Complete genome sequence of *Desulfarculus baarsii* type strain (2st14[T])

Hui Sun[1], Stefan Spring[2], Alla Lapidus[1], Karen Davenport[1,3], Tijana Glavina Del Rio[1], Hope Tice[1], Matt Nolan[1], Alex Copeland[1], Jan-Fang Cheng[1], Susan Lucas[1], Roxanne Tapia[1,3], Lynne Goodwin[1,3], Sam Pitluck[1], Natalia Ivanova[1], Ionna Pagani[1], Konstantinos Mavromatis[1], Galina Ovchinnikova[1], Amrita Pati[1], Amy Chen[1], Krishna Palaniappan[4], Loren Hauser[1,5], Yun-Juan Chang[1,5], Cynthia D. Jeffries[1,5], John C. Detter[1,3], Cliff Han[1,3], Manfred Rohde[6], Evelyne Brambilla[2], Markus Göker[2], Tanja Woyke[1], Jim Bristow[1], Jonathan A. Eisen[1,7], Victor Markowitz[4], Philip Hugenholtz[1], Nikos C Kyrpides[1], Hans-Peter Klenk[2*] and Miriam Land[1,5]

[1] DOE Joint Genome Institute, Walnut Creek, California, USA
[2] DSMZ - German Collection of Microorganisms and Cell Cultures GmbH, Braunschweig, Germany
[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] University of California Davis Genome Center, Davis, California, USA

*Corresponding author: Hans-Peter Klenk

Desulfarculus baarsii (Widdel 1981) Kuever *et al.* 2006 is the type and only species of the genus *Desulfarculus*, which represents the family *Desulfarculaceae* and the order *Desulfarculales*. This species is a mesophilic sulfate-reducing bacterium with the capability to oxidize acetate and fatty acids of up to 18 carbon atoms completely to CO_2. The acetyl-CoA/CODH (Wood-Ljungdahl) pathway is used by this species for the complete oxidation of carbon sources and autotrophic growth on formate. The type strain 2st14[T] was isolated from a ditch sediment collected near the University of Konstanz, Germany. This is the first completed genome sequence of a member of the order *Desulfarculales*. The 3,655,731 bp long single replicon genome with its 3,303 protein-coding and 52 RNA genes is a part of the *Genomic Encyclopedia of Bacteria and Archaea* project.

Introduction

Most sulfate reducing bacteria, available in pure culture, oxidize organic electron donors incompletely to acetate, whereas species that oxidize acetate and other carbon compounds completely to CO_2, using sulfate as an electron acceptor, are less frequently isolated. Sulfate reducers with the latter type of metabolism are of special interest, because it is assumed that they are dominant in anoxic marine sediments [1]. Sulfate reducing prokaryotes with the ability to mineralize organic compounds to CO_2 are phylogenetically dispersed and can be found within the *Proteobacteria*, *Firmicutes* and *Euryarchaeota*. At the time of writing, representa-

tives of this type of metabolism, for which a completely sequenced genome exists include *Desulfobacterium autotrophicum* [2], *Desulfotomaculum acetoxidans* [3] and *Archaeoglobus fulgidus* [4]. In the present work, the complete genome sequence of *Desulfarculus baarsii* a completely oxidizing sulfate reducing bacterium representing the order *Desulfarculales* within the *Deltaproteobacteria,* was determined. The original description of *D. baarsii* was based on strain 1st1 (= "Göttingen") [5], which was probably subsequently lost and replaced by the designated type strain 2st14[T] (= "Konstanz") [6]. Strain 2st14[T] (= DSM 2075 = ATCC 33931 =

LMG 7858) was enriched from anoxic mud from a ditch near the University of Konstanz, Germany, in a medium supplemented with stearate and sulfate and subsequently isolated in an anaerobic agar dilution series with formate plus sulfate [7,8]. *D. baarsii* strain 2st14^T is the first member of the family *Desulfarculaceae* within the order *Desulfarculales* with a sequenced genome. The presented sequence data will enable interesting genome comparisons with other sulfate reducing bacteria of the class *Deltaproteobacteria*.

