The complete genome sequences of sulfur-oxidizing Gammaproteobacteria *Sulfurifustis variabilis* skN76^T and *Sulfuricaulis limicola* HA5^T

Kazuhiro Umezawa, Tomohiro Watanabe, Aya Miura, Hisaya Kojima* and Manabu Fukui

### Abstract

*Sulfurifustis variabilis* and *Sulfuricaulis limicola* are autotrophic sulfur-oxidizing bacteria belonging to the family *Acidiferrobacteraceae* in the order *Acidiferrobacterales*. The type strains of these species, strain skN76^T and strain HA5^T, were isolated from lakes in Japan. Here we describe the complete genome sequences of *Sulfurifustis variabilis* skN76^T and *Sulfuricaulis limicola* HA5^T. The genome of *Sulfurifustis variabilis* skN76^T consists of one circular chromosome with size of 4.0 Mbp including 3864 protein-coding sequences. The genome of *Sulfuricaulis limicola* HA5^T is 2.9 Mbp chromosome with 2763 protein-coding sequences. In both genomes, 46 transfer RNA-coding genes and one ribosomal RNA operon were identified. In the genomes, redundancies of the genes involved in sulfur oxidation and inorganic carbon fixation pathways were observed. This is the first report to show the complete genome sequences of bacteria belonging to the order *Acidiferrobacterales* in the class *Gammaproteobacteria*.

**Keywords:** Bacteria, Gram-negative, Sulfur-oxidizing bacteria, *Acidiferrobacterales*, *Acidiferrobacteraceae*

### Introduction

*Sulfurifustis variabilis* skN76^T and *Sulfuricaulis limicola* HA5^T are gammaproteobacterial sulfur-oxidizing bacteria isolated from sediments of Lake Mizugaki and Lake Harutori, respectively [1, 2]. They both belong to the family *Acidiferrobacteraceae* within the order *Acidiferrobacterales*. In this order, only three species have been isolated in pure culture. They are all chemolithoautotrophs and can grow by oxidation of inorganic sulfur compounds. *Sulfurifustis variabilis* and *Sulfuricaulis limicola* are neutrophilic, whereas the other species, *Acidiferrobacter thiooxydans*, is acidophilic [3]. Taxonomy of *Acidiferrobacter thiooxydans* has been revised several times, and the family *Acidiferrobacteraceae* and order *Acidiferrobacterales* were recently established to accommodate the species [1, 3–5]. The members of the family *Acidiferrobacteraceae* have been frequently detected in various environments as gene sequences [2, 3, 6].

Here we show the complete genome sequences of *Sulfurifustis variabilis* skN76^T and *Sulfuricaulis limicola* HA5^T as the first genomes of the order *Acidiferrobacterales*.

### Organism information

#### Classification and features

The cells of *Sulfurifustis variabilis* skN76^T are rod-shaped or filamentous form with varying length, and 0.3–0.5 μm in width (Fig. 1a, Table 1). The cells of *Sulfuricaulis limicola* HA5^T are rod-shaped, 1.2–6.0 μm in length and 0.3–0.5 μm in width (Fig. 1b, Table 1). They are both Gram-stain-negative. *Sulfurifustis variabilis* and *Sulfuricaulis limicola* belong to the family *Acidiferrobacteraceae* within the class *Gammaproteobacteria* (Fig. 2). They both utilized thiosulfate, tetrathionate and elemental sulfur as electron donors for chemolithoautotrophic growth under aerobic conditions [1, 2].
Fig. 1 Phase-contrast micrographs of *Sulfurifustis variabilis* skN76T (a) and *Sulfuricaulis limicola* HAS² (b), grown with thiosulfate at 45 and 28 °C, respectively. Bars, 5 μm

