

Genome Sequence of the Thermophilic Cyanobacterium *Thermosynechococcus* sp. Strain NK55a

Sergey Stolyar,^{a*} Zhenfeng Liu,^{b*} Vera Thiel,^b Lynn P. Tomsho,^b Nicolas Pinel,^c William C. Nelson,^a Stephen R. Lindemann,^a Margie F. Romine,^a Shin Haruta,^d Stephan C. Schuster,^{b*} Donald A. Bryant,^b Jim K. Fredrickson^a

Pacific Northwest National Laboratory, Richland, Washington, USA^a; Department of Biochemistry and Molecular Biology, the Pennsylvania State University, University Park, Pennsylvania, USA^b; Institute for Systems Biology, Seattle, Washington, USA^c; Department of Biological Sciences, Graduate School of Science and Engineering, Tokyo Metropolitan University, Tokyo, Japan^d

* Present address: Sergey Stolyar, Institute for Systems Biology, Seattle, Washington, USA; Zhenfeng Liu, Dana and David Dornsife College of Letters, Arts and Sciences, University of Southern California, Los Angeles, California, USA; Stephan C. Schuster, School of Biological Sciences/College of Science, Nanyang Technological University, Singapore.

The genome of the unicellular cyanobacterium *Thermosynechococcus* sp. strain NK55a, isolated from the Nakabusa hot spring, Nagano Prefecture, Japan, comprises a single, circular, 2.5-Mb chromosome. The genome is predicted to contain 2,358 protein-encoding genes, including genes for all typical cyanobacterial photosynthetic and metabolic functions. No genes encoding hydrogenases or nitrogenase were identified.

Received 13 November 2013 Accepted 31 December 2013 Published 30 January 2014

Citation Stolyar S, Liu Z, Thiel V, Tomsho LP, Pinel N, Nelson WC, Lindemann SR, Romine MF, Haruta S, Schuster SC, Bryant DA, Fredrickson JK. 2014. Genome sequence of the thermophilic cyanobacterium *Thermosynechococcus* sp. strain NK55a. *Genome Announc.* 2(1):e01060-13. doi:10.1128/genomeA.01060-13.

Copyright © 2014 Stolyar et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Sergey Stolyar, sergey.stolyar@systemsbiology.org.

Cyanobacteria, oxygenic phototrophs, thrive in diverse environments—from hot, arid deserts to ice-bound, polar environments and from the surface waters of freshwater lakes or oceans to deeper waters and soils. They occur as free-living, single-celled organisms, but they also provide the foundation for complex communities, some forming symbioses with eukaryotic organisms. Cyanobacteria are found in microbial mats of alkaline hot springs on all continents. These mats represent highly organized communities, in which cyanobacteria are responsible for the majority of carbon fixation. The cyanobacterium *Thermosynechococcus* sp. strain NK55a (NBRC 108920) was isolated from a green microbial mat at the Nakabusa hot spring, Nagano Prefecture, Japan. NK55a and related strains are the major oxygenic photosynthetic organisms in mats growing at moderate temperatures of 52 to 60°C (1). The 16S rRNA gene of NK55a was almost identical (2 mismatches) in sequence to that of *Thermosynechococcus elongatus* BP-1, whose genome was previously described (2).

Purified genomic DNA was sequenced in a 454 pyrosequencer (GS FLX+, Roche) maintained in the Schuster laboratory at the Pennsylvania State University. A total of 229,966 reads averaging 282 bp were generated and then assembled with the Newbler assembler (Roche) into 19 contigs of at least 500 bp. The average read depth was ~26×. Further assembly and gap closing were managed using Sanger sequencing and the phred/phrap/consed package. The genome sequence was autoannotated using RAST (3) (<http://rast.nmpdr.org/>).

The genome of NK55a consists of a single circular chromosome of ~2,519,964 bp; we were unable to close a single, gene-internal gap that resulted from amplification and assembly problems arising from a repeat-rich region within a hemagglutinin-like gene. The genome of NK55a is slightly smaller than the genome of *T. elongatus* BP-1 (2), displaying 94% average nucleotide identity

for coding regions and containing 99 unique genes. Among them are several colocalized open reading frames comprising a sequence of about 8 kb with high homology to *Roseiflexus castenholzii*. Only 1 insertion element was detected in NK55a, while 77 are present in BP-1. Phages or genes associated with phages were not identified in either the NK55a or BP-1 genomes. The genome contains a single rRNA gene cluster.

The core metabolism, photosynthetic apparatus, and cellular processes encoded in the NK55a genome are similar to those found in other cyanobacteria (4, 5). The NK55a genome encodes a type I clustered regularly interspaced short palindromic repeat (CRISPR) immunity system, types I through III restriction-modification systems, and a single toxin-antitoxin system for defense against invasion by foreign DNA. Although hydrogenases commonly occur in cyanobacteria, they are missing in strains BP-1 and NK55a as well as in the type A and B genomes of *Synechococcus* spp., which occur in mat communities associated with hot springs in Yellowstone National Park (6). Unlike the latter strains, NK55a also lacks nitrogenase. Further analysis of this genome and comparative analysis with other genomes will provide further insights into the adaptive mechanisms and evolution of hot spring cyanobacteria.

Nucleotide sequence accession number. The genome sequence of *Thermosynechococcus* sp. NK55a is available in GenBank under accession number CP006735.

ACKNOWLEDGMENTS

This research was supported by the Genomic Science Program (GSP), Office of Biological and Environmental Research (OBER), U.S. Department of Energy (DOE), and is a contribution of the Pacific Northwest National Laboratory (PNNL) Foundational Scientific Focus Area.

REFERENCES

1. Everroad CR, Otaki H, Matsuura K, Haruta S. 2012. Diversification of bacterial community composition along a temperature gradient at a thermal spring. *Microbes Environ.* 27:374–381. <http://dx.doi.org/10.1264/jsm2.ME11350>.
2. Nakamura Y, Kaneko T, Sato S, Ikeuchi M, Katoh H, Sasamoto S, Watanabe A, Iriguchi M, Kawashima K, Kimura T, Kishida Y, Kiyokawa C, Kohara M, Matsumoto M, Matsuno A, Nakazaki N, Shimpo S, Sugimoto M, Takeuchi C, Yamada M, Tabata S. 2002. Complete genome structure of the thermophilic cyanobacterium *Thermosynechococcus elongatus* BP-1. *DNA Res.* 9:123–130. <http://dx.doi.org/10.1093/dnares/9.4.123>.
3. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. <http://dx.doi.org/10.1186/1471-2164-9-75>.
4. Beck C, Knoop H, Axmann IM, Steuer R. 2012. The diversity of cyanobacterial metabolism: genome analysis of multiple phototrophic microorganisms. *BMC Genomics* 13:56. <http://dx.doi.org/10.1186/1471-2164-13-56>.
5. Larsson J, Nylander JA, Bergman B. 2011. Genome fluctuations in cyanobacteria reflect evolutionary, developmental and adaptive traits. *BMC Evol. Biol.* 11:187. <http://dx.doi.org/10.1186/1471-2148-11-187>.
6. Bhaya D, Grossman AR, Steunou AS, Khuri N, Cohan FM, Hamamura N, Melendrez MC, Bateson MM, Ward DM, Heidelberg JF. 2007. Population level functional diversity in a microbial community revealed by comparative genomic and metagenomic analyses. *ISME J.* 1:703–713. <http://dx.doi.org/10.1038/ismej.2007.46>.