1   **The genome of the epsilonproteobacterial chemolithoautotroph**

2   ***Sulfurimonas denitrificans***

3

4   Stefan M. Sievert[‡][*][1], Kathleen M. Scott[‡][*][2], Martin G. Klotz[3], Patrick S. G. Chain[4,5],

5   Loren J. Hauser[6], James Hemp[7], Michael Hügler[1,8], Miriam Land[6], Alla Lapidus[5], Frank

6   W. Larimer[6], Susan Lucas[5], Stephanie A. Malfatti[4,5], Folker Meyer[9], Ian T. Paulsen[10#],

7   Qinghu Ren[10], Jörg Simon[11], and the USF Genomics Class[2, Δ]

8

9   *Biology Department, Woods Hole Oceanographic Institution, Woods Hole,*

10   *Massachusetts[1]; Biology Department, University of South Florida, Tampa, Florida[2];*

11   *Departments of Biology and Microbiology & Immunology, University of Louisville,*

12   *Louisville, Kentucky[3]; Lawrence Livermore National Laboratory, Livermore, California[4];*

13   *Joint Genome Institute, Walnut Creek, California[5]; Oak Ridge National Laboratory, Oak*

14   *Ridge, Tennessee[6]; Center for Biophysics and Computational Biology, University of*

15   *Illinois at Urbana-Champaign, Urbana, Illinois[7]; Leibniz-Institut für*

16   *Meereswissenschaften, Kiel, Germany[8]; Mathematics and Computer Science Division,*

17   *Argonne National Laboratory, Argonne, Illinois[9]; The Institute for Genomic Research,*

18   *Rockville, Maryland[10]; Institute of Molecular Biosciences, Johann Wolfgang Goethe*

19   *University, Frankfurt am Main, Germany[11]*

20

21   [#]Current address: Department of Chemistry and Biomolecular Sciences, Macquarie
22   University, Sydney, Australia
23
24   [‡]SMS and KMS contributed equally to this work.

25

26   ***Corresponding authors. SMS: Mailing address: Woods Hole Oceanographic Institution;

27   Watson Building 207, MS#52; Woods Hole, MA 02543. Phone: (508) 289-2305. Fax:

28   (508) 457-2076. E-mail: ssievert@whoi.edu. KMS: Mailing address: 4202 East Fowler

29   Avenue, SCA 110; Tampa, FL 33620. Phone: (813) 974-5173. Fax (813) 974-3263. E-

30   mail: kscott@cas.usf.edu.

31

32   ᐃKathryn Bailey, Erik Diaz, Kelly Ann Fitzpatrick, Bryan Glover, Natasha Gwatney,

33   Asja Korajkic, Amy Long, Jennifer M. Mobberley, Shara N. Pantry, Geoffrey Pazder,

34   Sean Peterson, Joshua D. Quintanilla, Robert Sprinkle, Jacqueline Stephens, Phaedra

35   Thomas, Roy Vaughn, M Joriane Weber, Lauren L. Wooten

36

37   Running title: *Sulfurimonas denitrificans* genome

38

39   Journal section: Evolutionary and Genomic Microbiology

40

41   (abstract)

42   **Sulfur-oxidizing epsilonproteobacteria are common in a variety of sulfidogenic**

43   **environments. These autotrophic and mixotrophic sulfur-oxidizing bacteria are**

44   **believed to contribute substantially to the oxidative portion of the global sulfur**

45   **cycle. In order to better understand the ecology and roles of sulfur-oxidizing**

46   **epsilonproteobacteria, in particular the widespread genus *Sulfurimonas*, in**

47   **biogeochemical cycles, the genome of *Sulfurimonas denitrificans* DSM1251 was**

48   **sequenced. This genome has many features, including a larger size (2.2 Mbp), that**

49   **suggest a greater degree of metabolic versatility or responsiveness to the**

50   **environment than most of the other sequenced epsilonproteobacteria. A branched**

51   **electron transport chain is apparent, with genes encoding complexes for the**