Table 1 Classification and general features of *Sulfurifustis variabilis* skN76T and *Sulfuricaulis limicola* HAS² according to MIGS recommendations

| MIGS ID | Property          | Term     | Evidence code | Term     | Evidence code |
|---------|-------------------|----------|---------------|----------|---------------|
|         | **Sulfurifustis variabilis skN76T** |          |               | **Sulfuricaulis limicola HAS²** |          |
|         | **Domain**        | Bacteria | TAS [23]      | **Domain** | Bacteria      | TAS [23]      |
|         | **Phylum**        | Proteobacteria | TAS [24]    | **Phylum** | Proteobacteria | TAS [24]    |
|         | **Class**         | Gammaproteobacteria | TAS [25] | **Class** | Gammaproteobacteria | TAS [25] |
|         | **Order**         | Acidiferrobacterales | TAS [1]  | **Order** | Acidiferrobacterales | TAS [1]  |
|         | **Family**        | Acidiferrobacteraceae | TAS [1] | **Family** | Acidiferrobacteraceae | TAS [1] |
|         | **Genus**         | Sulfurifustis | TAS [1]   | **Genus** | Sulfuricaulis   | TAS [2]   |
|         | **Species**       | *Sulfurifustis variabilis* | TAS [1] | **Species** | *Sulfuricaulis limicola* | TAS [2] |
|         | Type strain       | skN76    |               | HAS      |               |
|         | Gram stain        | negative | TAS [1]       | negative | TAS [2]       |
|         | Cell shape        | rod or filaments | TAS [1] | rod        | TAS [2] |
|         | Motility          | motile   | TAS [1]       | not reported |               |
|         | Sporulation       | not reported |               | not reported |               |
|         | Temperature range | 28–46 °C | TAS [1]       | 8–37 °C | TAS [2] |
|         | Optimum temperature | 42–45 °C | TAS [1] | 28–32 °C | TAS [2] |
|         | pH range; Optimum | 6.3–8.9; 6.8–8.2 | TAS [1] | 6.1–9.2; unknown | TAS [2] |
|         | Carbon source     | bicarbonate | TAS [1] | bicarbonate | TAS [2] |
|         | **MIGS-6**        | Habitat  | Sediment of a lake | Sediment of a lake | TAS [2] |
|         | **MIGS-6.3**      | Salinity | <2.6 % NaCl (w/v) | <1.2 % NaCl (w/v) | TAS [2] |
|         | **MIGS-22**       | Oxygen requirement | aerobic | aerobic | TAS [2] |
|         | **MIGS-15**       | Biotic relationship | free-living | free-living | TAS [2] |
|         | **MIGS-14**       | Pathogenicity | non-pathogen | non-pathogen | NAS |
|         | **MIGS-4**        | Geographic location | Lake Mizugaki, Japan | Lake Harutori, Japan | TAS [2] |
|         | **MIGS-5**        | Sample collection | November 30, 2010 | April 26, 2012 | NAS |
|         | **MIGS-4.1**      | Latitude | 35°51.5’ N | 42°58.4’ N | NAS |
|         | **MIGS-4.2**      | Longitude | 138°30.0’ E | 144°23.9’ E | NAS |
|         | **MIGS-4.4**      | Altitude | not reported | not reported | |

* Evidence codes—IDA Inferred from Direct Assay, 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 the Gene Ontology project.
Genome sequencing information

**Genome project history**

*Sulfurifustis variabilis* skN76<sup>T</sup> and *Sulfuricaulis limicola* HA5<sup>T</sup> were selected for sequencing as representatives of sulfur-oxidizing bacteria belonging to the order *Acidiferrobacterales*, to reveal characteristics of their genomes. A summary of the project information is shown in Table 2.

**Growth conditions and genomic DNA preparation**

*Sulfurifustis variabilis* skN76<sup>T</sup> and *Sulfuricaulis limicola* HA5<sup>T</sup> were grown with 20 mM thiosulfate as an energy source in a bicarbonate-buffered medium previously described [1], at 45 and 28 °C, respectively. Genomic DNA samples were prepared by using Wizard® genomic DNA purification kit (Promega, Madison, WI, USA) from approximately 0.2 ml (skN76) or 0.1 ml (HA5) of

![Phylogenetic tree](image)

**Fig. 2** Phylogenetic tree showing the relationships of *Sulfurifustis variabilis* skN76<sup>T</sup> and *Sulfuricaulis limicola* HA5<sup>T</sup> with other members of the class *Gammaproteobacteria* based on 16S rRNA gene sequences aligned by using CLUSTAL W. *Desulfatitalea tepidiphila* S28bFT was used as an outgroup. This tree was reconstructed using 1412 sites with the neighbor-joining method by using MEGA6 [27]. Percentage values of 1000 bootstrap resamplings are shown at nodes; values below 50 % were not shown.