Classification and features

The species *D. baarsii* represents a separate lineage within the *Deltaproteobacteria* which is only distantly related to most other members of this class. The closest relatives based on 16S rRNA gene sequence similarity values are the type strains of *Desulfomonile tiedjei* (87.6% sequence identity) and *Desulfomonile liminaris* (87.2%), both belonging to the family *Syntrophaceae* within the order *Syntrophobacterales* [9]. The most similar cloned 16S rRNA gene, EUB-42 [10] shared only 95.5% sequence similarity with *D. baarsii* and was retrieved from anaerobic sludge. Strain 2st14^T represents the only strain of this species available from a culture collection, thus far. Currently available data from cultivation independent studies (environmental screening and genomic surveys) did not surpass 86% sequence similarity, indicating that members of this species are restricted to distinct habitats which are currently undersampled in most environments or are in low abundance (status October 2010). The single genomic 16S rRNA sequence of strain 2st14^T was compared using BLAST with the most resent release of the Greengenes database [11] and the relative frequencies of taxa and keywords, weighted by BLAST scores, were determined. The five most frequent genera were *Desulfovibrio* (43.3%), *Syntrophobacter* (14.4%), *Desulfomonile* (11.8%), *Desulfarculus* (9.6%) and *Desulfatibacillum* (7.5%). The species yielding the highest score was *D. baarsii*. The five most frequent keywords within the labels of environmental samples which yielded hits were 'sediment' (4.5%), 'microbial' (4.5%), 'lake' (1.7%), 'depth' (1.7%) and 'sea' (1.6%). 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. baarsii* 2st14^T in a 16S rRNA based tree. The sequence of the single 16S rRNA gene in the genome differs by one nucleotide from the previously published 16S rRNA gene sequence generated from DSM 2075 (AF418174) which contains five ambiguous base calls. Genbank entry M34403 from 1989 is also annotated as 16S rRNA sequence of strain 2st14^T, but differs in 45 positions (3.2%) from the actual sequence. This difference probably reflects more the progress in sequencing technology than biological differences.
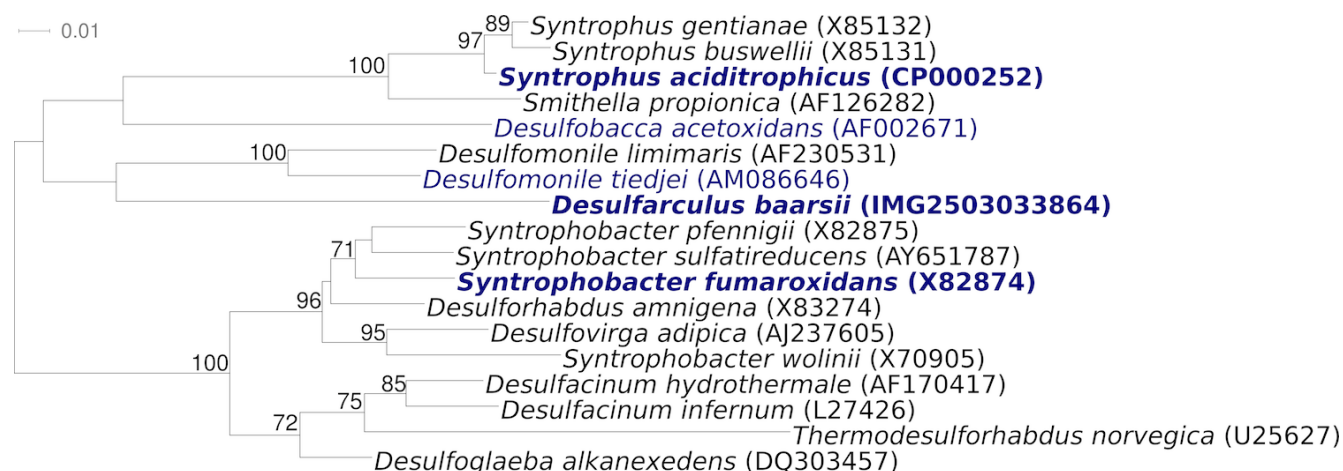

Figure 1. Phylogenetic tree highlighting the position of *D. baarsii* relative to the other type strains of related genera within the class *Deltaproteobacteria*. The tree was inferred from 1,465 aligned characters [12,13] of the 16S rRNA gene sequence under the maximum likelihood criterion [14] and 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, published genomes [17] and INSDC accession CP000478 for *Syntrophobacter fumaroxidans* in bold.

