

First Insights into the Genome Sequence of the Halophilic Archaeon *Halalkalicoccus paucihalophilus* (DSM 24557)

Anja Poehlein,^a Katharina Mucek,^b Marieke Enders,^b Frederik Pankok,^b Rolf Daniel^a

Genomic and Applied Microbiology & Göttingen Genomics Laboratory, Institute of Microbiology and Genetics, Georg-August University, Göttingen, Germany^a; Members of the Applied Bioinformatics in Microbiology Course of the Microbiology and Biochemistry M.Sc./Ph.D. program, Georg-August University, Göttingen, Germany^b

***Halalkalicoccus paucihalophilus* is an extremely halophilic, Gram-negative, and nonmotile coccus-like archaeon, which was originally isolated from the Lop Nur region in the northwest of China. The genome consists of a single replicon (3.98 Mbp). *H. paucihalophilus* is able to utilize mannose, which is unique for members of this genus.**

Received 24 March 2016 Accepted 29 March 2016 Published 19 May 2016

Citation Poehlein A, Mucek K, Enders M, Pankok F, Daniel R. 2016. First insights into the genome sequence of the halophilic archaeon *Halalkalicoccus paucihalophilus* (DSM 24557). *Genome Announc* 4(3):e00382-16 doi:10.1128/genomeA.00382-16.

Copyright © 2016 Poehlein et al. 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 Rolf Daniel, rdaniel@gwdg.de.

The strictly aerobic and halophilic archaeon *Halalkalicoccus paucihalophilus* was isolated from the Lop Nur region in Xinjiang Province in Northwest China (1). It is a Gram-negative and nonmotile coccus. *H. paucihalophilus* represents a new species within the genus *Halalkalicoccus*, which exhibited 99% 16S rRNA gene sequence identity to its closest relative *Halalkalicoccus tibetensis* (1). *H. paucihalophilus* (DSM 24557) was obtained from the German Collection of Cell Cultures (DSMZ; Braunschweig, Germany). The MasterPure complete DNA purification kit (Epicentre, Madison, WI, USA) was used for DNA isolation, which was followed by generation of Illumina shotgun sequencing libraries, according to the protocol of the supplier (Illumina, San Diego, CA, USA). A MiSeq reagent kit version 3 and a MiSeq instrument were used for paired-end sequencing, as recommended by the manufacturer (Illumina). Trimming of the recovered reads was performed with Trimmomatic version 0.3.2 (2) and yielded 4,625,882 paired-end reads. SPAdes version 3.6.2 (3) was employed for assembly, which resulted in 35 contigs (>500 bp), with an average coverage of 208.58×. With Qualimap version 2, the assembly was validated and the read coverage determined (4). The draft genome of *H. paucihalophilus* contained a single chromosome, with a G+C content of 61.67%. Automatic annotation, gene prediction, and identification of rRNA and tRNA genes were carried out with Prokka (5). The analyses revealed 4,163 predicted protein-coding genes, of which 2,508 had a predicted function and 1,655 were hypothetical proteins. The genome harbored 54 tRNA genes and one rRNA operon. In contrast to other members of the genus, it is reported that *H. paucihalophilus* is able to utilize mannose and galactose (1). The genome contains a putative *gmd-1* gene that encodes GDP-mannose 4, 6-dehydratase. In some eukaryotes, this enzyme converts GDP-mannose to GDP-4-dehydro-6-deoxy-D-mannose and H₂O (6). *H. paucihalophilus* is able to use galactose as a carbon source (1). Correspondingly, a gene coding for β-galactosidase and one coding for D-gluconate-D-galactonate dehydratase (*gad3*) were present in the genome. The *gad3* gene is part of the galactose catabolism via the Entner-Doudoroff pathway (7). In addition, several

