

Complete genome sequence of *Desulfurococcus mucosus* type strain (O7/1^T)

Reinhard Wirth¹, Olga Chertkov^{2,3}, Brittany Held^{2,3}, Alla Lapidus², Matt Nolan², Susan Lucas², Nancy Hammon², Shweta Deshpande², Jan-Fang Cheng², Roxane Tapia^{2,3}, Cliff Han^{2,3}, Lynne Goodwin^{2,3}, Sam Pitluck², Konstantinos Liolios², Pagani Ioanna², Natalia Ivanova², Konstantinos Mavromatis², Natalia Mikhailova², Amrita Pati², Amy Chen⁴, Krishna Palaniappan⁴, Miriam Land^{2,5}, Loren Hauser^{2,5}, Yun-Juan Chang^{2,5}, Cynthia D. Jeffries^{2,5}, Yvonne Bilek¹, Thomas Hader¹, Manfred Rohde⁶, Stefan Spring⁷, Johannes Sikorski⁷, Markus Göker⁷, Tanja Woyke², James Bristow², Jonathan A. Eisen^{2,8}, Victor Markowitz⁴, Philip Hugenholtz^{2,9}, Nikos C. Kyrpides², and Hans-Peter Klenk^{7*}

¹ University of Regensburg, Archaeozentrum, Regensburg, Germany

² DOE Joint Genome Institute, Walnut Creek, California, USA

³ Los Alamos National Laboratory, Bioscience Division, Los Alamos, New Mexico, USA

⁴ Biological Data Management and Technology Center, Lawrence Berkeley National Laboratory, Berkeley, California, USA

⁵ Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA

⁶ HZI – Helmholtz Centre for Infection Research, Braunschweig, Germany

⁷ DSMZ - German Collection of Microorganisms and Cell Cultures GmbH, Braunschweig, Germany

⁸ University of California Davis Genome Center, Davis, California, USA

⁹ Australian Centre for Ecogenomics, School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, Australia

*Corresponding author: Hans-Peter Klenk

Keywords: hyperthermophile, anaerobic, organotroph, sulfur respiration, spheroid-shaped, non-motile, extracellular enzymes, *Desulfurococcaceae*, GEBA

Desulfurococcus mucosus Zillig and Stetter 1983 is the type species of the genus *Desulfurococcus*, which belongs to the crenarchaeal family *Desulfurococcaceae*. The species is of interest because of its position in the tree of life, its ability for sulfur respiration, and several biotechnologically relevant thermostable and thermoactive extracellular enzymes. This is the third completed genome sequence of a member of the genus *Desulfurococcus* and already the 8th sequence from a member the family *Desulfurococcaceae*. The 1,314,639 bp long genome with its 1,371 protein-coding and 50 RNA genes is a part of the *Genomic Encyclopedia of Bacteria and Archaea* project.

Introduction

Strain O7/1^T (= DSM 2162 = ATCC 35584 = JCM 9187) is the type strain of the species *Desulfurococcus mucosus* [1], which is the type species of its genus *Desulfurococcus*. The genus currently consists of five species with a validly published name [2]. For the genus name the Neo-Latin 'desulfo-' meaning 'desulfurating', is used to characterize the dissimilatory sulfate-reducing feature of this spheroid-shaped 'coccus' [2]. The species epithet is derived from the Latin word 'mucosus' (slimy) [2]. Strain O7/1^T was isolated from an acidic hot

spring in Askja, Iceland and the name of the species was effectively published by Zillig *et al.* in 1982 [1]; valid publication of the name followed in 1983 [3]. The strain was an early target for phylogenetic studies of the domain *Archaea* (at that time termed '*Archaeobacteria*') via DNA-rRNA cross-hybridizations [4,5], as well as studies on the archaeal DNA-dependent RNA polymerase structure [6] and *Archaea*-specific quinones [7]. Subsequently, strain O7/1^T was used for studies on thermostable extracellular enzymes such as

proteinase [8] and pullulanase [9]. Here we present a summary classification and a set of features for *D. mucosus* strain O7/1^T, together with a description of the complete genome sequencing and annotation.

Classification and features

The single genomic 16S rRNA sequence of strain O7/1^T was compared using NCBI BLAST under default settings (e.g., considering only the high-scoring segment pairs (HSPs) from the best 250 hits) with the most recent release of the GreenGenes database [10] and the relative frequencies, weighted by BLAST scores, of taxa and keywords (reduced to their stem [11]) were determined. The five most frequent genera were *Sulfolobus* (27.8%), *Aeropyrum* (11.3%), *Desulfurococcus* (11.3%), *Ignicoccus* (6.5%) and *Vulcanisaeta* (6.2%) (100 hits in total). Regarding the five hits to sequences from other members of the genus, the average identity within HSPs was 96.7%, whereas the average cov-

erage by HSPs was 97.4%. Among all other species, the one yielding the highest score was *Desulfurococcus mobilis*, which corresponded to an identity of 100.0% and an HSP coverage of 100.0%. The highest-scoring environmental sequence was AB462558 ('Microbial production and energy source hyperthermophilic prokaryotes geothermal hot spring pool clone DDP-A01'), which showed an identity of 95.8% and a HSP coverage of 98.2%. The five most frequent keywords within the labels of environmental samples which yielded hits were 'spring' (9.2%), 'microbi' (6.8%), 'hot' (6.2%), 'nation/park/yellowston' (5.4%) and 'popul' (4.8%) (150 hits in total), indicating a good fit to the original habitat of *D. mucosus*. Environmental samples which yielded hits of a higher score than the highest scoring species were not found.

