Genomics Data 10 (2016) 30–32

Contents lists available at ScienceDirect

Genomics Data

journal homepage: www.elsevier.com/locate/gdata

Lab resource

Draft genome sequence of the extremely halophilic *Halorubrum* sp. SAH-A6 isolated from rock salts of the Danakil depression, Ethiopia

Ashagrie Gibtan [a], Mingyeong Woo [a], Dokyung Oh [a], Kyounghee Park [a], Han-Seung Lee [a], Jae Hak Sohn [a], Dong-Woo Lee [b], Jung-Kue Shin [c], Sang-Jae Lee [a,*]

[a] Major in Food Biotechnology and The Research Center for Extremophiles & Marine Microbiology, Silla University, Busan 46958, Republic of Korea
[b] School of Applied Bioscience, Kyungpook National University, Daegu 41566, Republic of Korea
[c] Department of Korean Cuisine and Food Industry Research Institute, Jeonju University, Jeonju 55069, Republic of Korea

ARTICLE INFO

Article history:
Received 7 August 2016
Received in revised form 20 August 2016
Accepted 22 August 2016
Available online 30 August 2016

Keywords:
Halorubrum sp. SAH-A6
Genome
Rock salts
Danakil depression
Ethiopia

ABSTRACT

The draft genome sequence of *Halorubrum* sp. SAH-A6, isolated from commercial rock salts of the Danakil depression, Ethiopia. The genome comprised 3,325,770 bp, with the G + C content of 68.0%. The strain has many genes which are responsible for secondary metabolites biosynthesis, transport and catabolism as compared to other *Halorubrum* archaea members. Abundant genes responsible for numerous transport systems, solute accumulation, and aromatic/sulfur decomposition were detected. The first genomic analysis encourages further research on comparative genomics, and biotechnological applications. The NCBI accession number for this genome is SAMN04278861 and ID: 4278861 and strain deposited with accession number KCTC 43215.

1. Resource table

Name of resource	*Halorubrum species* strain SAH-A6
Institution	This strain is available from Korean Collection for Type Cultures (KCTC) with the accession number KCTC 43215
Person who created resource	Ashagrie Gibtan, Mingyeong Woo, Dokyung Oh, Kyounghee Park, Han-Seung Lee, Jae Hak Sohn, Dong-Woo Lee, Jung-Kue Shin, Sang-Jae Lee*
Contact person and email	Sang-Jae Lee, sans76@silla.ac.kr
Date archived/stock date	June 2, 2016
Type of resource	Whole genome sequence data
Link to directly related literature that employed/validated this resource	http://www.ncbi.nlm.nih.gov/genome/14537?genome_assembly_id=263668
Information in public databases	http://www.ncbi.nlm.nih.gov/genome/14537?genome_assembly_id=263668

2. Resource details

The genus *Halorubrum* was originally proposed by McGenity and Grant [1] and constitutes a large group of extremely halophilic aerobic archaea belonging to the family *Halobacteriaceae.* At the time of writing, 31 species have been described in the genus *Halorubrum* [2] which are widely distributed in diverse natural and artificial hypersaline environments such as marine salterns, salt lakes, soda lakes, saline soils, salt fermented foods, and salt preserved food products [3,4]. Hence, more investigations at genomic level are required to improve our understanding of its ecology, physiology, genetics, and potentiality in biotechnological applications. *Halorubrum* sp. SAH-A6 strain was isolated from the commercial rock salt produced from the Danakil depression of Ethiopia. Currently, neither genome of this species nor *Halorubrum* genome from commercial rock salt of the Danakil depression of Ethiopia reported. To fill this gap, *Halorubrum* sp. SAH-A6 was chosen for genome sequencing.

3. Specifications

Organism	*Halorubrum* sp.
Strain	SAH-A6
Sequencer or array type	PacBio RS II
Data format	Analyzed
Experimental factors	Archaea strain

* Corresponding author.
E-mail address: sans76@silla.ac.kr (S.-J. Lee).

http://dx.doi.org/10.1016/j.gdata.2016.08.014

(*continued*)

Organism	*Halorubrum* sp.
Experimental features	Assembled and annotated whole genome
Consent	N/A
Sample source location	Rock salts of the Danakil depression, Ethiopia

The draft genome sequence of *Halorubrum* sp. SAH-A6, isolated from commercial rock salts of the Danakil depression, Ethiopia. The assembled genome comprised 3,325,770 bp, with high G + C content of 68.0% (Table 1). The strain has many genes which are responsible for secondary metabolites biosynthesis, transport, and catabolism as compared to other *Halorubrum* Archaea. In addition, strain SAH-A6 use universal strategies for extreme adaptation as indicated by the genome. Abundant genes responsible for numerous transport systems, solute accumulation, and aromatic/sulfur decomposition were detected. The subsystem category distribution statistics for *Halorubrum* sp. strain SAH-A6 were shown in Fig. 1.

