

1 **Expanded diversity of microbial groups that**
2 **shape the global sulfur cycle**

3
4 **Authors:** Karthik Anantharaman^{1*#}, Bela Hausmann², Sean P. Jungbluth³, Rose S. Kantor⁴, Adi
5 Lavy¹, Lesley A. Warren⁵, Michael S. Rappé⁶, Michael Pester⁷, Alexander Loy², Brian C.
6 Thomas¹, and Jillian F. Banfield^{1,8,9}

7 **Affiliations:**

8 ¹Department of Earth and Planetary Sciences, Berkeley, CA, USA

9 ²Division of Microbial Ecology, Department of Microbiology and Ecosystem Science, Research
10 Network Chemistry meets Microbiology, University of Vienna, Vienna, Austria

11 ³DOE Joint Genome Institute, Walnut Creek, CA, USA

12 ⁴Department of Plant and Microbial Biology, University of California, Berkeley, CA, USA

13 ⁵Department of Civil Engineering, University of Toronto, Ontario, Canada

14 ⁶Hawaii Institute of Marine Biology, University of Hawaii at Manoa, Kaneohe, HI, USA

15 ⁷Department Microorganisms, Leibniz Institute DSMZ – German Collection of Microorganisms
16 and Cell Cultures, Braunschweig, Germany

17 ⁸Department of Environmental Science, Policy, and Management, Berkeley, CA, USA

18 ⁹Earth and Environmental Sciences, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

19 # Present address: Department of Bacteriology, University of Wisconsin-Madison, WI, USA

20 *Corresponding author

21 Email: karthik@bact.wisc.edu

22
23
24
25 Correspondence and requests for materials should
26 be addressed to K.A. (karthik@bact.wisc.edu)

27
28
29
30
31
32

33 This PDF file includes:

34

35 Supplementary Figures 1 to 14

36

37 **Supplementary Fig. 1. Sequence alignment of DsrA proteins identified in this study with**
38 **reference sequences.**

39

40 **Supplementary Fig. 2. Sequence alignment of DsrB proteins identified in this study with**
41 **reference sequences.**

42

43 **Supplementary Fig. 3. Sequence alignment of AsrC proteins identified in this study with**
44 **reference sequences.**

45

46 **Supplementary Fig. 4. Genome summary showing individual sulfur oxidation, and sulfate**
47 **and sulfite reduction genes (dsr only) identified in this study.**

48

49 **Supplementary Fig. 5. Sequence alignment of DsrD proteins identified in this study with**
50 **reference sequences.**

51

52 **Supplementary Fig. 6. Phylogenetic analysis of dissimilatory sulfite reductase subunit D**
53 **genes inferred by maximum-likelihood.**

54

55 **Supplementary Fig. 7. Phylogenetic analysis of quinone-interacting membrane**
56 **oxidoreductase genes inferred by maximum-likelihood.**

57

58 **Supplementary Fig. 8. Phylogenetic analysis of adenosine phosphosulfate reductase genes**
59 **inferred by maximum-likelihood.**

60

61 **Supplementary Fig. 9. Phylogenetic analysis of sulfate adenylyltransferase (ATP**
62 **sulfurylase) genes inferred by maximum-likelihood.**

63

64 **Supplementary Fig. 10. Phylogenetic analysis of newly identified *dsrT* genes.**

65

66 **Supplementary Fig. 11. Example of observed *dsrT-dsrM* gene fusion and its comparison to**
67 **typical *dsrT*-containing operons.**

68

69 **Supplementary Fig. 12. Structural model of DsrT proteins.**

70

71 **Supplementary Fig. 13. Sequence alignment of DsrC proteins identified in this study with**
72 **reference sequences.**

73

74

75

76

77

78

79 Other Supplementary Files

80

81 **Data File S1. Complete RP tree in newick format.**

82

83 **Data File S2. Complete dsrAB tree in newick format.**

84

85 **Data File S3. Complete reductive dsrAB tree inferred by neighbor joining.**

86

87 **Data File S4. Complete reductive dsrAB tree inferred by maximum likelihood.**

88

89 **Data File S5. Complete 16S ribosomal RNA tree inferred by neighbor joining.**

90

91 **Data File S6. Complete 16S ribosomal RNA tree inferred by maximum likelihood.**