Figure 1 shows the phylogenetic neighborhood of *D. mucosus* in a 16S rRNA based tree. A 16S rRNA reference sequence for *D. mucosus* has not been previously published.

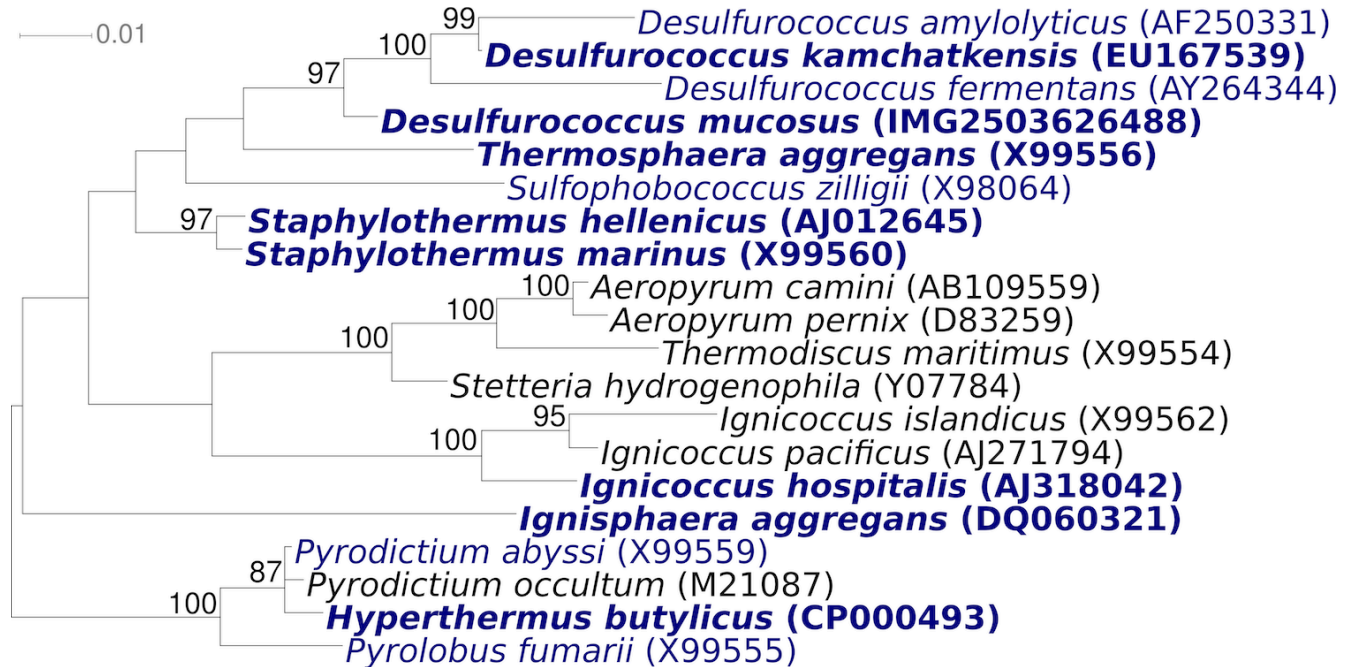


Figure 1. Phylogenetic tree highlighting the position of *D. mucosus* relative to the other type strains within the family *Desulfurococcaceae*. The tree was inferred from 1,334 aligned characters [12,13] of the 16S rRNA gene sequence under the maximum likelihood criterion [14] and rooted in accordance with the current taxonomy. The branches are scaled in terms of the expected number of substitutions per site. Numbers above branches are support values from 1,000 bootstrap replicates [15] if larger than 60%. Lineages with type strain genome sequencing projects registered in GOLD [16] are shown in blue, *Staphylothermus hellenicus* CP002051 and published genomes in bold [17-22].

The non-motile cells of strain 07/1^T are spheroid with diameters of 0.3 to 2.0 μm [1] (Figure 2), sometimes up to 10 μm [23], surrounded by a slimy mucoid layer, which covers the envelope and consists of neutral sugars and a small fraction of amino sugars [24] (Figure 2). In growing cultures, cells of strain 07/1^T were often found in pairs [2] (Table 1). Cells of strain 07/1^T can be differentiated from those of *D. mobilis*, the closest relative of *D. mucosus*, which are mobile by monopolar polytrichous flagella and devoid of the mucous polymer surrounding the *D. mucosus* cells [1,23]. Strain 07/1^T can utilize yeast extract and casein or its tryptic digests, but not casamino acids as the sole carbon source, by sulfur respiration with the production of H₂S and CO₂, or by fermentation [1]. Growing cultures synthesize a strong smelling uncharacterized product [1]. Cultures require little or no NaCl in growth media [1,23]. The temperature range for growth of strain 07/1^T is 76 to 93°C, with an optimum at 85°C [1,23]. At the optimal growth temperature, the generation time of strain 07/1^T was about four hours [1]. The pH range is 4.5 to 7.0, with an optimum at 6.0 [1,23]. Sugars, starch, glycogen, alcohols and intermediary metabolites are also not utilized [1]. Strain 07/1^T lacks an intron in the 23S RNA gene, which has been described for its close relative *D. mobilis* [35].