### Table 2 Project information

| MIGS ID | Property                      | *Sulfurifustis variabilis* skN76<sup>T</sup> Term | *Sulfuricaulis limicola* HA5<sup>T</sup> Term |
|---------|-------------------------------|-----------------------------------------------|---------------------------------------------|
| MIGS 31 | Finishing quality             | Completed                                     | Completed                                   |
| MIGS-28 | Libraries used                | 15–20 kb SMRTbell<sup>TM</sup> library         | 10–20 kb SMRTbell<sup>TM</sup> library      |
| MIGS 29 | Sequencing platforms          | PacBio RS II                                  | PacBio RS II                                |
| MIGS 31.2| Fold coverage                | 210 ×                                         | 142 ×                                       |
| MIGS 30 | Assemblers                    | RS_HGAP Assembly.2                           | RS_HGAP Assembly.3                          |
| MIGS 32 | Gene calling method           | Microbial Genome Annotation Pipeline          | Microbial Genome Annotation Pipeline        |
|         | Locus Tag                     | SVA                                           | SCL                                         |
|         | Genbank ID                    | AP014936                                      | AP014879                                    |
|         | GenBank Date of Release       | July 29, 2016                                 | July 29, 2016                               |
|         | BIOPROJECT                    | PRJDB4108                                     | PRJDB3927                                   |
| MIGS 13 | Source Material Identifier    | DSM 100313                                    | DSM 100373                                  |
|         | Project relevance             | Environmental                                 | Environmental                               |
cell pellets. Amounts of the obtained DNA assessed by spectrophotometry were ca. 270 μg (skN76) and 90 μg (HA5) respectively, and the UV absorption ratio of 260/280 nm was greater than 1.8 in both samples.

**Genome sequencing and assembly**

The genomic DNA was sheared into approximately 20 kb using g-TUBE (Covaris, Inc., Woburn, MA, USA). The SMRTbell™ templates were prepared from the fragments using SMRTbell™ Template Prep Kit 1.0 (Pacific Biosciences, Menlo Park, CA, USA). The size-selected libraries for sequencing were prepared by using BluePippin (Sage Science, Beverly, MA, USA). The libraries were sequenced on a PacBio RS II instrument (Pacific Biosciences) with P6-C4 chemistry (for *Sulfurifustis variabilis* skN76ᵀ) or P5-C3 chemistry (for *Sulfuricaulis limicola* HA5ᵀ). De novo assembly was performed by using RS_HGAP Assembly.3 (for *Sulfurifustis variabilis* skN76ᵀ) or RS_HGAP Assembly.2 (for *Sulfuricaulis limicola* HA5ᵀ), implemented within the SMRT Analysis v2.3 (Pacific Biosciences) software environment. By assembling 79,017 subreads (837,333,548 bp) of *Sulfurifustis variabilis* skN76ᵀ, two contigs with the lengths of ca. 4.0 Mbp and 5.4 kbp were obtained. The shorter one was identical to a partial sequence of the larger one, and a circular chromosome was manually constructed from the larger contig by finding self-overlapping regions using the *in silico* Molecular Cloning (R) Genomic Edition (In Silico Biology, Inc., Yokohama, Japan) application. As for *Sulfuricaulis limicola* HA5ᵀ, a single contig (ca. 2.9 Mbp) was obtained by assembling 61,565 subreads (409,124,339 bp), and circular chromosome was manually constructed in the same manner.