92

93

94

95

96

97

98

99

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

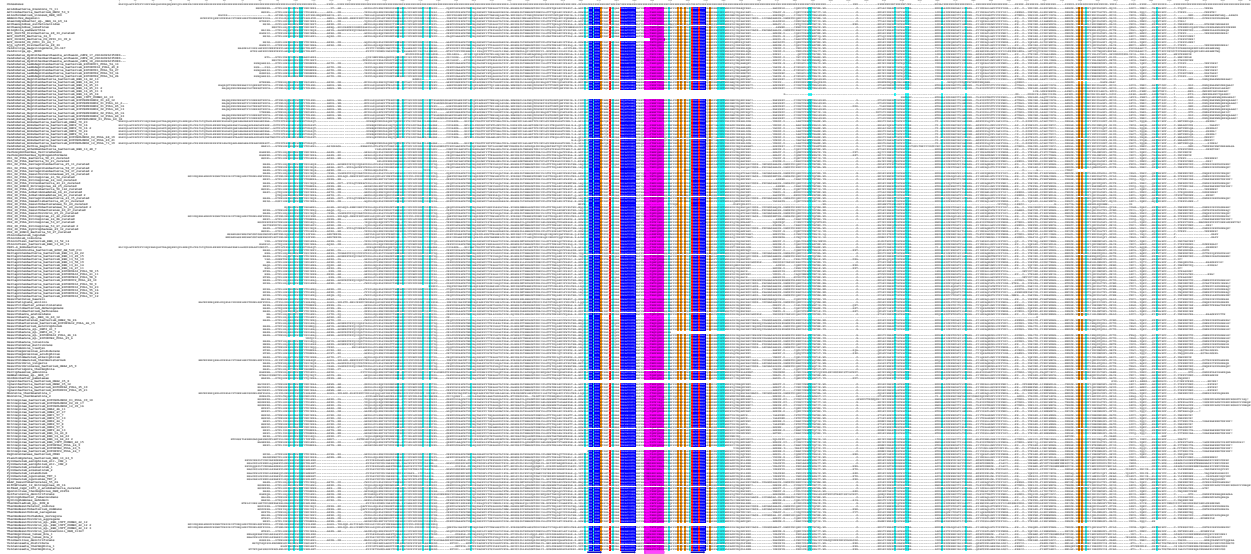
121

122

123

124

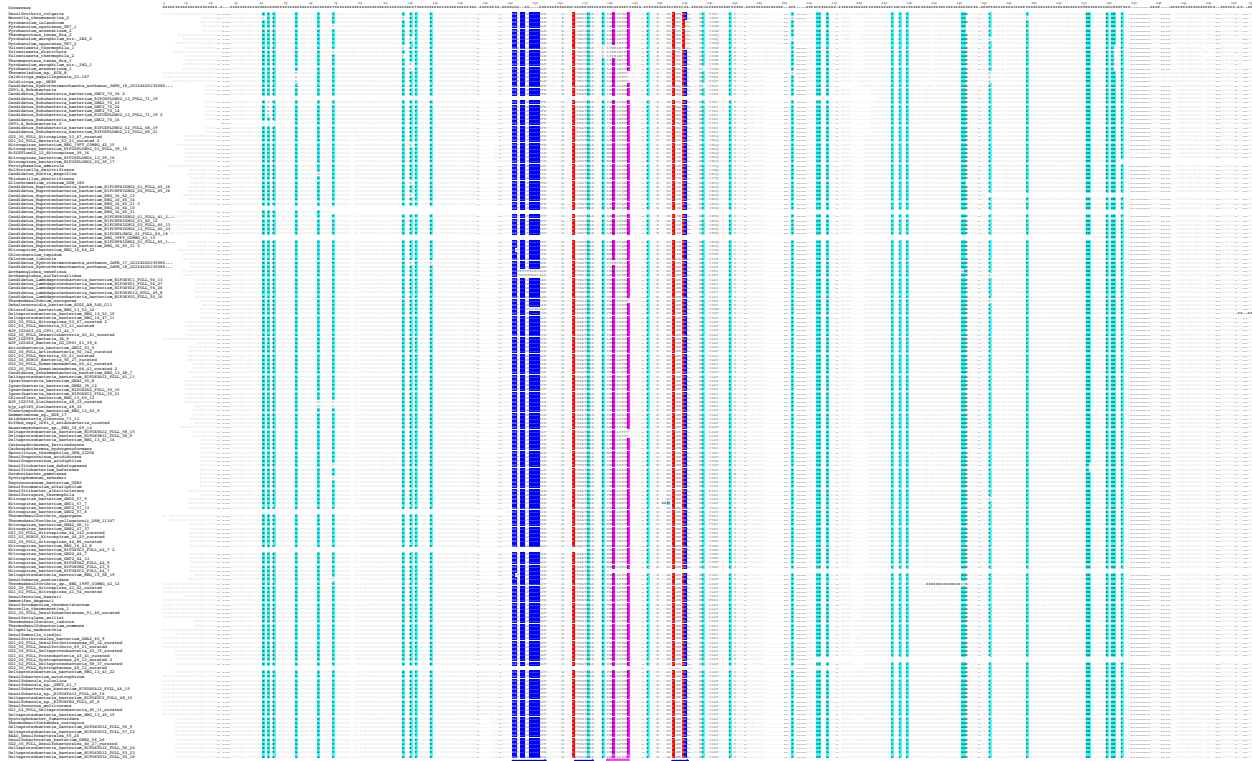
FIG. S1



125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156

Supplementary Fig. 1. Sequence alignment of DsrA proteins identified in this study with reference sequences. Conserved residues are highlighted as follows: (i) Red – Strictly conserved binding sites for the siroheme-[4Fe4S] (2) (ii) Orange – Substrate binding sites in dsrA (2) (iii) Cyan – Highly conserved residues of the same type (9). The secondary structure elements are indicated below the alignments with blue arrows depicting α -helices and pink cylinders indicating β -sheets.

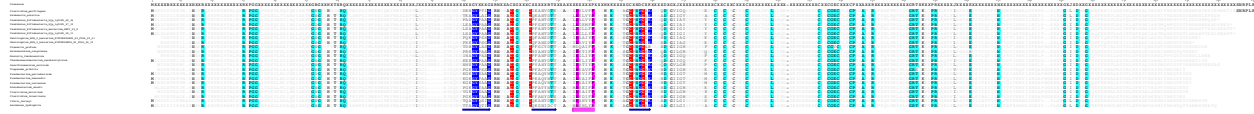
Fig. S2



157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180

Supplementary Fig. 2 Sequence alignment of DsrB proteins identified in this study with reference sequences. Conserved residues are highlighted as follows: (i) Red – Strictly conserved binding sites for the siroheme-[4Fe4S] (2) (ii) Cyan – Highly conserved residues of the same type (9). The secondary structure elements are indicated below the alignments with blue arrows depicting α -helices and pink cylinders indicating β -sheets.