Table 1. Classification and general features of *D. baarsii* strain 2st14 in according to the MIGS recommendations [18].

MIGS ID	Property	Term	Evidence code
	Current classification	Domain *Bacteria*	TAS [19]
		Phylum *Proteobacteria*	TAS [20]
		Class *Deltaproteobacteria*	TAS [21,22]
		Order *Desulfarculales*	TAS [21,23]
		Family *Desulfarculaceae*	TAS [7,21,23, 24]
		Genus *Desulfarculus*	TAS [7,21]
		Species *Desulfarculus baarsii*	TAS [6,7,21]
		Type strain 2st14	TAS [6]
	Gram stain	negative	TAS [5]
	Cell shape	vibrio-shaped	TAS [5]
	Motility	motile (single polar flagellum)	TAS [5]
	Sporulation	non-sporulating	TAS [5]
	Temperature range	20-39°C	TAS [5]
	Optimum temperature	35°C	TAS [5]
	Salinity	optimum growth at 7–20 g/l NaCl	TAS [5,7]
MIGS-22	Oxygen requirement	strictly anaerobic	TAS [5]
	Carbon source	CO_2, formate, acetate, propionate, butyrate, higher fatty acids	TAS [5]
	Energy source	formate, acetate, propionate, butyrate, higher fatty acids	TAS [5]
MIGS-6	Habitat	anoxic freshwater or brackish sediments	TAS [5]
MIGS-15	Biotic relationship	free living	NAS
MIGS-14	Pathogenicity	none	TAS [25]
	Biosafety level	1	TAS [25]
	Isolation	mud from a ditch	TAS [7]
MIGS-4	Geographic location	Konstanz, Germany	TAS [7]
MIGS-5	Sample collection time	1981 or before	NAS
MIGS-4.1	Latitude	47.7	NAS
MIGS-4.2	Longitude	9.2	
MIGS-4.3	Depth	not reported	
MIGS-4.4	Altitude	about 406 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 [26]. If the evidence code is IDA, then the property was directly observed by one of the authors or an expert mentioned in the acknowledgements.

The cells of *D. baarsii* 2st14[T] are vibrioid, Gram-negative and 0.5-0.7 by 1.5–4 μm in size (Figure 2, Table 1). Motility is conferred by a single polar flagellum (not visible in Figure 2) [5]. The temperature range for growth is 20-39°C with an optimum around 35°C. The pH range for growth is 6.5–8.2, with an optimum at 7.3. The strain grows optimally in the presence of 7–20 g/l NaCl and 1.2–3g/l $MgCl_2 \times 6\ H_2O$, but growth is nearly as rapid at lower concentrations [7]. *D. baarsii* strain 2st14[T] is a strictly anaerobic, non-fermentative, chemoorganotrophic sulfate-reducer that oxidizes organic substrates completely to CO_2. Sulfate, sulfite and thiosulfate serve as terminal electron acceptors and are reduced to H_2S, but sulfur, fumarate and nitrate cannot be utilized. The following compounds are utilized as electron donors and carbon sources: formate, acetate, propionate, butyrate, iso-butyrate, 2-methylbutyrate, valerate, iso-valerate, and higher fatty acids up to 18 carbon atoms. Growth on formate does not require an additional organic carbon source [5,7]. A high activity of carbon monoxide dehydrogenase is observed in *D. baarsii*, indicating the operation of the anaerobic C_1-pathway (Wood-Ljungdahl pathway) for formate assimilation and CO_2 fixation or complete oxidation of acetyl-CoA [27].

The oxygen detoxification system of *D. baarsii* was analyzed in some detail. It could be shown that a genomic region encoding a putative rubredoxin oxidoreductase (*rbo*) and rubredoxin (*rub*) of *D. baarsii* is able to suppress deleterious effects of reactive oxygen species (ROS) in *Escherichia coli* mutants lacking superoxide dismutase [28]. The cloned genes were identified in the whole genome sequence as Deba_2049 (*rub*) and Deba_2050 (*rbo*) and found in close proximity to a gene encoding rubrerythrin (Deba_2051), which is supposed to play an important role in the oxygen tolerance of anaerobic bacteria [29]. The product of the recombinant *rbo* gene of *D. baarsii* was later further characterized and designated as desulfoferrodoxin (Dfx), because no evidence for a rubredoxin oxidoreductase could be demonstrated. Instead, a function as superoxide reductase was proposed [30].