**Genome annotation**

The genomes were annotated automatically using the Microbial Genome Annotation Pipeline [7]. Further manual annotation of the predicted protein-coding sequences was performed on the basis of BLASTP searches against the NCBI nonredundant database. CDSs were annotated as hypothetical protein-coding genes when they met any of the following four criteria in the top hit of the BLASTP analysis: (1) E-value >1e-8, (2) length coverage <60 % against query sequence (3) sequence identity <30 % or (4) function of the hit was unidentified. The WebMGA server was used to assign the genes to Clusters of Ortholog Groups and Protein family domains [8–11]. The Phobius server was used to predict signal peptides and transmembrane helices [12]. Clustered Regularly Interspaced Short Palindromic Repeat loci were detected using CRISPRfinder [13].

**Genome properties**

The basic statistics of the genomes are shown in Table 3. Both genomes contained 46 tRNA genes and one rRNA operon. The genome size of *Sulfurifustis variabilis* skN76ᵀ was approximately 1.4 times larger than that of *Sulfuricaulis limicola* HA5ᵀ. CRISPR loci were found only in the genome of *Sulfurifustis variabilis* skN76ᵀ (Table 3). The distribution of genes into COGs functional categories is presented in Table 4.

| Table 3 | Genome statistics of *Sulfurifustis variabilis* skN76ᵀ and *Sulfuricaulis limicola* HA5ᵀ |
|---------|-----------------------------------------------------------------------------------------|
|         | *Sulfurifustis variabilis* skN76ᵀ | *Sulfuricaulis limicola* HA5ᵀ |
| Attribute | Value | % of Total | Value | % of Total |
| Genome size (bp) | 3,958,814 | 100.00 | 2,864,672 | 100.00 |
| DNA coding (bp) | 3,565,567 | 90.06 | 2,567,493 | 89.63 |
| DNA G+C (bp) | 2,670,566 | 67.46 | 1,759,557 | 61.42 |
| DNA scaffolds | 1 | 100.00 | 1 | 100.00 |
| Total genes | 3913 | 100.00 | 2812 | 100.00 |
| Protein coding genes | 3864 | 98.75 | 2763 | 98.26 |
| RNA genes | 49 | 1.25 | 49 | 1.74 |
| Pseudo genes | unknown | unknown | unknown | unknown |
| Genes in internal clusters | unknown | unknown | unknown | unknown |
| Genes with function prediction | 2930 | 75.83 | 2036 | 73.69 |
| Genes assigned to COGs | 2921 | 75.60 | 2165 | 78.36 |
| Genes with Pfam domains | 2970 | 76.86 | 2208 | 79.91 |
| Genes with signal peptides | 893 | 23.11 | 562 | 20.34 |
| Genes with transmembrane helices | 845 | 21.87 | 622 | 22.51 |
| CRISPR repeats | 6 | | 0 | |
Insights from the genome sequences
In both the genomes of *Sulfurifustis variabilis* skN76T and *Sulfuricaulis limicola* HA5T, genes involved in the sulfur oxidation pathway were identified. The genomes of both strains contain genes of the DSR system related to the oxidation of elemental sulfur to sulfite [14, 15]. They contain a *dsr* gene cluster of identical composition, *dsrABEFHCMKL-JOPNR* (SVA_1954-1967, SCL_1274-1261). There are some *dsr* genes outside of the gene cluster, *dsrAB* (SVA_0258-0259, SCL_0256-0257), *dsrS* (SVA_2921, SCL_0781) and *dsrC* (SVA_0281, SVA_0284, SVA_0917, SVA_0969, SVA_1205, SVA_1793, SVA_1949, SVA_2832, SVA_3655; SCL_0275, SCL_0524, SCL_0785, SCL_1279, SCL_1423, SCL_2646).

As genes encoding proteins involved in oxidation of sulfite to sulfate in the cytoplasm, both genomes contain two copies of the *aprAB* genes encoding an adenosine-5’-phosphosulphate reductase (SVA_0067-0066, SVA_3594-3595; SCL_0078-0077) and the *sqr* gene encoding a sulfide:quinone oxidoreductase (SVA_1781, SVA_2675, SVA3205).