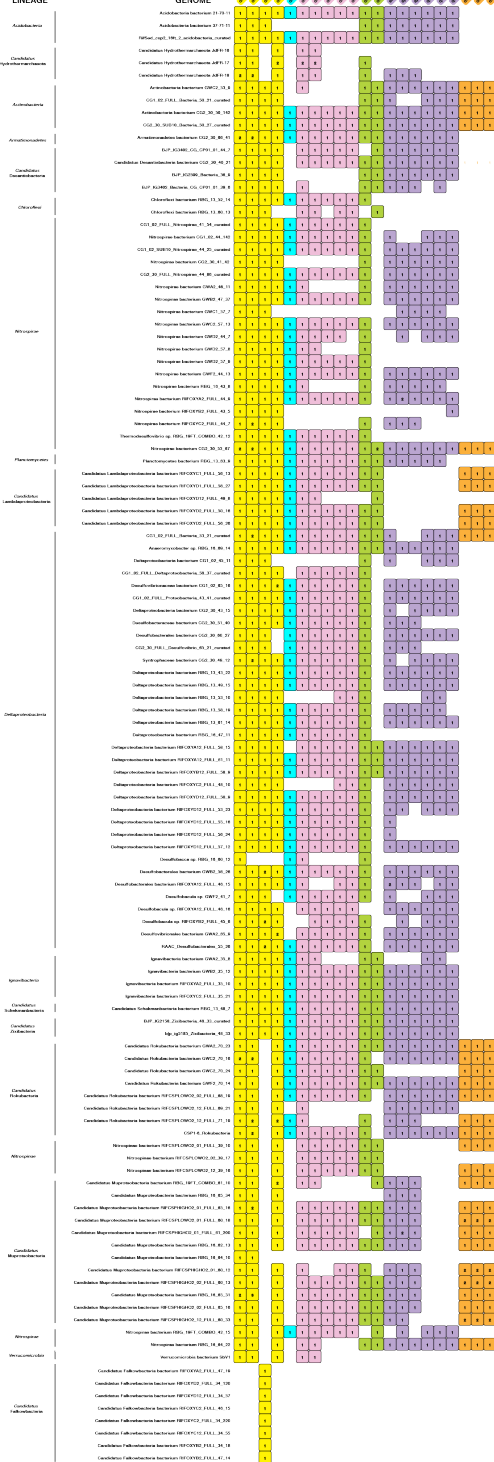
Fig. S3



181
182 **Supplementary Fig. 3. Sequence alignment of AsrC proteins identified in this study with**
183 **reference sequences.** Conserved residues are highlighted as follows: (i) Red – Strictly conserved
184 binding sites for the siroheme-[4Fe4S] (2) (ii) Cyan – Highly conserved residues of the same
185 type (9). The secondary structure elements are indicated below the alignments with blue arrows
186 depicting α -helices and pink cylinders indicating β -sheets.

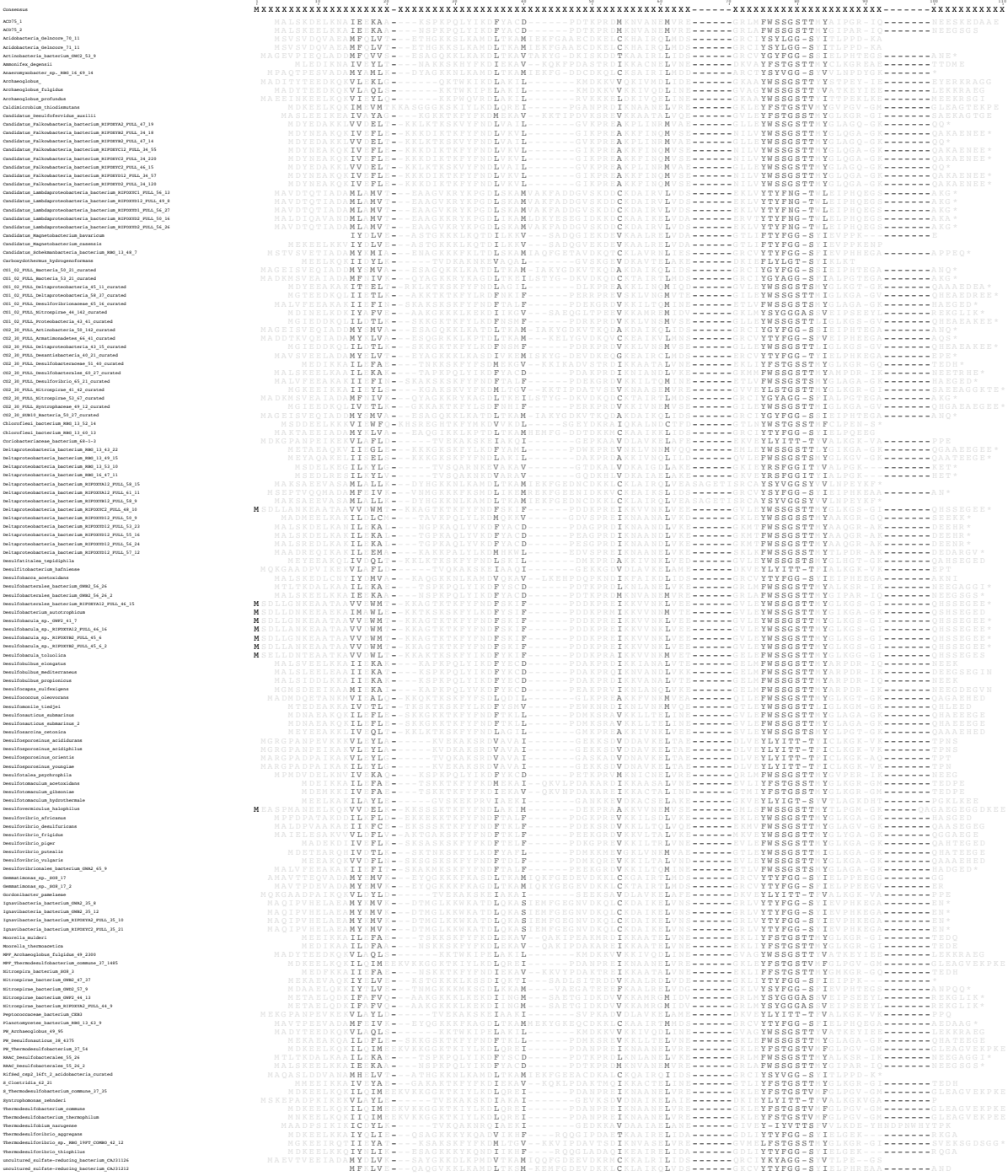
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224