Chemotaxonomy

The cellular fatty acid pattern of *D. baarsii* strain $2st14^T$ is dominated by saturated straight chain fatty acids (43.0% $C_{14:0}$, 9.9% $C_{16:0}$, and 2.3% $C_{18:0}$), followed by saturated iso- and anteiso-branched fatty acids (21.3% i-$C_{14:0}$, 12.3% ai-$C_{15:0}$, and 2.8% i-$C_{15:0}$). Occurrence of the dimethylacetal (DMA) derivates $C_{15:0\ DMA}$ (1.8%) and i-$C_{15:0\ DMA}$ (0.6%) represents a distinctive trait of this strain, because these compounds are rarely found in *Desulfovibrio* species [31]. A comparison of the fatty acid profiles of *D. baarsii* and various Gram-negative sulfate-reducers by cluster analysis indicated a separate position of *D. baarsii* [31], corroborating the distinct phylogenetic position of the species as shown based on the 16S rRNA sequence analysis (Figure 1). Unfortunately, besides the cellular fatty acid composition no further chemotaxonomic data are available for this species.
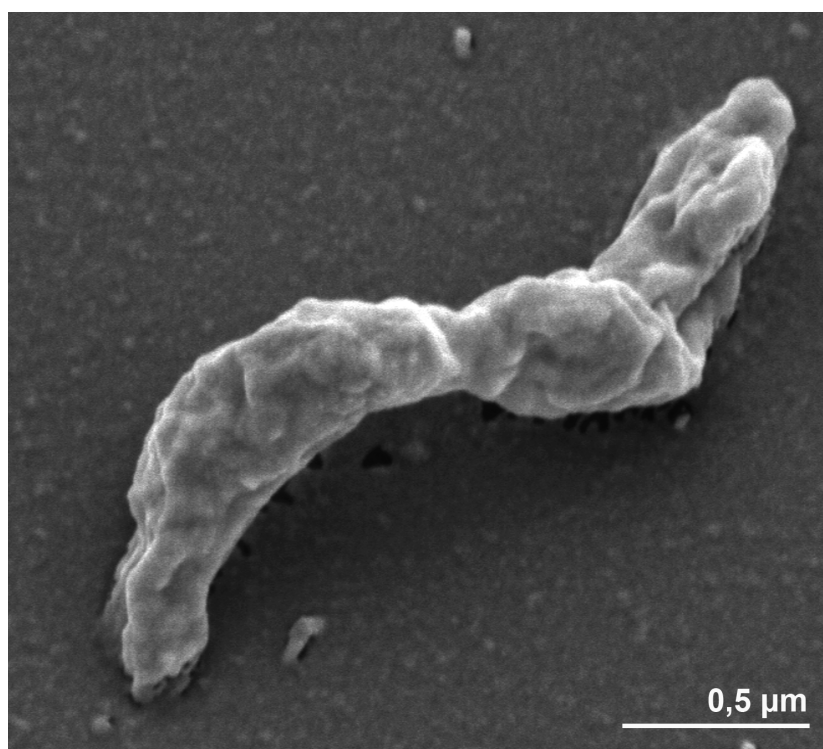

Figure 2. Scanning electron micrograph of *D. baarsii* $2st14^T$

Genome sequencing and annotation
Genome project history

This organism was selected for sequencing on the basis of its phylogenetic position [32], and is part of the ***Genomic Encyclopedia of Bacteria and Archaea*** project [33]. The genome project is deposited in the Genome OnLine 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 summary of the project information is shown in Table 2.

Table 2. Genome sequencing project information.