Chemotaxonomy

According to Zillig *et al.* 1982 [1], the cell envelope of the strain 07/1^T is flexible and probably composed of two layers of which at least the outer one appears to consist of subunits perpendicular to the surface [1]. Scarce information is available regarding the lipid composition of *D. mucosus*. The lipids in the strain 07/1^T are composed of phytanol and C40 polyisoprenoid dialcohols [1]. The polar lipid profile of the closely related *D. mobilis* has been studied and the structure of its three complex lipids has been elucidated in detail [36].

Genome sequencing and annotation

Genome project history

This organism was selected for sequencing on the basis of its phylogenetic position [37], and is part of the *Genomic Encyclopedia of Bacteria and Archaea* project [38]. The genome project is deposited in the Genomes On Line Database [16] and the complete genome sequence is deposited in GenBank. Sequencing, finishing and annotation were performed by the DOE Joint Genome Institute (JGI). A sum-

mary of the project information is shown in Table 2.

Growth conditions and DNA isolation

D. mucosus strain 07/1^T, DSM 2162, was grown anaerobically in DSMZ medium 184 (*Desulfurococcus* medium) [39] at 85°C. DNA was isolated from 0.5-1 g of cell paste using Qiagen Genomic 500 DNA kit (Qiagen 10262) following the standard protocol as recommended by the manufacturer, with no modification. DNA is available through the DNA Bank Network [40].

Genome sequencing and assembly

The genome was sequenced using a combination of Illumina and 454 sequencing platforms. All general aspects of library construction and sequencing can be found at the JGI website [41]. Pyrosequencing reads were assembled using the Newbler assembler version 2.5-internal-10Apr08-1-threads (Roche). The initial Newbler assembly consisting of three contigs in one scaffold was converted into a phrap assembly [42] by making fake reads from the consensus, to collect the read pairs in the 454 paired end library. Illumina GAii sequencing data (99.5 Mb) were assembled with Velvet [43] and the consensus sequences were shredded into 1.5 kb overlapped fake reads and assembled together with the 454 data. The 454 draft assembly was based on 546.5 Mb 454 draft data and all of the 454 paired end data. Newbler parameters are -consed -a 50 -l 350 -g -m -ml 20. The Phred/Phrap/Consed software package [42] was used for sequence assembly and quality assessment in the subsequent finishing process. After the shotgun stage, reads were assembled with parallel phrap (High Performance Software, LLC). Possible mis-assemblies were corrected with gapResolution [41], Dupfinisher [44], or sequencing cloned bridging PCR fragments with subcloning or transposon bombing (Epicentre Biotechnologies, Madison, WI). Gaps between contigs were closed by editing in Consed, by PCR and by Bubble PCR primer walks (J.-F.Chang, unpublished). A total of 12 additional reactions were necessary to close gaps and to raise the quality of the finished sequence. Illumina reads were also used to correct potential base errors and increase consensus quality using a software Polisher developed at JGI [45]. The error rate of the completed genome sequence is less than 1 in 100,000. Together, the combination of the Illumina and 454 sequencing platforms provided 120.5 × coverage of the genome. The final assembly contained 264,988 pyrosequence and 1,310,055 Illumina reads.

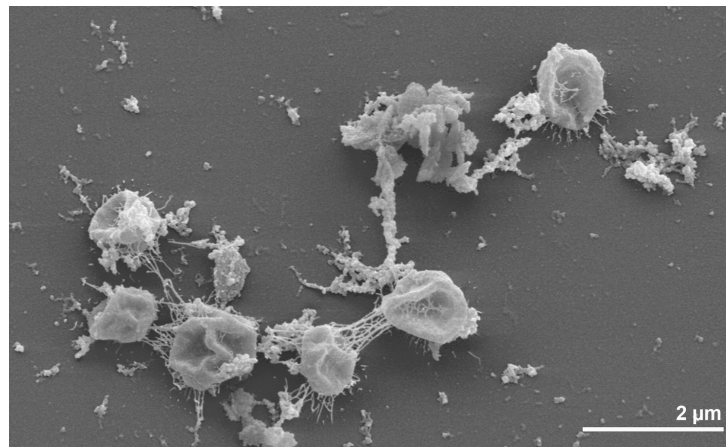


Figure 2. Scanning electron micrograph of *D. mucosus* strain O7/1^T

Table 1. Classification and general features of *D. mucosus* O7/1^T according to the MIGS recommendations [25].