The genomic analysis showed that the overall central metabolism of SAH-A6 seems to be similar to other *Halorubrum* species. All members of *Halorubrum* share the same genes that are responsible for full glycolysis/gluconeogenesis, citrate cycle, pentose phosphate, and pyruvate pathways and sugars metabolism. This shows a horizontal gene transfer within the genus. However, metabolic differences were predicted in many other pathways. Among them, for example, the number of genes which are responsible for secondary metabolites biosynthesis, transport, and catabolism are very high in SAH-A6 strain as compared to other *Halorubrum* groups.

Further genomic analysis of strain SAH-A6 showed the genetic capacity for adaptation to harsh environments. Unlike other *Halorubrum* groups, SAH-A6 has much more genes responsible for inorganic ion transport, energy conversion, amino acid transport, and metabolism which can help it to cope with the hot, saline, and nutrient limited environments. Strain SAH-A6 revealed the presence of numerous ionic regulation genes, including magnesium, and copper transport, arsenic pump-driving, ABC transporters, cobalt-zinc-cadmium resistance, and P-type ATPase. These genes help SAH-A6 and other microbes to overcome the high metallic ion in rock salt as compared to other saline environments. Apart from this, strain SAH-A6 is using genes such as stress response, heat shock proteins, DNA repair systems, maintenance of membrane fluidity, and accumulation of compatible solutes as indicated by the genome. In addition, SAH-A6 also has other unique feature for adaptation in slow growth in nutrient-poor commercial rock salt in that it possesses a single rRNA operon. However, fewer genes encoding transposase, lipid transport, and metabolism were found in the SAH-A6 genome, compared with other *Halorubrum* members.

Phylogenetic tree was built based on neighbor joining tree with the alignment of the 16S rRNA gene sequences (~1470 bp) showing the relationship between *Halorubrum* genomes available at the EzTaxon data base and SAH-A6 using MEGA6 [5] (Supplementary Fig. 1).

4. Materials and methods

The genome sequencing was performed using a single molecule real-time (SMRT) sequencing platform on the PacBio RS II (Pacific Biosciences, Menlo Park, CA) [6]. Genomic DNA was extracted using a standard genomic DNA isolation kit (Promega, USA). The whole genome sequence of strain SAH-A6 was performed using single SMRT cell with a single 180 min movie (Pacific Biosciences) with P6C4 chemistry. The open reading frames of the assembled genome were predicted and annotated using the hierarchical genome-assembly process (HGAP) [7] protocol RS HGAP Assembly 2 in SMRT analysis version 2.3.0 (Pacific Biosciences; https://github.com/PacificBiosciences/SMRT-Analysis), IMG-ER [8], NCBI COG function [9], Pfam information [10], and EzTaxon [11] database. The rRNA and tRNA genes were identified using RNAmmer 1.2 [12] and tRNA scan-SE 1.23 [13], respectively. The whole genome sequence of SAH-A6 was annotated using the Rapid Annotation System Technology (RAST) server. The pie chart showed the count of each subsystem feature and the subsystem coverage.

5. Direct link to deposited data

http://www.ncbi.nlm.nih.gov/genome/14537?genome_assembly_id=263668

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gdata.2016.08.014.

Conflict of interest

The authors have nothing to disclose.

Verification and authentication

The whole draft genomic sequence of *Halorubrum* sp. SAH-A6 (Bio project PRJNA302707) has been deposited at NCBI GenBank database under accession numbers SAMN04278861 and ID: 4278861. This strain is available from Korean Collection for Type Cultures (KCTC) with the accession number KCTC 43215.

Acknowledgements

This work was supported by the Human Resource Training Program for Regional Innovation & Creativity through the Ministry of Education (MOE) & National Research Foundation of Korea (NRF-2014H1C1A1066945) and by the Basic Science Research Program

Table 1
Comparison of the genomic feature of *Halorubrum* sp. SAH-A6 strain with various halophilic *Halorubrum* strains. The information of the reference genomes was obtained from NCBI data base.

Organism	BioProject	Resource	Genome size	Contigs	G + C (%)	r + tRNA
H. sp. SAH-A6[a]	PRJNA302707	Danakil depression, Ethiopia	3,325,770	3	68.0	6 + 45
H. lipolyticum DSM 21995	PRJNA188614	Xin-Jiang, China	3,425,042	41	68.0	3 + 44
H. aidingense JCM 13560	PRJNA188616	Xin-Jiang, China	3,108,525	37	67.2	4 + 49
H. kocurii JCM 14978	PRJNA188615	Inner Mongolia, China	3,619,738	105	66.9	1 + 46
H. lacusprofundi ATCC 49239	PRJNA58807	Deep Lake, Antarctica	3,692,576	3	64.0	9 + 51
H. saccharovorum DSM 1137	PRJNA188612	California, USA	3,423,703	72	66.9	2 + 445
H. coriense DSM 10284	PRJNA188619	Geelong, Australia	3,645,313	69	67.0	3 + 48
H. distributum JCM 10118	PRJNA188621	Turkmenistan	3,306,135	68	68.1	4 + 43

[a] This study.