*Sulfurifustis variabilis* skN76T (SCL_2523-2520), but that of *Sulfurifustis variabilis* skN76T does not. The AprM and Hdr complex are thought to have similar function that interacts with the adenosine-5’-phosphosulphate reductase [16–18]. The genomes also contain the *soeABC* genes (SVA_2734, SVA_2736-2737; SCL_0523-0521), encoding a membrane-bound polysulfide reductase-like iron-sulfur molybdenoprotein, which is suspected to be involved in sulfite oxidation in the cytoplasm [19]. Further, the genome of *Sulfurifustis variabilis* skN76T contains the *sorAB* genes (SVA_1391-1390) related to the direct oxidation of sulfite to sulfate in the periplasm [20].

For thiosulfate oxidation, both genomes contain the *soxXYZAB* gene cluster (SVA_2999-3003, SCL_2229-2333). Although sulfide oxidation by these bacteria has not been demonstrated, genes related to sulfide oxidation were identified; the *fccAB* (soxEF) genes encoding a flavocytochrome *c*–sulfide dehydrogenase (SVA_0067-0066, SVA_3594-3595; SCL_0078-0077) and the *sqr* gene encoding a sulfide:quinone oxidoreductase (SVA_1781, SVA_2675, SVA3205).

*Sulfurifustis variabilis* skN76T and *Sulfuricaulis limicola* HA5T are autotrophic bacteria. They both have two

### Table 4 Number of genes associated with general COG functional categories

| Code | *Sulfurifustis variabilis* skN76T | *Sulfuricaulis limicola* HA5T | Description |
|------|----------------------------------|------------------------------|-------------|
|      | Value %age | Value %age | Translation, ribosomal structure and biogenesis |
| J    | 164 4.24   | 159 5.75   | RNA processing and modification |
| A    | 5 0.13     | 2 0.07     | Transcription |
| K    | 191 4.94   | 130 4.71   | Replication, recombination and repair |
| L    | 154 3.99   | 117 4.23   | Chromatin structure and dynamics |
| B    | 1 0.03     | 1 0.04     | Cell cycle control, Cell division, chromosome partitioning |
| D    | 36 0.93    | 31 1.12    | Defense mechanisms |
| V    | 43 1.11    | 29 1.05    | Signal transduction mechanisms |
| T    | 283 7.32   | 218 7.89   | Cell wall/membrane biogenesis |
| M    | 265 6.86   | 210 7.60   | Cell motility |
| N    | 66 1.71    | 64 2.32    | Intracellular trafficking and secretion |
| U    | 123 3.18   | 98 3.55    | Posttranslational modification, protein turnover, chaperones |
| O    | 185 4.79   | 142 5.14   | Energy production and conversion |
| C    | 265 6.86   | 192 6.95   | Carbohydrate transport and metabolism |
| G    | 148 3.83   | 101 3.66   | Amino acid transport and metabolism |
| E    | 201 5.20   | 150 5.43   | Nucleotide transport and metabolism |
| F    | 63 1.63    | 59 2.14    | Coenzyme transport and metabolism |
| H    | 167 4.32   | 129 4.67   | Lipid transport and metabolism |
| I    | 90 2.33    | 65 2.35    | Inorganic ion transport and metabolism |
| P    | 189 4.89   | 127 4.60   | Secondary metabolites biosynthesis, transport and catabolism |
| Q    | 56 1.45    | 35 1.27    | General function prediction only |
| R    | 394 10.20  | 247 8.94   | Function unknown |
| S    | 346 8.95   | 230 8.32   | Not in COGs |
| -    | 943 24.40  | 598 21.64  | |
Fig. 3 (See legend on next page.)
copies of the \textit{rbcL} and \textit{rbcS} genes, encoding large and small subunits of ribulose bisphosphate carboxylase/oxygenase (\textit{SVA\_3460-3459}, \textit{SVA\_3471-3470}; \textit{SCL\_2417-2416}, \textit{SCL\_2425-2424}), which is the key enzyme in the Calvin-Benson-Bassham cycle to catalyze inorganic carbon fixation. The two copies of RuBisCO in each genome are phylogenetically distinct, and belong to lineages referred to as green-like form IA and red-like form IC (Fig. 3) [21]. In the form IC RuBisCO coded by \textit{rbcL} gene (\textit{SVA\_3460}, \textit{SCL\_2417}), \textit{Sulfurijustis variabilis} skN76\textsuperscript{T} and \textit{Sulfuricaulis limicola} HAS\textsuperscript{5} have six-amino-acid inserts at the same position where a similar insert was reported from \textit{Nitrosospira} sp. 40KI [22]. There are two other RuBisCO sequences which have six-amino-acid inserts at the same position, and these sequences with inserts formed a monophyletic cluster in the tree of RuBisCO (Fig. 3). In general, RuBisCO of form IA and IC have different properties which are thought to be advantageous to fix inorganic carbon under different concentrations of carbon dioxide and/or oxygen [21]. Possession of the genes for these two distinct RuBisCO forms may be beneficial to cope with changing environmental conditions, or to thrive in various types of ecosystems.