Fig. S4



225
 226 **Supplementary Fig. 4. Genome summary showing individual sulfur oxidation, and sulfate**
 227 **and sulfite reduction genes (dsr only) identified in this study.** Colors indicate specific
 228 functions of the dsr genes: Yellow: Core genes (*dsrABCD*); Blue: Putative regulation (*dsrT*);
 229 Pink: Electron transport transmembrane complex (*dsrMKJOP*); Green: Ancillary proteins;
 230 Purple: Sulfate activation and reduction to sulfite; Orange: sulfur trafficking proteins;
 231 Presence/Absence of traits are shown in binary form - Presence: ≥ 1 , Absence: 0.

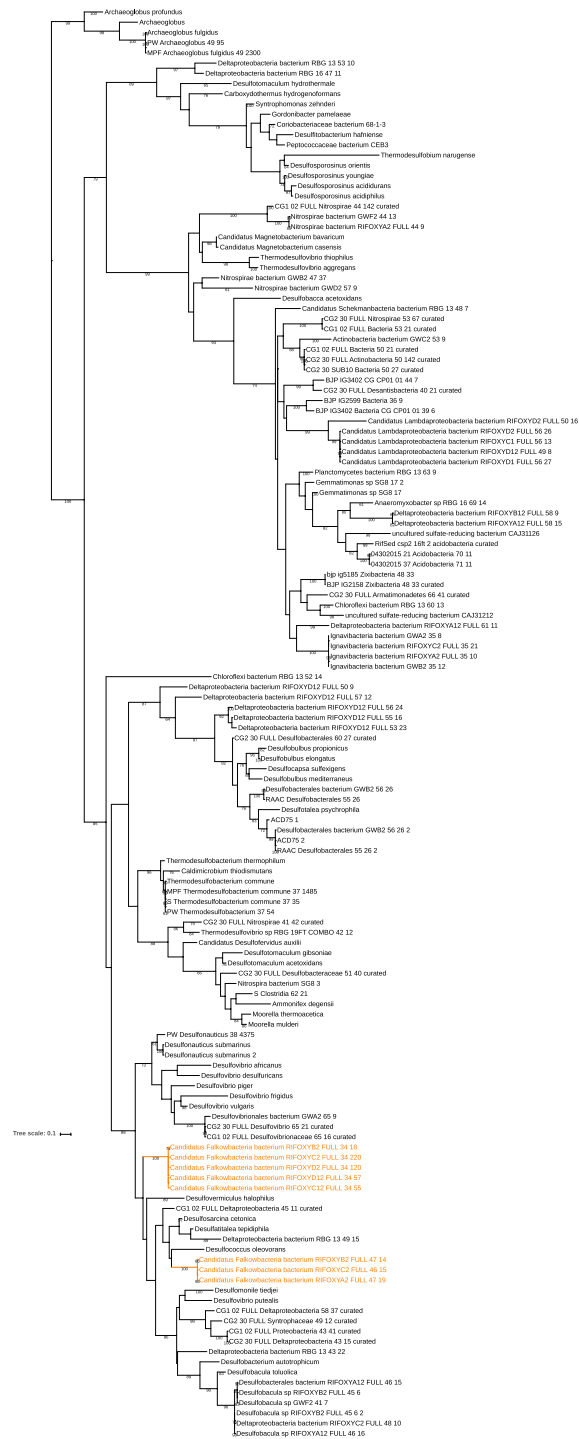
Fig. S5



232
233
234
235
236
237
238

Supplementary Fig. 5. Sequence alignment of DsrD proteins identified in this study with reference sequences. Conserved residues are highlighted as follows: (i) Red – highly conserved residues in the DsrD family (10) (ii) Cyan – Conserved residues of the same type.