MIGS ID	Property	Term	Evidence code
		Domain <i>Archaea</i>	TAS [26]
		Phylum <i>Crenarchaeota</i>	TAS [27,28]
		Class <i>Thermoprotei</i>	TAS [27,29]
	Current classification	Order <i>Desulfurococcales</i>	TAS [27,30]
		Family <i>Desulfurococcaceae</i>	TAS [2,3,31]
		Genus <i>Desulfurococcus</i>	TAS [1,3,32]
		Species <i>Desulfurococcus mucosus</i>	TAS [1,3]
		Type strain O7/1	TAS [1]
	Gram stain	negative	TAS [1]
	Cell shape	spheroid, often in pairs	TAS [1]
	Motility	non-motile	TAS [1]
	Sporulation	none	NAS
	Temperature range	76°C-93°C	TAS [23]
	Optimum temperature	85°C	TAS [1,23]
	Salinity	around 0	TAS [23]
MIGS-22	Oxygen requirement	strictly anaerobic	TAS [1]
	Carbon source	yeast extract, casein or its tryptic digest	TAS [1]
	Energy metabolism	organotroph	TAS [1]
MIGS-6	Habitat	fresh water, sulfur spring	TAS [1]
MIGS-15	Biotic relationship	free living	TAS [1]
MIGS-14	Pathogenicity	none	NAS
	Biosafety level	1	TAS [33]
	Isolation	acidic hot spring	TAS [1]
MIGS-4	Geographic location	Askja, Iceland	TAS [1]
MIGS-5	Sample collection time	1981 or before	TAS [1]
MIGS-4.1	Latitude	65.05	NAS
MIGS-4.2	Longitude	-16.8	NAS
MIGS-4.3	Depth	not reported	NAS
MIGS-4.4	Altitude	approx. 1,053 m	NAS

Evidence codes - IDA: Inferred from Direct Assay (first time in publication); TAS: Traceable Author Statement (i.e., a direct report exists in the literature); NAS: Non-traceable Author Statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence). These evidence codes are from of the Gene Ontology project [34]. If the evidence code is IDA, then the property was directly observed by one of the authors or an expert mentioned in the acknowledgements.

Table 2. Genome sequencing project information

MIGS ID	Property	Term
MIGS-31	Finishing quality	Finished
MIGS-28	Libraries used	Three genomic libraries: one 454 pyrosequence standard library, one 454 PE library (13 kb insert size), one Illumina library
MIGS-29	Sequencing platforms	Illumina GAii, 454 GS FLX Titanium
MIGS-31.2	Sequencing coverage	75.7 × Illumina; 44.8 × pyrosequence
MIGS-30	Assemblers	Newbler version 2.5-internal-10Apr08-1-threads, Velvet, phrap
MIGS-32	Gene calling method	Prodigal 1.4, GenePRIMP
	INSDC ID	CP002363
	Genbank Date of Release	January 20, 2011
	GOLD ID	Gc02914
	NCBI project ID	48641
	Database: IMG-GEBA	2503538025
MIGS-13	Source material identifier	DSM 2162
	Project relevance	Tree of Life, GEBA

Genome annotation

Genes were identified using Prodigal [46] as part of the Oak Ridge National Laboratory genome annotation pipeline, followed by a round of manual curation using the JGI GenePRIMP pipeline [47]. The predicted CDSs were translated and used to search the National Center for Biotechnology Information (NCBI) nonredundant database, UniProt, TIGR-Fam, Pfam, PRIAM, KEGG, COG, and InterPro databases. Additional gene prediction analysis and functional annotation were performed within the Integrated Microbial Genomes - Expert Review (IMG-ER) platform [48].

Genome properties

The genome consists of a 1,314,639 bp long chromosome with a G+C content of 53.1% (Figure 3 and Table 3). Of the 1,421 genes predicted, 1,371 were protein-coding genes, and 50 RNAs; 26 pseudogenes were also identified. The majority of the protein-coding genes (65.5%) were assigned with a putative function while the remaining ones were annotated as hypothetical proteins. The distribution of genes into COGs functional categories is presented in Table 4.

Table 3. Genome Statistics

Attribute	Value	% of Total
Genome size (bp)	1,314,639	100.00%
DNA coding region (bp)	1,186,810	90.28%
DNA G+C content (bp)	698,621	53.14%
Number of replicons	1	
Extrachromosomal elements	0	
Total genes	1,421	100.00%
RNA genes	50	3.52%
rRNA operons	1	
Protein-coding genes	1,371	96.48%
Pseudo genes	26	1.83%
Genes with function prediction	931	65.52%
Genes in paralog clusters	103	7.25%
Genes assigned to COGs	1,001	70.44%
Genes assigned Pfam domains	1,010	71.08%
Genes with signal peptides	146	10.27%
Genes with transmembrane helices	296	20.83%
CRISPR repeats	3	