**Conclusion**

This is the first report on complete genome sequences of bacteria belonging to the order \textit{Acidiferrobacterales}. The genome analysis of \textit{Sulfurijustis variabilis} skN76\textsuperscript{T} and \textit{Sulfuricaulis limicola} HAS\textsuperscript{5} revealed that they have similar sets of genes involved in sulfur oxidation pathways. In the both genomes, redundancies of the genes for sulfur oxidation and inorganic carbon fixation were observed, as represented by multiple copies of \textit{dsrAB}, \textit{aprAB} and \textit{rbcLS}. Such redundancies may provide physiological flexibility to the chemolithotrophic sulfur oxidizers which are fully depending on these functions to obtain energy and carbon source for growth.

**Abbreviations**

MiGAP: Microbial Genome Annotation Pipeline; Hdr: Heterodisulfide reductase; DSR: Dissimilatory sulfite reductase

**Acknowledgements**

This study was supported by JSPS KAKENHI Grant Number 15 K07209 to H. Kojima. We thank R. Tokizawa and A. Shinohara for their technical assistance.

**Authors’ contribution**

MF and HK designed the study. HK characterized the strains and prepared genomic DNA. KU, TW and AM performed the bioinformatics analysis. KU and HK wrote the draft of manuscript. All authors read and approved the final manuscript.

**Competing interests**

The authors declare that they have no competing interests.