52   **oxidation of hydrogen, reduced sulfur compounds, and formate, and the reduction**

53   **of nitrate and oxygen. Genes are present for a complete, autotrophic reductive citric**

54   **acid cycle. Many genes are present that could facilitate growth in the spatially and**

55   **temporally heterogeneous sediment habitat from where *Sulfurimonas denitrificans***

56   **was originally isolated. Many resistance-nodulation-development-family transporter**

57   **genes (11 total) are present, several of which are predicted to encode heavy metal**

58   **efflux transporters. An elaborate arsenal of sensory and regulatory protein-**

59   **encoding genes is in place, as well as genes necessary to prevent and respond to**

60   **oxidative stress.**

61   Epsilonproteobacteria are represented as molecular isolates from a vast array of

62   habitats, including brackish, marine, and subsurface (e.g., (3, 11, 25, 39, 45, 71, 74);

63   reviewed in (6)). Over the last years, quite a few cultured representatives of this group

64   have been obtained from these environments, and currently all cultured members are

65   chemolithoautotrophs or mixotrophs, capable of using sulfur and/or hydrogen as electron

66   donors, while denitrifying and/or growing as microaerophiles (reviewed in (6)). These

67   organisms use the reductive citric acid cycle for carbon fixation (24, 67). Given their

68   abundance, sulfur-oxidizing epsilonproteobacteria, in particular members of the genus

69   *Sulfurimonas*, are believed to be relevant to the function of the global sulfur cycle (6).

70   Genome data from these organisms would be key to metagenome sequencing

71   efforts in habitats where they are abundant, and would also, by comparison to other

72   epsilonproteobacteria, be helpful for determining the traits unique to a free-living,

73   autotrophic lifestyle versus a host-associated, heterotrophic lifestyle. Recently, the

74   genome sequences of *Sulfurovum* sp. NBC37-1 and *Nitratiruptor* sp. SB155-2, two

75   sulfur-oxidizing epsilonproteobacteria from deep-sea hydrothermal vents were published,

76   which revealed that these organisms share many features with their pathogenic (e.g.,

77   *Campylobacter* and *Helicobacter* spp.) epsilonproteobacterial relatives (41). Given the

78   remarkable variety of habitats where sulfur-oxidizing epsilonproteobacteria are found, it

79   was of great interest to also conduct these analyses on non-vent epsilonproteobacteria. To

80   represent the abundant sulfur-oxidizing epsilonproteobacteria present in coastal marine

81   sediments, we chose to sequence and analyze the genome of the sulfur-oxizing

82   chemolithoautotroph *Sulfurimonas denitrificans* DSM1251. Based on its phenotype, *S.*

83   *denitrificans* was originally named *Thiomicrospira denitrificans* (70). Subsequent

84   sequencing revealed the polyphyletic nature of the *Thiomicrospiras,* with members from

85   both the gammaproteobacteria and epsilonproteobacteria (40). As a result,

86   *Thiomicrospira denitrificans* was eventually removed from the genus *Thiomicrospira* and

87   placed within the genus *Sulfurimonas* (68). In addition to marine sediments, bacteria

88   belonging to this genus have been isolated or detected in a variety of sulfidogenic

89   environments, including deep-sea hydrothermal vents, the oxic-anoxic interface of

90   marine anoxic basins, and oilfield (6, 18, 25, 32), making organisms of this genus

91   globally significant.

92

93   **MATERIALS AND METHODS**

94

95   **Genome sequencing, annotation, and analysis.** DNA libraries were created and

96   sequenced to approximately 13X depth of coverage at the Production Genomics Facility

97   of the Joint Genome Institute (JGI) using the whole-genome shotgun method as

98   previously described (7, 57). Gaps were closed and base quality problems were

99   addressed by sequencing finishing reads, and PHRED/PHRAP/CONSED were used for

100 assembly (12, 13, 20). Automated and manual annotation were conducted by ORNL

101 similarly to (7, 57). Results were collated and presented via GenDB (37) for manual

102 verification. The prediction of membrane transporters was based on a transporter

103 annotation pipeline that uses several predictive approaches such as BLAST, COG, PFAM

104 and TIGRFAM HMM searches, transmembrane topology prediction algorithms, and

105 takes advantage of a curated database of transporters. Details of this pipeline and

106 database have been published in (50, 51, 57). The main limitation of this approach is the

107 ability to accurately predict precise transporter specificities. Based on both internal and

108 external testing, this methodology is highly successful at identifying putative transporters

109 and predicting approximate substrate specificity. However, making precise substrate

110 predictions, eg., serine transport rather than amino acid transport, is more problematic.