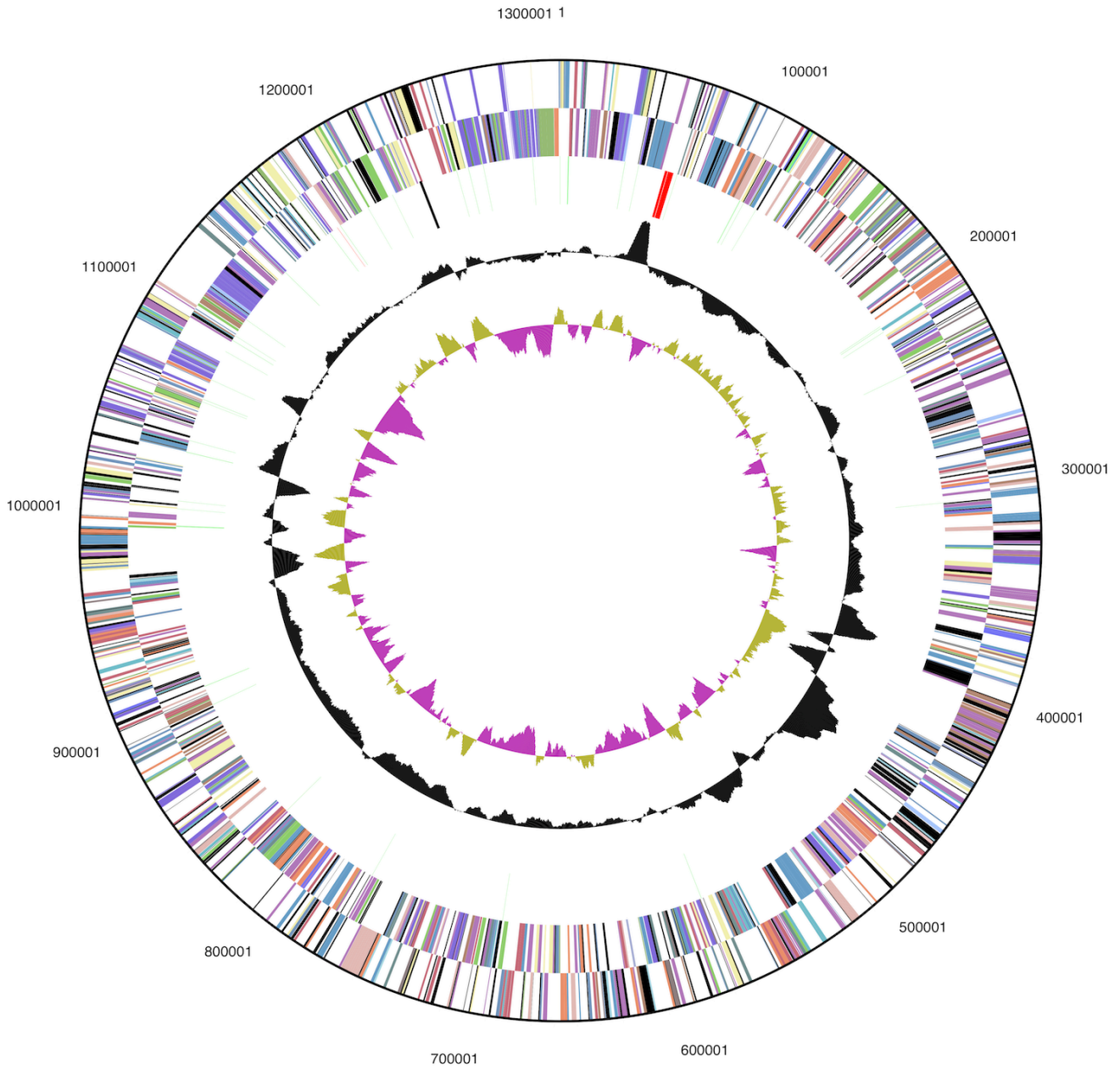


Figure 3. Graphical circular map of genome. From outside to the center: Genes on forward strand (color by COG categories), Genes on reverse strand (color by COG categories), RNA genes (tRNAs green, rRNAs red, other RNAs black), GC content, GC skew.

Table 4. Number of genes associated with the general COG functional categories

Code	value	%age	Description
J	148	13.9	Translation, ribosomal structure and biogenesis
A	2	0.2	RNA processing and modification
K	50	4.7	Transcription
L	62	5.8	Replication, recombination and repair
B	1	0.1	Chromatin structure and dynamics
D	7	0.7	Cell cycle control, cell division, chromosome partitioning
Y	0	0.0	Nuclear structure
V	10	0.9	Defense mechanisms
T	14	1.3	Signal transduction mechanisms
M	37	3.5	Cell wall/membrane/envelope biogenesis
N	4	0.4	Cell motility
Z	0	0.0	Cytoskeleton
W	0	0.0	Extracellular structures
U	10	0.9	Intracellular trafficking, secretion, and vesicular transport
O	45	4.2	Posttranslational modification, protein turnover, chaperones
C	97	9.1	Energy production and conversion
G	52	4.9	Carbohydrate transport and metabolism
E	77	7.2	Amino acid transport and metabolism
F	39	3.7	Nucleotide transport and metabolism
H	45	4.2	Coenzyme transport and metabolism
I	14	1.3	Lipid transport and metabolism
P	81	7.6	Inorganic ion transport and metabolism
Q	3	0.3	Secondary metabolites biosynthesis, transport and catabolism
R	170	16.0	General function prediction only
S	96	9.0	Function unknown
-	420	29.6	Not in COGs

Acknowledgements

We would like to gratefully acknowledge the help of Olivier D. Ngatchou-Djao (HZI) in preparing the manuscript. This work was performed under the auspices of the US Department of Energy Office of Science, Biological and Environmental Research Program, and by the University of California, Lawrence Berkeley National Laboratory under contract No. DE-AC02-05CH11231,

Lawrence Livermore National Laboratory under Contract No. DE-AC52-07NA27344, and Los Alamos National Laboratory under contract No. DE-AC02-06NA25396, UT-Battelle and Oak Ridge National Laboratory under contract DE-AC05-000R22725, as well as German Research Foundation (DFG) INST 599/1-2.