**References**

1. Kojima H, Shinohara A, Fukui M. Sulfurijustis variabilis gen. nov., sp. nov., a sulfur oxidizer isolated from a lake, and proposal of \textit{Acidiferrobacteriales} fam. nov. and \textit{Acidiferrobacteraceae} ord. nov. Int J Syst Evol Microbiol. 2015;65:3709–13.
2. Kojima H, Watanabe T, Fukui M. Sulfuricaulis limicola gen. nov., sp. nov., a sulfur oxidizer isolated from a lake. Int J Syst Evol Microbiol. 2016;66:266–70.
3. Hallberg KB, Hedrich S, Johnson DB. Acidiferrobacter thiooxydans, gen. nov. sp. nov., an acidophilic, thermo-tolerant, facultatively anaerobic iron- and sulfur-oxidizer of the family \textit{Estuarihodospiraceae}. ExtremeMicrobes. 2011;15:271–9.
4. Kelly DP, Wood AP. Reclassification of some species of \textit{Thiobacillus} to the newly designated genera \textit{Acidithioaculis} gen. nov., \textit{Halothiaobacillus} gen. nov. and \textit{Thermithiobacillus} gen. nov. Int J Syst Evol Microbiol. 2000;50:511–6.
5. Williams KP, Kelly DP. Proposal for a new class within the phylum \textit{Proteobacteria}. \textit{Acidithiobacillus} classis nov., with the type order \textit{Acidithiobacillales}, and emended description of the class \textit{Gamma proteobacteria}. Int J Syst Evol Microbiol. 2013;63:2901–6.
6. Dykoma S, Bischof K, Fuchs BM, Hoffmann K, Meier D, Meyerdierks A, et al. Ubiquitous \textit{Gammaproteobacteria} dominate dark carbon fixation in coastal sediments. ISME J. 2011;60:199–212.
7. Sugawara H, Ohyama A, Mori H, Kurokawa K. Microbial Genome Annotation Pipeline (MiGAP) for diverse users. The 20th International Conference on Genome Informatics (GIW2009) Poster and Software Demonstrations (Yokohama), 2009;S001-1-2.
8. Wu S, Zhu Z, Fu L, Niu B, Li W. WebMGA: a customizable web server for fast metagenomic sequence analysis. BMC Genomics. 2011;12:444.
9. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215:403–10.
10. Eddy SR. Profile hidden Markov models. Bioinformatics. 1998;14:755–63.
11. Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, et al. The Pfam protein families database. Nucleic Acids Res. 2010;38:D218–22.
12. Kall L, Krogh A, Sonnhammer ELL. Advantages of combined transmembrane topology and signal peptide prediction-the Phobius web server. Nucleic Acids Res. 2007;35:W429–32.
13. Grissa I, Vergnaud G, Pourcel C. CRISPRfinder: a web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Res. 2007;35:W52–7.
14. Pott AS, Dahl C. Sirohaem sulfite reductase and other proteins encoded by genes at the \textit{dsr} locus of \textit{Chromatium vinosum} are involved in the oxidation of intracellular sulfur. Microbiology. 1998;144:1881–94.
15. Dahl C, Engels S, Pott-Sperling AS, Schulte A, Sander J, Lübbe Y, et al. Novel genes at the \textit{dsr} locus of the phototrophic sulfur bacterium \textit{Allochromatium vinosum}–phosphosulfate reductase-encoding genes. Nucleic Acids Res. 2010;38:D211.
16. Parey K, Demmer U, Warkentin E, Wynen A, Ermler U, Dahl C. Structural, biochemical and genetic analysis of dissimilatory adenosine-5'-phosphosulfate reductase-encoding genes (\textit{aprB}) among sulfur-oxidizing prokaryotes. Microbiology. 2007;153:3478–98.
17. Dahle C, Franz B, Hensen D, Kesselheim A, Ziegenn R. Sulfite oxidation in the purple sulfur bacterium \textit{Allochromatium vinosum}: identification of SoeABC as a major player and relevance of SoxYZ in the process. Microbiology. 2013;159(26):26–38.
20. Kappler U, Bennett B, Rethmeier J, Schwarz G, Deutzmann R, McEwan AG, et al. Sulfite:cytochrome c oxidoreductase from *Thiobacillus novellus* — purification, characterization and molecular biology of a heterodimeric member of the sulfite oxidase family. J Biol Chem. 2000;275:13202–12.

21. Badger MR, Bek EJ. Multiple Rubisco forms in proteobacteria: their functional significance in relation to CO₂ acquisition by the CBB cycle. J Exp Bot. 2008;59:1525–41.

22. Utåker JB, Andersen K, Aakra Å, Moen B, Nes IF. Phylogeny and functional expression of ribulose 1,5-Bisphosphate carboxylase/oxygenase from the autotrophic ammonia-oxidizing bacterium *Nitrosospira* sp. Isolate 40K. J Bacteriol. 2002;184:468–78.

23. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci U S A. 1990;87:4576–9.

24. Garrity GM, Bell JA, Lilburn T. Phylum XIV. *Proteobacteria* phyl. nov. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT, editors. Bergey’s Manual of Systematic Bacteriology, Volume 2, Part B. 2nd ed. New York: Springer; 2005. p. 1.

25. Garrity GM, Bell JA, Lilburn T. Class III. Gammaproteobacteria class. nov. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT, editors. Bergey’s Manual of Systematic Bacteriology, Volume 2, Part B. 2nd ed. New York: Springer; 2005. p. 1.

26. Kojima H, Ivata T, Fukui M. DNA-based analysis of planktonic methanotrophs in a stratified lake. Freshw Biol. 2009;54:1501–9.

27. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGAs: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 2013;30:2725–9.