111 The other related limitation is that the approach is dependent on comparison with known

112 experimentally characterized transporters, so completely novel transporters which have

113 never had homologues experimentally characterized will not be predicted by this

114 methodology.

115 　　　　To uncover genes involved in oxidative stress, the *S. denitrificans* genome was

116 examined with a series of BLAST queries, using genes known to be involved in oxidative

117 stress response in *H. pylori* (73).

118 **　　　　Identification of genes encoding signal transduction and regulatory proteins.**

119 The complement of genes that encode signal transduction and regulatory proteins were

120 compared among *S. denitrificans* DSM1251, *Thiomicrospira crunogena* XCL-2, and

121 *Nitrosococcus oceani* ATCC 19707.   To compare signal transduction and regulatory

122 protein genes among these obligate autotroph genomes, genes were identified by

123 querying the predicted gene products to the InterPro (PRINTS, PFAM, TIGRFAM,

124 PRODOM and SMART) and COGs databases (via HMM search – InterPro, or RPSblast

125 – COGs) to identify domains indicative of a role in these processes (e.g., EAL, GGDEF,

126 PAS/PAC). Genes with predicted domains above the trusted cutoff score (InterPro) or an

127 e value less than e^{-5} (COGs) were assigned a product description and classified using a

128   set of rules based on the domain architecture of the protein. The final results were

129   manually verified.

130   **Nucleotide sequence accession number.** The complete sequence of the *S.*

131   *denitrificans* genome is available from GenBank (accession number NC_007575).

132

133

134   **RESULTS AND DISCUSSION**

135

136   **Genome structure.** The *S. denitrificans* DSM1251 genome is one of the largest

137   epsilonproteobacterial genomes yet sequenced, consisting of a single 2.2 Mbp

138   chromosome (Table 1). The coding density and G+C content are similar to the other

139   epsilonproteobacteria (Table 1). Four rRNA operons are present, which, due to their

140   elevated G+C content (~50%) relative to the genome average (34.5%), are visible as

141   positive G+C content anomalies on the genome map (Fig. 1). Three of these operons

142   (*16S-tRNA^{Ala}-23S-5S*) are 100% identical, and are oriented in the same direction, while

143   the fourth (*16S-tRNA^{Ile}-23S-5S*) is in the opposite orientation, and its 5S and 23S genes

144   each have a single nucleotide substitution compared to the others. The free-living *S.*

145   *denitrificans, Sulfurovum* sp. NBC37-1, and *Nitratiruptor* sp. SB155-2 have more rRNA

146   operons than those epsilonproteobacteria that are known to be exclusively host-related

147   (Table 1), which is likely a reflection of an adaptation to fluctuating environmental

148   conditions and the necessity for versatility (30, 41, 65).

149   Two large (17,627 bp), identical transposons are apparent as negative G+C

150   content anomalies (30.0%; Fig. 1). Flanked by identical 12 bp inverted repeats, these

151 transposons (*Suden_0690 – 0702; Suden_1587 - 1599*) include genes encoding

152 transposases, as well as proteins similar to the TniB (46%) and TniQ (47%) transposase

153 accessory proteins found in mercury resistance-transposons in *Xanthomonas* sp. W17 and

154 other systems (29). These transposons also include genes encoding a type I restriction-

155 modification methyltransferase and restriction enzyme (see Supplemental Materials).