References

- Zillig W, Stetter KO, Prangishvilli D, Schäfer W, Wunderl S, Janekovic D, Holz I, Palm P. *Desulfurococcaceae*, the second family of the extremely thermophilic, anaerobic, sulfur-respiring Thermoproteales. *Zentralbl Bakteriol* 1982; **3**:304-317.
- Garrity G. NamesforLife. BrowserTool takes expertise out of the database and puts it right in the browser. *Microbiol Today* 2010; **37**:9.
- List Editor. Validation List no. 10. Validation of the publication of new names and new combinations previously effectively published outside the

- IJSB. *Int J Syst Bacteriol* 1983; **33**:438-440. [doi:10.1099/00207713-33-2-438](https://doi.org/10.1099/00207713-33-2-438)
4. Tu JK, Prangishvilli D, Huber H, Wildgruber G, Zillig W, Stetter KO. Taxonomic relations between *Archaeobacteria* including 6 novel genera examined by cross hybridization of DNAs and 16S rRNAs. *Mol Evol* 1982; **18**:109-114. [PubMed](https://pubmed.ncbi.nlm.nih.gov/2311111/) [doi:10.1007/BF01810829](https://doi.org/10.1007/BF01810829)
 5. Klenk HP, Haas B, Schwass V, Zillig W. Hybridization homology: a new parameter for the analysis of phylogenetic relations, demonstrated with the urkingdom of the *Archaeobacteria*. *J Mol Evol* 1986; **24**:167-173. [doi:10.1007/BF02099964](https://doi.org/10.1007/BF02099964)
 6. Prangishvilli D, Zillig W, Gierl A, Biesert L, Holz I. DNA-dependent RNA polymerase of thermoacidophilic archaeobacteria. *Eur J Biochem* 1982; **122**:471-477. [PubMed](https://pubmed.ncbi.nlm.nih.gov/1011111/) [doi:10.1111/j.1432-1033.1982.tb06461.x](https://doi.org/10.1111/j.1432-1033.1982.tb06461.x)
 7. Thurl S, Witke W, Buhrow I, Schäfer W. Quinones from archaeobacteria, II. Different types of quinones from sulphur-dependent archaeobacteria. *Biol Chem Hoppe Seyler* 1986; **367**:191-197. [PubMed](https://pubmed.ncbi.nlm.nih.gov/101515/) [doi:10.1515/bchm3.1986.367.1.191](https://doi.org/10.1515/bchm3.1986.367.1.191)
 8. Cowan DA, Smolenski KA, Daniel RM, Morgan HW. An extracellular proteinase from a strain of the archaeobacterium *Desulfurococcus* growing at 88 degrees C. *Biochem J* 1987; **247**:121-133. [PubMed](https://pubmed.ncbi.nlm.nih.gov/101128/)
 9. Duffner F, Bertoldo C, Andersen JT, Wagner K, Antranikian G. A new thermoactive pullulanase from *Desulfurococcus mucosus*: cloning, sequencing, purification, and characterization of the recombinant enzyme after expression in *Bacillus subtilis*. *J Bacteriol* 2000; **182**:6331-6338. [PubMed](https://pubmed.ncbi.nlm.nih.gov/101128/) [doi:10.1128/JB.182.22.6331-6338.2000](https://doi.org/10.1128/JB.182.22.6331-6338.2000)
 10. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie E, Keller K, Huber T, Dalevi D, Hu P, Andersen G. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* 2006; **72**:5069-5072. [PubMed](https://pubmed.ncbi.nlm.nih.gov/101128/) [doi:10.1128/AEM.03006-05](https://doi.org/10.1128/AEM.03006-05)
 11. Porter MF. An algorithm for suffix stripping. *Program: electronic library and information systems* 1980; **14**:130-137. [doi:10.1108/eb046814](https://doi.org/10.1108/eb046814)
 12. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 2000; **17**:540-552. [PubMed](https://pubmed.ncbi.nlm.nih.gov/101128/)
 13. Lee C, Grasso C, Sharlow MF. Multiple sequence alignment using partial order graphs. *Bioinformatics* 2002; **18**:452-464. [PubMed](https://pubmed.ncbi.nlm.nih.gov/101093/) [doi:10.1093/bioinformatics/18.3.452](https://doi.org/10.1093/bioinformatics/18.3.452)
 14. Stamatakis A, Hoover P, Rougemont J. A rapid bootstrap algorithm for the RAxML Web servers. *Syst Biol* 2008; **57**:758-771. [PubMed](https://pubmed.ncbi.nlm.nih.gov/101080/) [doi:10.1080/10635150802429642](https://doi.org/10.1080/10635150802429642)
 15. Pattengale ND, Alipour M, Bininda-Emonds ORP, Moret BME, Stamatakis A. How many bootstrap replicates are necessary? *Lect Notes Comput Sci* 2009; **5541**:184-200. [doi:10.1007/978-3-642-02008-7_13](https://doi.org/10.1007/978-3-642-02008-7_13)
 16. Liolios K, Chen IM, Mavromatis K, Tavernarakis N, Hugenholtz P, Markowitz VM, Kyrpides NC. The Genomes On Line Database (GOLD) in 2009: status of genomic and metagenomic projects and their associated metadata. *Nucleic Acids Res* 2009; **38**:D346-D354. [PubMed](https://pubmed.ncbi.nlm.nih.gov/101093/) [doi:10.1093/nar/gkp848](https://doi.org/10.1093/nar/gkp848)
 17. Ravin NV, Mardanov AV, Beletsky AV, Kublanov IV, Kolganova TV, Lebedinsky AV, Chernyh NA, Bonch-Osmolovskaya EA, Skryabin KG. Complete genome sequence of the anaerobic, protein-degrading hyperthermophilic crenarchaeon *Desulfurococcus kamchatkensis*. *J Bacteriol* 2009; **191**:2371-2379. [PubMed](https://pubmed.ncbi.nlm.nih.gov/101128/) [doi:10.1128/JB.01525-08](https://doi.org/10.1128/JB.01525-08)
 18. Anderson IJ, Sun H, Lapidus A, Copeland A, Glavina Del Rio T, Tice H, Dalin E, Lucas S, Barry K, Land M, et al. Complete genome sequence of *Staphylothermus marinus* Stetter and Fiala 1986 type strain F1. *Stand Genomic Sci* 2009; **1**:183-188. [PubMed](https://pubmed.ncbi.nlm.nih.gov/104056/) [doi:10.4056/sigs.30527](https://doi.org/10.4056/sigs.30527)
 19. Brügger K, Chen L, Stark M, Zibat A, Redder P, Ruepp A, Awayez M, She Q, Garrett RA, Klenk HP. The genome of *Hyperthermus butylicus*: a sulfur-reducing, peptide fermenting, neutrophilic crenarchaeote growing up to 108 °C. *Archaea* 2007; **2**:127-135. [PubMed](https://pubmed.ncbi.nlm.nih.gov/101155/) [doi:10.1155/2007/745987](https://doi.org/10.1155/2007/745987)
 20. Göker M, Held B, Lapidus A, Nolan M, Spring S, Yasawong M, Lucas S, Glavina Del Rio T, Tice H, Cheng JF, et al. Complete genome sequence of *Ignisphaera aggregans* type strain (AQ1.S1¹). *Stand Genomic Sci* 2010; **3**:66-75. [PubMed](https://pubmed.ncbi.nlm.nih.gov/104056/) [doi:10.4056/sigs.1072907](https://doi.org/10.4056/sigs.1072907)
 21. Podar M, Anderson I, Makarova KS, Elkins JG, Ivanova N, Wall MA, Lykidis A, Mavromatis K, Sun H, Hudson ME, et al. A genomic analysis of the archaeal system *Ignicoccus hospitalis*-*Nanoarchaeum equitans*. *Genome Biol* 2008; **9**:R158. [PubMed](https://pubmed.ncbi.nlm.nih.gov/101186/) [doi:10.1186/gb-2008-9-11-r158](https://doi.org/10.1186/gb-2008-9-11-r158)