MIGS ID	Property	Term
MIGS-31	Finishing quality	Finished
MIGS-28	Libraries used	Two 454 pyrosequence libraries, standard and pairs end (13 kb insert size) and one Illumina standard library
MIGS-29	Sequencing platforms	454 Titanium, Illumina GAii
MIGS-31.2	Sequencing coverage	43.1 × 454 Titanium; 73.2 × Illumina
MIGS-30	Assemblers	Newbler, Velvet, phrap
MIGS-32	Gene calling method	Prodigal
	INSDC ID	CP002085
	GenBank Date of Release	August 6, 2010
	GOLD ID	Gc01335
	NCBI project ID	37955
	Database: IMG-GEBA	2502957037
MIGS-13	Source material identifier	DSM 2075
	Project relevance	GEBA

Growth conditions and DNA isolation

D. baarsii, strain 2st14T, DSM 2075, was grown anaerobically in DSMZ medium 208 (*Desulfovibrio baarsii* medium) [34] at 37°C. DNA was isolated from 0.5-1 g of cell paste using Jetflex Genomic DNA Purification Kit (GENOMED 600100) following the manufacturer's instructions, but with 30 min incubation at 58°C with an additional 10 μl proteinase K for cell lysis.

Genome sequencing and assembly

The genome of 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 [35]. Pyrosequencing reads were assembled using the Newbler assembler version 2.1-PreRelease-4-28-2009-gcc-3.4.6-threads (Roche). The initial Newbler assembly consisted of 42 contigs in two scaffolds and was converted into a phrap assembly by making fake reads from the consensus, collecting the read pairs in the 454 paired end library. Illumina GAii sequencing data (267.7Mb) were assembled with Velvet [36] 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 157.7 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 [37] was used for sequence assembly and quality assessment in the following finishing process: After the shotgun stage, reads were assembled with parallel phrap (High Performance Software, LLC). Possible misassemblies were corrected with gapResolution [35], Dupfinisher, or sequencing cloned bridging PCR fragments with subcloning or transposon bombing (Epicentre Biotechnologies, Madison, WI) [38]. Gaps between contigs were closed by editing in Consed, by

PCR and by Bubble PCR primer walks (J.-F.Chang, unpublished). A total of 344 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 [39]. 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 116.3 × coverage of the genome. Final assembly contained 431,804 pyrosequence and 7,436,430 Illumina reads.

Genome annotation

Genes were identified using Prodigal [40] as part of the Oak Ridge National Laboratory genome annotation pipeline, followed by a round of manual curation using the JGI GenePRIMP pipeline [41]. 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 was performed within the Integrated Microbial Genomes - Expert Review platform [42].