156 Interestingly, one of the copies of this transposon interrupts a flagellin biosynthetic

157 operon, which may explain why, unlike close relatives (25, 68), this strain of *S.*

158 *denitrificans* is nonmotile (70). Other transposase and integrase genes are described in

159 Supplemental Materials.

160

161   **Transporters.** *S. denitrificans* has a modest complement of genes (97 total)

162 predicted to encode transporters. This number is similar to other sequenced heterotrophic

163 epsilonproteobacteria (75-124 genes), which is surprising given that nutrient

164 requirements for *S. denitrificans,* believed to be an obligate autotroph, are simple,

165 compared to the others. This similarity in transporter numbers is due in part to the

166 expansion of a few transporter families in this species compared to other

167 epsilonproteobacteria. Amt-family transporters, are encoded in the *S. denitrificans*

168 (*Suden_0641* and *0643*), *Sulfurovum* sp. NBC37-1 (2 copies), *Nitratiruptor* sp. SB155-2

169 (1 copy) and *Wolinella succinogenes* (1 copy) genome, but not in *Helicobacter pylori,*

170 *Helicobacter hepaticus,* or *Campylobacter jejuni,* and are likely to facilitate ammonium

171 uptake. Perhaps their absence in *Helicobacter* and *Campylobacter* spp. is due to nitrogen

172 requirements for these species being met primarily from exogenous urea and/or amino

173 acids (35, 54, 69). A FNT (formate-nitrite transporter) gene is present in *S. denitrificans*

174　(*Suden_0716),* and absent from other sequenced epsilonproteobacteria. Neither gene

175　context nor sequence comparison clarifies the substrate for this transporter. Also notable

176　is the presence of an abundance of Resistance-Nodulation-Cell Division (RND)

177　Superfamily genes (11 genes; *Suden_0270; Suden_0536; Suden_0799; Suden_0876;*

178　*Suden_0877; Suden_0883; Suden_1281; Suden_1440; Suden_1499; Suden_1574;*

179　*Suden_1575; Suden_2011*), compared to other epsilonproteobacteria, including the two

180　hydrothermal vent species (2-6 genes). Many of these are predicted to encode

181　transporters for metal efflux in *S. denitrificans*. As in *W. succinogenes* and the two

182　hydrothermal vent epsilonproteobacteria, an apparent operon is present that encodes a

183　cytoplasmic arsenate reductase (*Suden_0314)*, arsenite permease (*Suden_0313)*, and

184　regulatory protein ArsR (*Suden_0315;* (41, 61). Apparently the sediment ecosystem

185　inhabited by *S. denitrificans* requires a similar or perhaps enhanced level of resistance to

186　metals and other toxins, than the digestive tract habitats and hydrothermal vents favored

187　by the other sequenced species.

188

189　**Electron donors.** *S. denitrificans* was originally isolated in a chemostat with thiosulfate

190　as electron-donor and nitrate as electron-acceptor (70). However, prior to this study, the

191　pathways and complexes involved were not identified. Neutrophilic sulfur-oxidizing

192　bacteria use two types of sulfur oxidation pathways: one involving a multienzyme

193　complex catalyzing the complete oxidation of reduced sulfur compounds to sulfate (Sox

194　pathway) (15, 28), and another implementing sulfite and elemental sulfur as important

195　intermediates (27, 47, 59). The genome of *S. denitrificans* reveals that the oxidation of

196　reduced sulfur compounds proceeds via the Sox pathway (Fig. 2). Homologs for genes

197   encoding all components that are required for a fully functional complex in vitro, i.e.,

198   SoxB, SoxXA, SoxYZ, and SoxCD (15), could be identified. As in other obligate

199   sequenced autotrophs (41, 57), the *sox* genes in *S. denitrificans* do not occur in one

200   cluster, as in the model organism *Paracoccus pantotrophus* GB17 (15), but in different

201   parts of the genome. *S. denitrificans* has basically two clusters, one containing *soxXYZAB*

202   (*Suden_260-264*) and another one containing *soxZYCD* (*Suden_2057-2060*). SoxZY are

203   known to interact with both SoxAB and SoxCD and their duplication could possibly