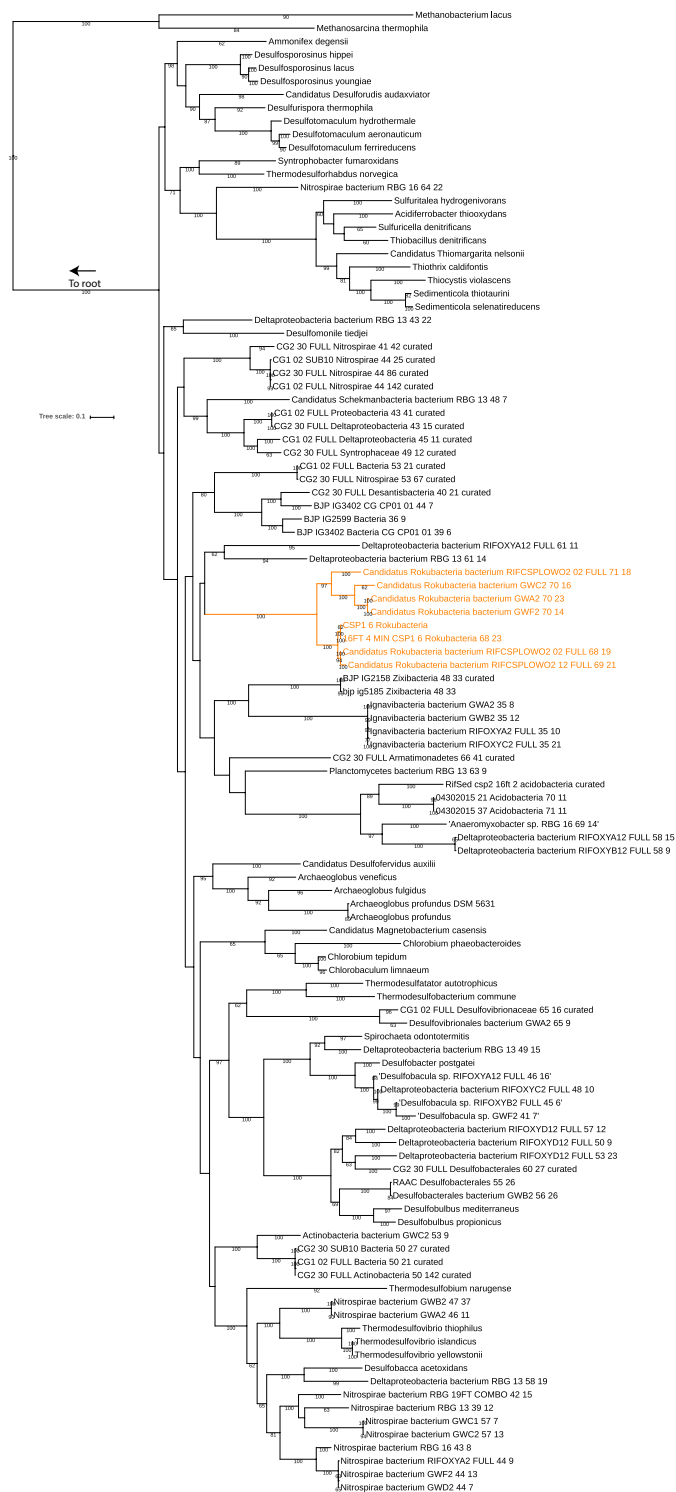
Fig. S6



239
240
241
242
243

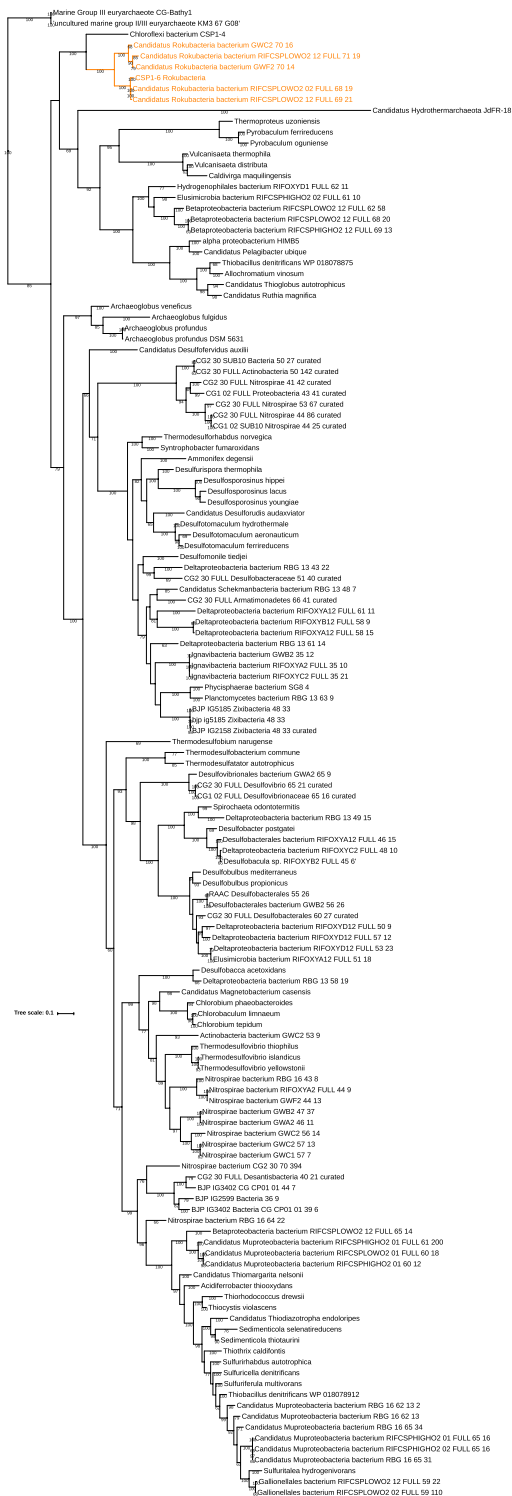
Supplementary Fig. 6. Phylogenetic analysis of dissimilatory sulfite reductase subunit D genes inferred by maximum-likelihood. Sequences in Orange are derived from the Candidate phylum Falkowbacteria. Only bootstrap values above 60 are shown.