Genome properties

The genome is 3,655,731 bp long and comprises one main circular chromosome with an overall GC content of 65.7% (Table 3 and Figure 3). Of the 3,355 genes predicted, 3,303 were protein-coding genes, and 52 RNAs; 26 pseudogenes were also identified. The majority of the protein-coding genes (73.4%) were assigned a putative function while those remaining 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)	3,655,731	100.00%
DNA coding region (bp)	3,313,356	90.63%
DNA G+C content (bp)	2,401,943	65.70%
Number of replicons	1	
Extrachromosomal elements	0	
Total genes	3,355	100.00%
RNA genes	52	1.55%
rRNA operons	1	
Protein-coding genes	3,303	98.45%
Pseudo genes	26	0.77%
Genes with function prediction	2,463	73.41%
Genes in paralog clusters	481	14.34%
Genes assigned to COGs	2,466	73.50%
Genes assigned Pfam domains	2,613	77.88%
Genes with signal peptides	686	20.45%
Genes with transmembrane helices	768	22.89%
CRISPR repeats	3	
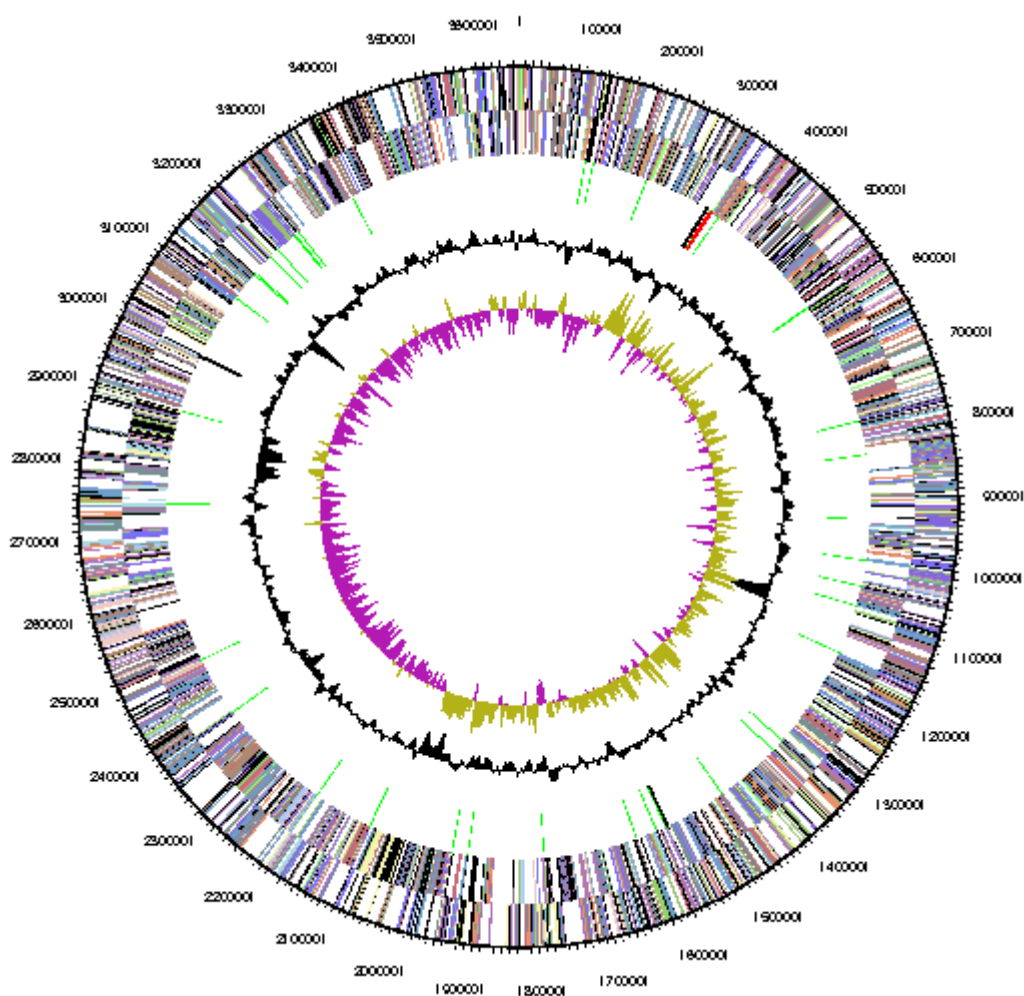
Figure 3. Graphical circular map of the genome. From outside to the center: Genes on forward strand (color by COG categories), Genes on reverse strand (color by COG categories), RNA genes (tRNAs green, rRNAs red, other RNAs black), GC content, GC skew.

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

Code	value	%age	Description
J	155	5.7	Translation, ribosomal structure and biogenesis
A	1	0.0	RNA processing and modification
K	137	5.0	Transcription
L	109	4.0	Replication, recombination and repair
B	3	0.1	Chromatin structure and dynamics
D	32	1.2	Cell cycle control, cell division, chromosome partitioning
Y	0	0.0	Nuclear structure
V	39	1.4	Defense mechanisms
T	260	9.6	Signal transduction mechanisms
M	210	7.7	Cell wall/membrane biogenesis
N	103	3.8	Cell motility
Z	0	0.0	Cytoskeleton
W	0	0.0	Extracellular structures
U	74	2.7	Intracellular trafficking and secretion, and vesicular transport
O	84	3.1	Posttranslational modification, protein turnover, chaperones
C	228	8.4	Energy production and conversion
G	89	3.3	Carbohydrate transport and metabolism
E	180	6.6	Amino acid transport and metabolism
F	62	2.3	Nucleotide transport and metabolism
H	149	5.5	Coenzyme transport and metabolism
I	130	4.9	Lipid transport and metabolism
P	124	4.6	Inorganic ion transport and metabolism
Q	56	2.1	Secondary metabolites biosynthesis, transport and catabolism
R	312	11.5	General function prediction only
S	182	6.7	Function unknown
-	889	26.5	Not in COGs

Acknowledgements

References

1. Jørgensen BB. Mineralization of organic matter in the sea bed - the role of sulphate reduction. *Nature* 1982; **296**:643-645. doi:10.1038/296643a0