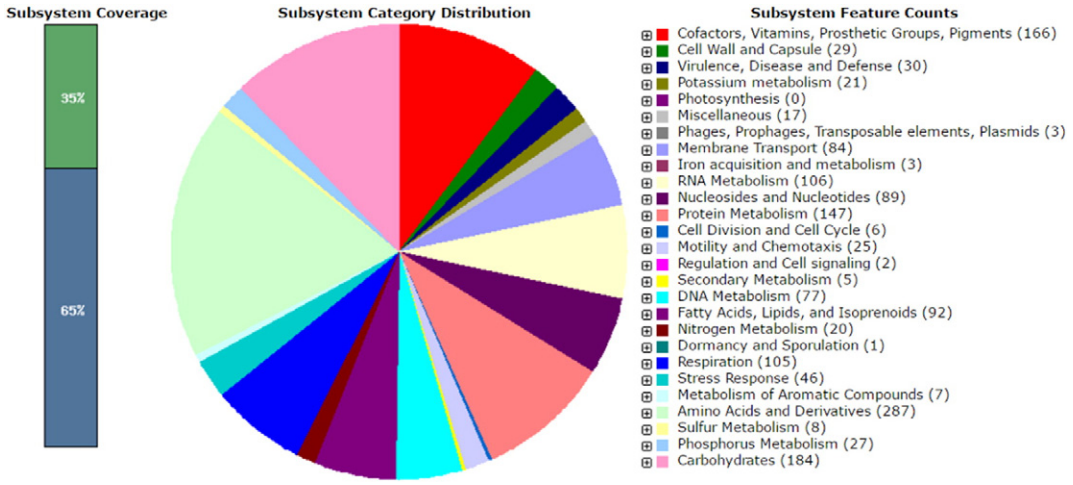

Fig. 1. The subsystem category distribution statistics for *Halorubrum* sp. strain SAH-A6. The whole genome sequence of SAH-A6 was annotated using the Rapid Annotation System Technology (RAST) server. The pie chart showed the count of each subsystem feature and the subsystem coverage.

through the National Research Foundation of Korea, which is funded by the Ministry of Education, Science and Technology (NRF-2014R1A1A1006415).

References

[1] T.J. McGenity, W.D. Grant, Transfer of *Halobacterium saccharovorum*, *Halobacterium sodomense*, *Halobacterium trapanicum* NRC 34021 and *Halobacterium lacusprofundi* to the genus *Halorubrumgen.* nov., as *Halorubrum saccharovorum* comb. nov., *Halorubrum sodomense* comb. nov., *Halorubrum trapanicum* comb. nov., and *Halorubrum lacusprofundi* comb. nov. Int. J. Syst. Bacteriol. 18 (1995) 237–243.
[2] A.C. Parte, LPSN-list of prokaryotic names with standing in nomenclature. Nucleic Acids Res. 42 (Database issue) (2016) D613–D616.
[3] M.S. Fullmer, S.M. Soucy, K.S. Swithers, A.M. Makkay, R. Wheeler, A. Ventosa, R.T. Papke, Population and genomic analysis of the genus *Halorubrum*. Front. Microbiol. 5 (140) (2014) 1–15.
[4] H.S. Lee, Diversity of halophilic archaea in fermented foods and human intestines and their application. J. Microbiol. Biotechnol. 23 (2013) 1645–1653.
[5] K. Tamura, G. Stecher, D. Peterson, A. Filipski, S. Kumar, MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol. Biol. Evol. 30 (2013) 2725–2729.
[6] J. Eid, A. Fehr, J. Gray, K. Luong, J. Lyle, G. Otto, S. Turner, Real-time DNA sequencing from single polymerase molecules. Science 323 (5910) (2009) 133–138.
[7] C.S. Chin, D.H. Alexander, P. Marks, A.A. Klammer, J. Drake, C. Heiner, J. Korlach, Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat. Methods 10 (6) (2013) 563–569.
[8] V.M. Markowitz, K. Mavromatis, N.N. Ivanova, I.M. Chen, K. Chu, N.C. Kyrpides, IMG ER: a system for microbial genome annotation expert review and curation. Bioinformatics 25 (17) (2009) 2271–2278.
[9] R.L. Tatusov, M.Y. Galperin, D.A. Natale, E.V. Koonin, The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. 28 (1) (2000) 33–36.
[10] R.D. Finn, A. Bateman, J. Clements, P. Coggill, R.Y. Eberhardt, S.R. Eddy, M. Punta, Pfam: the protein families database. Nucleic Acids Res. 42 (2014) D222–D230.
[11] O.S. Kim, Y.J. Cho, K. Lee, S.H. Yoon, M. Kim, H. Na, J. Chun, Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. Int. J. Syst. Evol. Microbiol. 62 (Pt 3) (2012) 716–721.
[12] K. Lagesen, P. Hallin, E.A. Rødland, H.H. Staerfeldt, T. Rognes, D.W. Ussery, RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. 35 (9) (2007) 3100–3108.
[13] T.M. Lowe, S.R. Eddy, tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25 (5) (1997) 955–964.