Fig. S7



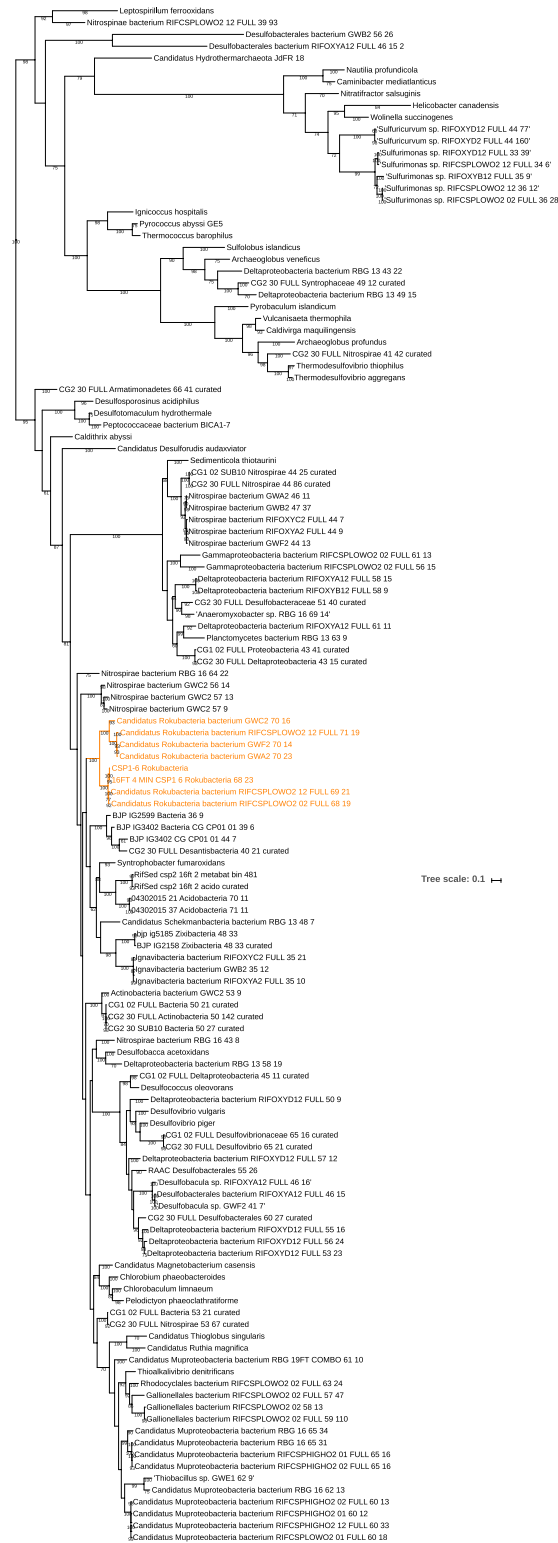
244
 245 **Supplementary Fig. 7. Phylogenetic analysis of quinone-interacting membrane**
 246 **oxidoreductase genes inferred by maximum-likelihood.** The phylogenetic tree is based on
 247 concatenated quinone-interacting membrane oxidoreductase subunits A and B. Sequences in
 248 Orange are derived from the Candidate phylum Rokubacteria. Only bootstrap values above 60
 249 are shown.

Fig. S8



250
251 **Supplementary Fig. 8. Phylogenetic analysis of adenosine phosphosulfate reductase genes**
252 **inferred by maximum-likelihood.** The phylogenetic tree is based on concatenated adenosine
253 phosphosulfate reductase subunits A and B. Sequences in Orange are derived from the Candidate
254 phylum Rokubacteria. Only bootstrap values above 60 are shown.

Fig. S9



255
256
257
258

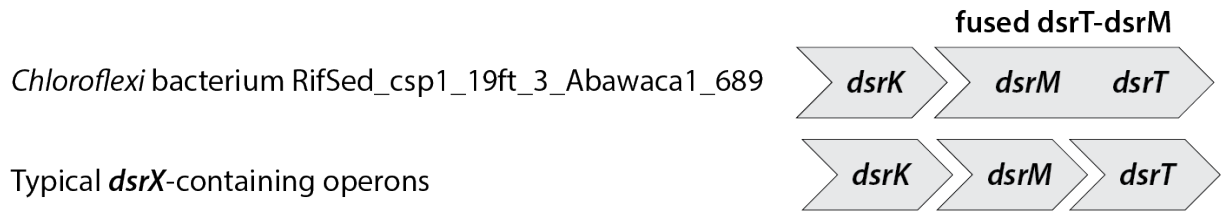
Supplementary Fig. 9. Phylogenetic analysis of sulfate adenylyltransferase (ATP sulfurylase) genes inferred by maximum-likelihood. Sequences in Orange are derived from the Candidate phylum Rokubacteria. Only bootstrap values above 60 are shown.

Fig. S10



259
260
261
262

Supplementary Fig. 10. Phylogenetic analysis of newly identified *dsrT* genes. Sequences in Orange are derived from the Candidate phylum Rokubacteria. Only bootstrap value above 60 are shown.