2. Strittmatter AW, Liesegang H, Rabus R, Decker I, Amann J, Sönke A, Henna A, Fricke WF, Martinez-Arias R, Bartels D, *et al*. Genome sequence of *Desulfobacterium autotrophicum* HRM2, a marine sulfate reducer oxidizing organic carbon completely to carbon dioxide. *Environ Microbiol* 2009; **11**:1038-1055. PubMed doi:10.1111/j.1462-2920.2008.01825.x

3. Spring S, Lapidus A, Schröder M, Gleim D, Sims D, Meincke L, Glavina Del Rio T, Tice H, Copeland A, Cheng JF, *et al*. Complete genome sequence of *Desulfotomaculum acetoxidans* type strain (5575T). *Stand Genomic Sci* 2009; **1**:242-253. doi:10.4056/sigs.39508

4. Klenk HP, Clayton RA, Tomb JF, White O, Nelson KE, Ketchum KA, Dodson RJ, Gwinn M, Hickey EK, Peterson JD, *et al*. The complete genome sequence of the hyperthermophilic, sulphate-

reducing archaeon *Archaeoglobus fulgidus*. *Nature* 1997; **390**:364-370. [PubMed](doi:10.1038/37052) doi:10.1038/37052

5. Widdel F. Anaerober Abbau von Fettsäuren und Benzoesäure durch neu isolierte Arten Sulfat-reduzierender Bakterien. Dissertation. Georg-August-Universität zu Göttingen, Göttingen 1980.

6. Validation of the publication of new names and new combinations previously effectively published outside the IJSB. Validation List no. 7. *Int J Syst Bacteriol* 1981; **31**:382-383. doi:10.1099/00207713-31-3-382

7. Kuever J, Rainey FA, Widdel F. Genus I. *Desulfarculus* gen. nov. *In*: DJ Brenner, NR Krieg, JT Staley and GM Garrity (*eds*), Bergey's Manual of Systematic Bacteriology, second edition, vol. 2 (The *Proteobacteria*), part C (The *Alpha-*, *Beta-*, *Delta-*, and *Epsilonproteobacteria*), Springer, New York, 2005, p. 1004-1005.

8. Jansen K, Thauer RK, Widdel F, Fuchs G. Carbon assimilation pathways in sulfate reducing bacteria. Formate, carbondioxide, carbon monoxide, and acetate by *Desulfovibrio baarsii*. *Arch Microbiol* 1984; **138**:257-262. doi:10.1007/BF00402132

9. Chun J, Lee JH, Jung Y, Kim M, Kim S, Kim BK, Lim YW. EzTaxon: a web-based tool for the identification of prokaryotes based on 16S ribosomal RNA gene sequences. *Int J Syst Evol Microbiol* 2007; **57**:2259-2261. [PubMed](doi:10.1099/ijs.0.64915-0) doi:10.1099/ijs.0.64915-0

10. Xu K, Liu H, Li X, Chen J, Wang A. Typical methanogenic inhibitors can considerably alter bacterial populations and affect the intereaction between fatty acid degraders and homoacetogens. *Appl Microbiol Biotechnol* 2010; **87**:2267-2279. [PubMed](doi:10.1007/s00253-010-2708-y) doi:10.1007/s00253-010-2708-y

11. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* 2006; **72**:5069-5072. [PubMed](doi:10.1128/AEM.03006-05) doi:10.1128/AEM.03006-05

12. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 2000; **17**:540-552. [PubMed]

13. Lee C, Grasso C, Sharlow MF. Multiple sequence alignment using partial order graphs. *Bioinformatics* 2002; **18**:452-464. [PubMed](doi:10.1093/bioinformatics/18.3.452) doi: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](doi:10.1080/10635150802429642) doi: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

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](doi:10.1093/nar/gkp848) doi:10.1093/nar/gkp848

17. McInerney MJ, Rohlin L, Mouttaki H, Kim U, Krupp RS, Rios-Hernandez L, Sieber J, Struchtemeyer CG, Bhattacharyya A, Campbell JW, Gunsalus RP. The genome of *Syntrophus aciditrophus*: life at the thermodynamic limit of microbial growth. *Proc Natl Acad Sci USA* 2007; **104**:7600-7605. [PubMed](doi:10.1073/pnas.0610456104) doi:10.1073/pnas.0610456104

18. 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) doi:10.1038/nbt1360

19. 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) doi:10.1073/pnas.87.12.4576

20. Garrity GM, Holt JG. The Road Map to the Manual. In: Garrity GM, Boone DR, Castenholz RW (eds), Bergey's Manual of Systematic Bacteriology, Second Edition, Volume 1, Springer, New York, 2001, p. 119-169.