22. Spring S, Rachel R, Lapidus A, Davenport K, Tice H, Copeland A, Cheng JF, Lucas S, Chen F, Nolan M, et al. Complete genome sequence of *Thermosphaera aggregans* type strain (M11TL^T). *Stand Genomic Sci* 2010; **2**:245-259. [PubMed](#) [doi:10.4056/sigs.821804](https://doi.org/10.4056/sigs.821804)
23. Huber H, Stetter KO. 2006. *Desulfurococcales* In: M Dworkin, S Falkow, E Rosenberg, KH Schleifer E Stackebrandt (eds), *The Prokaryotes*, 3. ed, vol. 3. Springer, New York, p. 52-68.
24. Stetter KO, Zillig W. 1985. *Thermoplasma* and the thermophilic sulfur-dependent archaebacteria. In: Woese CR, and Wolfe RS (eds) *The Bacteria*. Academic Press. New York, NY. **8**:100-201.
25. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, Thomson N, Allen MJ, Angiuoli SV, et al. The minimum information about a genome sequence (MIGS) specification. *Nat Biotechnol* 2008; **26**:541-547. [PubMed](#) [doi:10.1038/nbt1360](https://doi.org/10.1038/nbt1360)
26. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains *Archaea*, *Bacteria*, and *Eucarya*. *Proc Natl Acad Sci USA* 1990; **87**:4576-4579. [PubMed](#) [doi:10.1073/pnas.87.12.4576](https://doi.org/10.1073/pnas.87.12.4576)
27. Validation list 85: Validation of publication of new names and new combinations previously effectively published outside the IJSEM. *Int J Syst Evol Microbiol* 2002; **52**:685-690. [PubMed](#) [doi:10.1099/ijs.0.02358-0](https://doi.org/10.1099/ijs.0.02358-0)
28. Garrity GM, Holt JG. 2001. Phylum AI. *Crenarchaeota* phy. nov. In: Garrity GM, Boone DR, Castenholz RW (eds), *Bergey's Manual of Systematic Bacteriology*, Second Edition, Volume 1, Springer, New York, p. 169-210.
29. Reysenbach AL. 2001. Class I. *Thermoprotei* class. nov. In: Garrity GM, Boone DR, Castenholz RW (eds), *Bergey's Manual of Systematic Bacteriology*, Second Edition, Volume 1, Springer, New York, p.169.
30. Huber H, Stetter O. 2001. Order II. *Desulfurococcales* ord. nov. In: Garrity GM, Boone DR, Castenholz RW (eds), *Bergey's Manual of Systematic Bacteriology*, Second Edition, Volume 1, Springer, New York, p. 179-180.
31. Burggraf S, Huber H, Stetter KO. Reclassification of the crenarchaeal orders and families in accordance with 16S rRNA sequence data. *Int J Syst Bacteriol* 1997; **47**:657-660. [PubMed](#) [doi:10.1099/00207713-47-3-657](https://doi.org/10.1099/00207713-47-3-657)
32. Perevalova AA, Svetlichny VA, Kublanov IV, Chernyh NA, Kostrikina NA, Tourova TP, Kuznetsov BB, Bonch-Osmolovskaya EA. *Desulfurococcus fermentans* sp. nov., a novel hyperthermophilic archaeon from a Kamchatka hot spring, and emended description of the genus *Desulfurococcus*. *Int J Syst Evol Microbiol* 2005; **55**:995-999. [PubMed](#) [doi:10.1099/ijs.0.63378-0](https://doi.org/10.1099/ijs.0.63378-0)
33. Classification of *Bacteria* and *Archaea* in risk groups. [www.baua.de TRBA 466](http://www.baua.de/TRBA/466).
34. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. Gene Ontology: tool for the unification of biology. *Nat Genet* 2000; **25**:25-29. [PubMed](#) [doi:10.1038/75556](https://doi.org/10.1038/75556)
35. Kjems J, Garrett RA. Novel splicing mechanism for the ribosomal RNA intron in the archaebacterium *Desulfurococcus mobilis*. *Cell* 1988; **54**:693-703. [PubMed](#) [doi:10.1016/S0092-8674\(88\)80014-X](https://doi.org/10.1016/S0092-8674(88)80014-X)
36. Lanzotti V, De Rosa M, Trincone A, Basso AL, Gambacorta A, Zillig W. Complex lipids from *Desulfurococcus mobilis*, a sulfur-reducing archaebacterium. *Biochim Biophys Acta* 1987; **922**:95-102.
37. Klenk HP, Göker M. En route to a genome-based classification of *Archaea* and *Bacteria*? *Syst Appl Microbiol* 2010; **33**:175-182. [PubMed](#) [doi:10.1016/j.syapm.2010.03.003](https://doi.org/10.1016/j.syapm.2010.03.003)
38. Wu D, Hugenholz P, Mavromatis K, Pukall R, Dalin E, Ivanova NN, Kunin V, Goodwin L, Wu M, Tindall BJ, et al. A phylogeny-driven genomic encyclopaedia of *Bacteria* and *Archaea*. *Nature* 2009; **462**:1056-1060. [PubMed](#) [doi:10.1038/nature08656](https://doi.org/10.1038/nature08656)
39. List of growth media used at DSMZ: http://www.dsmz.de/microorganisms/media_list.php.
40. Gemeinholzer B, Dröge G, Zetsche H, Haszprunar G, Klenk HP, Güntsch A, Berendsohn WG, Wägele JW. The DNA Bank Network: the start from a German initiative. *Biopreservation and Biobanking* 2011; **9**:51-55. [doi:10.1089/bio.2010.0029](https://doi.org/10.1089/bio.2010.0029)
41. DOE Joint Genome Institute. <http://www.jgi.doe.gov>
42. Phrap and Phred for Windows, MacOS, Linux, and Unix. <http://www.phrap.com>
43. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs.