204   indicate differential regulation of these two loci. SoxCD has homologies to sulfite

205   dehydrogenase (SorAB), but has been shown to act as a sulfur dehydrogenase (15). In

206   addition, it has recently been shown that organisms that lack *soxCD*, but do have *soxB*,

207   *soxXA*, and *soxYZ* use the Sox system to oxidize thiosulfate to sulfur, which is stored

208   either inside the cell or excreted (21). However, elemental sulfur formation by *S.*

209   *denitrificans* has not been reported. Recently, sulfur oxidation enzymes were also

210   measured in the closely related bacteria *Sulfurimonas autotrophica* and *Sulfurimonas*

211   *paralvinellae* (67). In this case, sulfite dehydrogenase was detected using an assay that

212   would not be expected to measure such activity were these organisms to use the Sox

213   system (C. G. Friedrich, pers. comm.), indicating that other *Sulfurimonas spp.* might

214   either not use the Sox system or use a modified version of it. In this regard it is

215   interesting to note that the SoxC sequence identities of *S. denitrificans* to sequences of

216   those organisms that have a contiguous *sox* gene set are significantly lower (44%) than

217   when SoxC sequences from organisms in which *sox* genes occur in one cluster are

218   compared among themselves (>63%). Both *soxB* and *soxC* genes exhibit highest

219   similarities with genes from *Sulfurovum* sp. NBC37-1 (41), which suggests that both

220   clusters of *sox* genes are not recent additions to this epsilonproteobacterial lineage. In

221   fact, a phylogenetic analysis based on a large number of SoxB sequences from a variety

222   of sulfur-oxidizing bacteria is even suggestive of an origin of the sox system in

223   epsilonproteobacteria (36).

224   Besides the Sox system, *S. denitrificans* also has a gene encoding a

225   sulfide:quinone oxidoreductase (SQR; *Suden_0619*). SQR catalyzes the oxidation of

226   sulfide to elemental sulfur in *Rhodobacter capsulatus* (56), leading to the deposition of

227   sulfur outside the cells. At present its role in *S. denitrificans* is unclear as this species has

228   not been shown to deposit elemental sulfur, though this possibility has not been

229   exhaustively explored with differing cultivation conditions.

230   The genome also provided evidence for the ability to also use H_2 and formate as

231   electron donors (Fig. 2). Based on this information, *S. denitrificans* was successfully

232   cultivated with H_2 as its electron donor and nitrate as the electron acceptor (Sievert and

233   Molyneaux, unpublished data). The *S. denitrificans* genome encodes two Ni-Fe

234   hydrogenase systems: one cytoplasmic, and one membrane-bound hydrogenase complex.

235   The genes encoding the two subunits of the cytoplasmic enzyme (which lack TAT

236   motifs; *Suden_1437-8*) are adjacent to genes encoding the periplasmic hydrogenase (its

237   small subunit has a TAT motif, and a b-type cytochrome subunit would function to

238   anchor it to the membrane and shunt electrons to the quinone pool; *Suden_1434-6*). The

239   small subunit of the cytoplasmic hydrogenase of *S. denitrificans* forms a cluster with

240   sequences from the two deep-sea hydrothermal vent epsilonproteobacteria and *Aquifex*

241   *aeolicus* and is distantly related to H_2-sensing hydrogenases of alphaproteobacteria and

242   cyanobacteria (41). Nakagawa et al. (41) suggest that the cytoplasmic hydrogenase acts

243 as a H_2-sensing hydrogenase in *Sulfurovum* NBC37-1 and *Nitratiruptor* SB155-2.