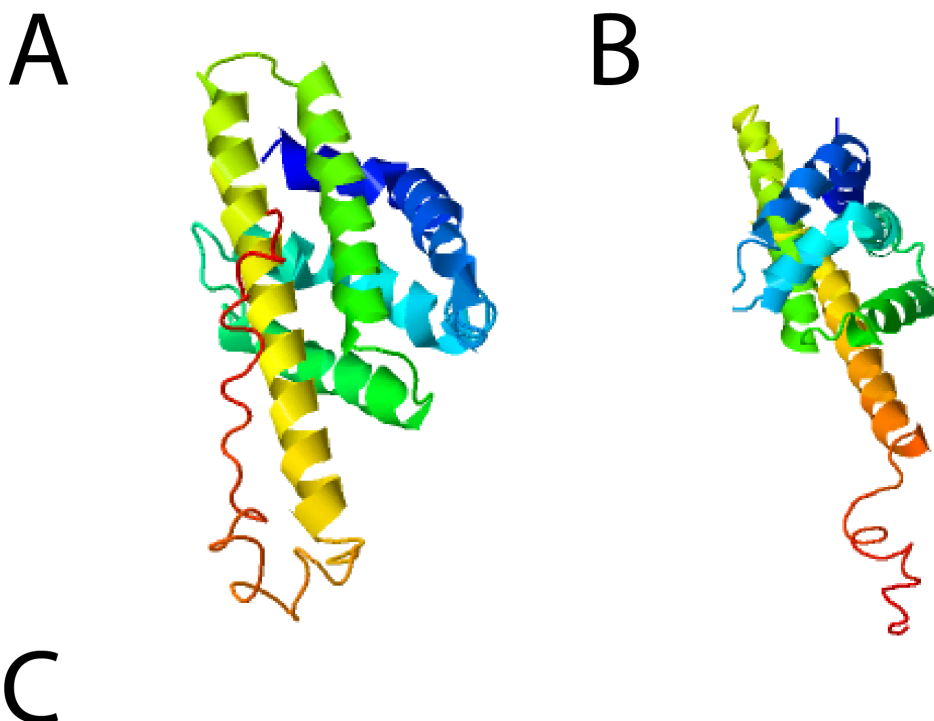


Chloroflexi bacterium RifSed_csp1_19ft_3_Abawaca1_689

Typical *dsrX*-containing operons

263
264
265
266
267
268
269
270
271
272
273
274
275
276
277

Supplementary Fig. 11. Example of observed *dsrT-dsrM* gene fusion and its comparison to typical *dsrT*-containing operons.



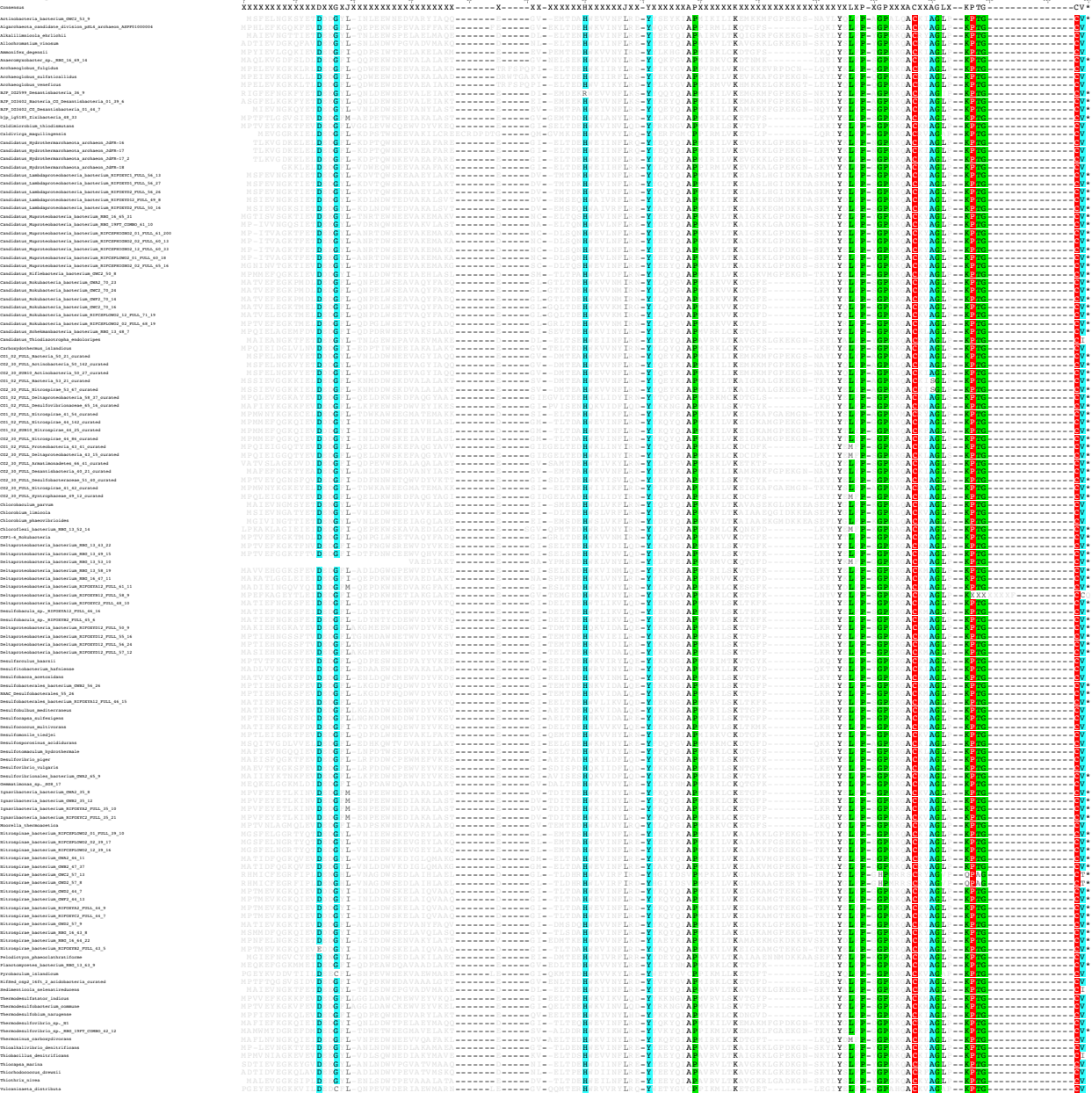
Rank	PDB Hit	Protein	TM-score	RMSD	IDEN	Cov
1	3pmdA	Sporulation inhibitor pXO1-118 from Bacillus anthracis	0.762	1.97	0.085	0.836
2	3pmcA	Sporulation inhibitor pXO2-61 from Bacillus anthracis	0.674	1.62	0.114	0.721
3	4uiqA	Globin domain of the Bordetella pertussis globin-coupled sensor	0.636	2.4	0.12	0.754
4	4zvaA	Globin domain of the E. coli DosC - form I (ferric)	0.636	2.15	0.087	0.732
5	2w31A	Globin domain of Geobacter Sulfurreducens globin-coupled sensor	0.626	2.43	0.115	0.732
6	2c71B	Low temperature structure of phycoerythrocyanin from Mastigocladus Laminosus	0.623	2.79	0.092	0.76
7	5aqdM	Phormidium Phycoerythrin at pH 8.5	0.621	3.31	0.129	0.792
8	2vm1D	Phycocyanin from Gloeobacter Violaceus	0.621	2.78	0.1	0.754
9	3ztaA	Bacterial stressosome from Moorella Theroacetica	0.621	2.56	0.089	0.738
10	1qgwD	Phycoerythrin 545 from Rhodomonas CS24	0.621	2.97	0.09	0.776