- Genome Res* 2008; **18**:821-829. [PubMed doi:10.1101/gr.074492.107](#)
44. Han C, Chain P. 2006. Finishing repeat regions automatically with Dupfinisher. *In*: Proceeding of the 2006 international conference on bioinformatics & computational biology. Edited by Hamid R. Arabnia & Homayoun Valafar, CSREA Press. June 26-29, 2006: 141-146.
45. Lapidus A, LaButti K, Foster B, Lowry S, Trong S, Goltsman E. POLISHER: An effective tool for using ultra short reads in microbial genome assembly and finishing. AGBT, Marco Island, FL, 2008.
46. Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 2010; **11**:119. [PubMed doi:10.1186/1471-2105-11-119](#)
47. Pati A, Ivanova NN, Mikhailova N, Ovchinnikova G, Hooper SD, Lykidis A, Kyrpides NC. Gene-PRIMP: a gene prediction improvement pipeline for prokaryotic genomes. *Nat Methods* 2010; **7**:455-457. [PubMed doi:10.1038/nmeth.1457](#)
48. Markowitz VM, Ivanova NN, Chen IMA, Chu K, Kyrpides NC. IMG ER: a system for microbial genome annotation expert review and curation. *Bioinformatics* 2009; **25**:2271-2278. [PubMed doi:10.1093/bioinformatics/btp393](#)