putative genes for glycotransferases, such as *aglD*, *mshA*, *aglI*, *gtf1*, and *aglG*, were identified. Among these, an archaea-specific glycotransferase-encoding gene (*aglI*) involved in the synthesis of glycolipids was detected (8). Furthermore, a tRNA (tRNA-*selC*) that is required for the incorporation of selenocysteine into proteins was identified. Additionally, the genes *selA* and *selD*, which encode enzymes that are also necessary for the incorporation of selenocysteine, were present in the genome (9). However, *selB*, which plays an essential role in selenocysteine incorporation, was not identified in the genome sequence. *H. paucihalophilus* might have lost this ability, as proteins that typically contain selenocysteine in archaea (10) were either not present or were replaced by cysteine-containing derivatives in this organism.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. [LTAZ00000000](https://www.ncbi.nlm.nih.gov/nuclink/LTAZ00000000). The version described in this paper is LTAZ01000000.

ACKNOWLEDGMENTS

This work was supported by the Nds. Ministerium für Wissenschaft und Kultur (MWK) of Lower Saxony (Germany).

We thank Kathleen Gollnow for technical support.

FUNDING INFORMATION

This work, including the efforts of Anja Poehlein, Katharina Mucek, Marieke Enders, Frederik Pankok, and Rolf Daniel, was funded by Nds. Ministerium für Wissenschaft und Kultur (MWK), Germany.

The funders (MWK) had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

REFERENCES

- Liu BB, Tang SK, Zhang YG, Lu XH, Li L, Cheng J, Zhang YM, Zhang LL, Li WJ. 2013. *Halalkalicoccus paucihalophilus* sp. nov., a halophilic archaeon from Lop nur region in Xinjiang, northwest of China. *Antonie van Leeuwenhoek* 103:1007–1014. <http://dx.doi.org/10.1007/s10482-013-9880-x>.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer

- for Illumina sequence data. *Bioinformatics* 30:2114–2120. <http://dx.doi.org/10.1093/bioinformatics/btu170>.
3. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <http://dx.doi.org/10.1089/cmb.2012.0021>.
 4. García-Alcalde F, Okonechnikov K, Carbonell J, Cruz LM, Götz S, Tarazona S, Dopazo J, Meyer TF, Conesa A. 2012. Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics* 28: 2678–2679. <http://dx.doi.org/10.1093/bioinformatics/bts503>.
 5. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <http://dx.doi.org/10.1093/bioinformatics/btu153>.
 6. Bonin CP, Freshour G, Hahn MG, Vanzin GF, Reiter WD. 2003. The GMD1 and GMD2 genes of *Arabidopsis* encode isoforms of GDP-D-mannose 4,6-dehydratase with cell type-specific expression patterns. *Plant Physiol* 132:883–892. <http://dx.doi.org/10.1104/pp.103.022368>.
 7. Kim S, Lee SB. 2005. Identification and characterization of *Sulfolobus solfataricus* D-gluconate dehydratase: a key enzyme in the non-phosphorylated Entner-Doudoroff pathway. *Biochem J* 387:271–280. <http://dx.doi.org/10.1042/BJ20041053>.
 8. Kaminski L, Abu-Qarn M, Guan Z, Naparstek S, Ventura VV, Raetz CR, Hitchen PG, Dell A, Eichler J. 2010. AglJ adds the first sugar of the N-linked pentasaccharide decorating the *Haloferax volcanii* S-layer glycoprotein. *J Bacteriol* 192:5572–5579. <http://dx.doi.org/10.1128/JB.00705-10>.
 9. Leinfelder W, Forchhammer K, Zinoni F, Sawers G, Mandrand-Berthelot MA, Böck A. 1988. *Escherichia coli* genes whose products are involved in selenium metabolism. *J Bacteriol* 170:540–546. <http://dx.doi.org/10.1002/jtra.1033>.
 10. Rother M, Resch A, Wilting R, Böck A. 2001. Selenoprotein synthesis in archaea. *Biofactors* 14:75–83. <http://dx.doi.org/10.1002/biof.5520140111>.