the family Thermococ-
caceae, species in the genus Thermococcus generally grow at temperatures between 75°C and 88°C (13). They require elemental sulfur or protons as an electron acceptor and release hydrogen, hydrogen sulfide, or carbon dioxide as an end product. Most of the species have been isolated from marine hydrothermal systems, whereas only two species have been isolated from freshwater hot springs (6, 12). Thermococcus zilligii, reported in 1982 (16), is the first of the two.

AN1 (DSM 2770), the type strain of T. zilligii, was isolated from a terrestrial geothermal spring in Kuirau Park, Rotorua, New Zealand, an environment rich in microbial biomass (9, 19). It is an obligately anaerobic and chemoorganotrophic sulfur-reducing microorganism and absolutely requires a low concentration of sodium or lithium ion for growth. A number of enzymes from AN1 have been investigated to understand the physiology and the evolution of thermophilic archaea (2, 4, 17, 18, 22). Also, the strain has been regarded as an attractive reservoir of enzymes, such as thermostable proteases and hemicellulases, for biotechnological applications (15, 21). In spite of scientific and industrial importance, the genome of this strain has not been analyzed.

The genome sequence of AN1 was determined using a massive parallel sequencing technology. Approximately 6.45 Gb of reads (~3,398-fold genome coverage) were produced from a 400-bp paired-end library using Illumina/Solexa HiSeq 2000 (NICEM, Republic of Korea). The paired reads were de novo assembled with the CLC Genomics Workbench (CLC bio, Inc.). Twenty-four contigs were orderly rearranged in five scaffolds by SSPACE (1). In silico gap closing was performed by Image (20) and Perl scripts developed in-house. All tRNA and rRNA genes were predicted by tRNAscan-SE and RNAmmer, respectively. A total of 1,958 protein-coding sequences (91.84% of the coding density) were predicted by Glimmer and Gene Marks. All genes were annotated using the information from GenBank, Pfam, TIGRFam, KEGG, and RAST. Clustered, regularly interspaced short palindromic repeats (CRISPRs) were predicted by a web-based program, CRISPRFinder (7).

The final assembly consists of five contigs of 1,764,559 bp (54.48% G+C content). The genome has a single 16S-23S rRNA operon, two distantly located 5S rRNA genes, and 46 tRNA genes.

Like other Thermococcus spp., the genome contains genes encoding two sulfhydrogenase complexes, three [NiFe]-hydrogenase complexes, and sodium-dependent transporters (5, 11, 14). However, 10 CRISPR loci composed of 5 to 58 short direct repeats were detected in the AN1 genome, whereas otherwise completely sequenced genomes of Thermococcus contain fewer than six CRISPR loci. The average length of the direct repeats is 30 bases, and spacer regions between repeats range from 32 to 70 bases. Thirty-five genes encoding CRISPR-associated protein families that belong to the core and two subtypes, Mtube and RAMP module, are found near six CRISPR loci (3, 8). These CRISPR-related genes may contribute to controlling viral infection or plasmid conjugation in microbial biomass-rich aquatic habitats.

Nucleotide sequence accession number. The draft genome sequence of AN1 has been deposited in GenBank under the accession number AIFL00000000. The sequence is also available from the Genome Encyclopedia of Microbes (GEM; http://www.gem.re.kr) (10).

ACKNOWLEDGMENTS

This project was conducted in part as the Microbial Genome Sequencing and Analysis course at the University of Science and Technology in fall 2011. We are grateful to students in the class for critical comments on the manuscript and to members of GEM and the KRIBB sequencing team for technical help.

This work was supported by the National Research Foundation of Korea (2011-0017670 to J.F.K.), the Global Frontier Biosystems Design and Engineering Center (2011-0031952) of the Ministry of Education, Science and Technology, and the Development of Biohydrogen Production Technology Using Hyperthermophilic Archaean program (to S.G.K.) of the Ministry of Land, Transport, and Maritime Affairs, Republic of Korea.

Received 18 April 2012. Accepted 30 April 2012.

Address correspondence to Jihyun F. Kim, jfk1@yonsei.ac.kr, or Sung Gyun Kang, sgkang@kordi.re.kr.

B.K.K. and S.H.L. contributed equally to this study.

Copyright © 2012, American Society for Microbiology. All Rights Reserved.

doi:10.1128/JB.00655-12
REFERENCES