21. Validation List No. 107. List of new names and new combinations previously effectively, but not validly, published. *Int J Syst Evol Microbiol* 2006; **56**:1-6. [PubMed](doi:10.1099/ijs.0.64188-0) doi:10.1099/ijs.0.64188-0

22. Kuever J, Rainey FA, Widdel F. Class IV. *Deltaproteobacteria* class. nov. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT (eds), Bergey's Manual of Systematic Bacteriology, Second Edition, Volume 2, Part C, Springer, New York, 2005, p. 922.

23. Kuever J, Rainey FA, Widdel F. Order IV. *Desulfarcales* ord. nov. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT (eds), Bergey's Manual of Sys-

tematic Bacteriology, Second Edition, Volume 2, Part C, Springer, New York, 2005, p. 1003.

24. Kuever J, Rainey FA, Widdel F. Family I. *Desulfarculaceae* fam. nov. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT (eds), Bergey's Manual of Systematic Bacteriology, Second Edition, Volume 2, Part C, Springer, New York, 2005, p. 1003.

25. Classification of bacteria and archaea in risk groups. http://www.baua.de TRBA 466.

26. 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

27. Schauder R, Eikmanns B, Thauer RK, Widdel F, Fuchs G. Acetate oxidation to CO_2 in anaerobic bacteria via a novel pathway not involving reactions of the citric acid cycle. *Arch Microbiol* 1986; **145**:162-172. doi:10.1007/BF00446775

28. Pianzzola MJ, Soubes M, Touati D. Overproduction of the *rbo* gene product from *Desulfovibrio* species suppresses all deleterious effects of lack of superoxide dismutase in *Escherichia coli. J Bacteriol* 1996; **178**:6736-6742. PubMed

29. Lumppio HL, Shenvi NV, Summers AO, Voordouw G, Kurtz DM, Jr. Rubrerythrin and rubredoxin oxidoreductase in *Desulfovibrio vulgaris*: a novel oxidative stress protection system. *J Bacteriol* 2001; **183**:101-108. PubMed doi:10.1128/JB.183.1.101-108.2001

30. Lombard M, Fontecave M, Touati D, Niviere V. Reaction of the desulfoferrodoxin from *Desulfoarculus baarsii* with superoxide anion. Evidence for a superoxide reductase activity. *J Biol Chem* 2000; **275**:115-121. PubMed doi:10.1074/jbc.275.1.115

31. Vainshtein M, Hippe H, Kroppenstedt RM. Cellular fatty acid composition of *Desulfovibrio* species and its use in classification of sulfate-reducing bacteria. *Syst Appl Microbiol* 1992; **15**:554-566.

32. 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

33. Wu D, Hugenholtz P, Mavromatis K, Pukall R, Dalin E, Ivanova N, Kunin V, Goodwin L, Wu M, Tindall BJ, *et al*. A phylogeny-driven genomic encyclopedia of *Bacteria* and *Archaea. Nature* 2009; **462**:1056-1060. PubMed doi:10.1038/nature08656

34. List of growth media used at DSMZ: http://www.dsmz.de/microorganisms/media_list.php

35. DOE Joint Genome Institute. http://www.jgi.doe.gov

36. 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

37. Phrap and Phred for Windows, MacOS, Linux, and Unix. http://www.phrap.com

38. Sims D, Brettin T, Detter JC, Han C, Lapidus A, Copeland A, Glavina Del Rio T, Nolan M, Chen F, Lucas S, *et al*. Complete genome sequence of *Kytococcus sedentarius* type strain (541[T]). *Stand Genomic Sci* 2009; **1**:12-20. doi:10.4056/sigs.761

39. 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.

40. 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

41. Pati A, Ivanova NN, Mikhailova N, Ovchinnikova G, Hooper SD, Lykidis A, Kyrpides NC. GenePRIMP: a gene prediction improvement pipeline for prokaryotic genomes. *Nat Methods* 2010; **7**:455-457. PubMed doi:10.1038/nmeth.1457

42. 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