278
279

280 **Supplementary Fig. 12 Structural model of DsrT proteins** from (A) Candidatus Rokubacteria
 281 CSP1-6, (B) *Desulfovibrio vulgaris*. (C) Top 10 structural homologs of DsrT in RCSB Protein
 282 Data Bank as identified by TM-Align. Rank is based on TM-Align scores. RMSD is the RMSD
 283 between structurally aligned residues. IDEN is the percentage sequence identity in the
 284 structurally aligned region. Cov represents the number of structurally aligned residues divided by
 285 length of the query protein.

286
287
288
289
290
291
292

Fig. S13



293
294
295
296
297
298
299
300
301
302
303
304
305

Supplementary Fig. 13. Sequence alignment of DsrC proteins identified in this study with reference sequences. Conserved residues are highlighted as follows: (i) Red – Strictly conserved cysteines (11) (ii) Cyan – Highly conserved residues of the same type.

306 **Supplementary References**

- 307
- 308 1. S. S. Venceslau, Y. Stockdreher, C. Dahl, I. A. C. Pereira, The “bacterial heterodisulfide”
309 DsrC is a key protein in dissimilatory sulfur metabolism. *Biochim. Biophys. Acta -*
310 *Bioenerg.* **1837**, 1148–1164 (2014).
 - 311 2. T. F. Oliveira *et al.*, The Crystal Structure of *Desulfovibrio vulgaris* Dissimilatory Sulfite
312 Reductase Bound to DsrC Provides Novel Insights into the Mechanism of Sulfate
313 Respiration. *J. Biol. Chem.* **283**, 34141–34149 (2008).
 - 314 3. F. Grein, A. R. Ramos, S. S. Venceslau, I. A. C. Pereira, Unifying concepts in anaerobic
315 respiration: Insights from dissimilatory sulfur metabolism. *Biochim. Biophys. Acta -*
316 *Bioenerg.* **1827**, 145–160 (2013).
 - 317 4. Y. Stockdreher *et al.*, Cytoplasmic sulfurtransferases in the purple sulfur bacterium
318 *Allochromatium vinosum*: evidence for sulfur transfer from DsrEFH to DsrC. *PLoS One.*
319 **7**, e40785 (2012).
 - 320 5. F. Grimm, B. Franz, C. Dahl, Regulation of Dissimilatory Sulfur Oxidation in the Purple
321 Sulfur Bacterium *Allochromatium Vinosum*. *Front. Microbiol.* **2**, 51 (2011).
 - 322 6. H. Laue, M. Friedrich, J. Ruff, A. M. Cook, Dissimilatory sulfite reductase
323 (desulfovirdin) of the taurine-degrading, non-sulfate-reducing bacterium *Bilophila*
324 *wadsworthia* RZATAU contains a fused DsrB-DsrD subunit. *J. Bacteriol.* **183**, 1727–33
325 (2001).
 - 326 7. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods.* **9**,
327 357–9 (2012).
 - 328 8. J. Besemer, M. Borodovsky, GeneMark: web software for gene finding in prokaryotes,
329 eukaryotes and viruses. *Nucleic Acids Res.* **33**, W451–W454 (2005).
 - 330 9. A. Dhillon, S. Goswami, M. Riley, A. Teske, M. Sogin, Domain Evolution and Functional
331 Diversification of Sulfite Reductases. *Astrobiology.* **5**, 18–29 (2005).
 - 332 10. N. Mizuno, G. Voordouw, K. Miki, A. Sarai, Y. Higuchi, Crystal Structure of
333 Dissimilatory Sulfite Reductase D (DsrD) Protein—Possible Interaction with B- and Z-
334 DNA by Its Winged-Helix Motif. *Structure.* **11**, 1133–1140 (2003).
 - 335 11. K. Anantharaman *et al.*, Sulfur oxidation genes in diverse deep-sea viruses. *Science.* **344**,
336 757–60 (2014).

337