244 However, an alternative, and in our view more likely, function for the cytoplasmic

245 enzyme as a catalytically active hydrogenase is suggested by the sequence similarity of

246 both subunits to the enzyme from *A. aeolicus*. In *A. aeolicus,* the cytoplasmic

247 hydrogenase can reduce electron acceptors with very negative redox midpoint potentials,

248 and therefore has been suggested to provide electrons with low potential electrons to the

249 reductive citric acid cycle (4), which would circumvent the necessity for reverse electron

250 transport and thus increase its growth efficiency, similar to what has been found in

251 certain Knallgas bacteria using the Calvin cycle for CO_2-fixation. Further experiments are

252 needed to confirm the actual role of the cytoplasmic hydrogenase. Following these

253 hydrogenase genes are several genes encoding hydrogenase-assembly related functions

254 (*Suden_1433-24).*

255   A formate dehydrogenase complex is encoded in an operon similar in gene order to

256 one found in *W. succinogenes* (*Suden_0816-24).* Formate dehydrogenase α subunits

257 contain a selenocysteine residue (26) which is encoded by a stop codon. A putative

258 selenocysteine codon (TGA) followed by a palindromic region was found between two

259 open reading frames (ORFS) that are homologous to the amino and carboxy ends of

260 formate dehydrogenase; accordingly, these ORFS have been combined into a single

261 coding sequence for the α subunit of this enzyme, *Suden_0820*, which includes the

262 molybdopterin-binding and iron-sulfur cluster domains typically found in this subunit, as

263 well as the TAT-pathway signal sequence which would shunt this subunit to the

264 periplasm. The β and γ subunits are encoded by *Suden_0819* and *Suden_0818,*

265 respectively, with the latter having an unprecedented N-terminal addition with two more

266   predicted transmembrane segments (making it six in total). *Suden_0824* encodes a

267   ferredoxin, which may shuttle the electrons from formate oxidation to cellular processes.

268   Formate dehydrogenase maturation is likely facilitated by the products of *Suden_0823,*

269   which encodes a TorD family protein that functions in molybdoprotein formation, and

270   *Suden_0817*, which encodes a FdhD/NarQ family maturation protein. *Suden_0816*

271   encodes a protein belonging to the aminotransferase class V PFAM, as does the SelA

272   protein, which catalyzes a step in selenocysteinyl-tRNA synthesis. Other genes likely to

273   be involved in selenocysteine synthesis are encoded nearby (*Suden_0830:*

274   selenophosphate synthase; *Suden_0831:* L-seryl-tRNA selenium transferase;

275   *Suden_0832:* selenocysteine-specific translation elongation factor SelB). Interestingly,

276   the hydrothermal vent epsilonproteobacteria do not appear to have this complex. Though

277   a homolog to the α subunit of formate dehydrogenase was present in both *Nitratiruptor*

278   SB155-2 and *Sulfurovum* sp. NBC37-1, genes encoding the other subunits were not

279   apparent from BLASTp searches of their genomes. In addition, *S. denitrificans* has a gene

280   encoding the large subunit of a formate dehydrogenase H (*fdhF*, *Suden_1902*), which is

281   most similar to the one in *W. succinogenes* (WS0126). However, its function and

282   substrate are not apparent based on its sequence or genomic context.

283

284   **Electron acceptors.** All genes required for the complete reduction of nitrate to N_2

285   are present (Fig. 2). However, *S. denitrificans* has some notable modifications compared

286   to the canonical denitrification pathway. Similar to *Bradyrhizobium japonicum*, *S.*

287   *denitrificans* appears to have only a periplasmic nitrate reductase (Nap) and not a

288   cytoplasmic membrane-bound nitrate reductase (Nar) (2), which is present in most

289   organisms producing N_2 from nitrate (52). The *nap* gene cluster (*NapABHGFLD;*

290   *Suden_1514-1519, 1521*) has the same arrangement as the one identified in *W.*

291   *succinogenes*, which is unusual in that it lacks a gene encoding the NapC subunit (64).

292   Possibly, *S. denitrificans* is able to denitrify under microaerobic conditions, as the Nap

293   enzyme has been implicated in aerobic denitrification (38). Nitrite reduction to nitric

294   oxide is likely catalyzed by a cytochrome cd_1-dependent nitrite reductase (*nirS, nirF*;

295   *Suden_1985, 1988*), whose genes are present in a gene cluster (*Suden_1976-1989*) that

296   also includes siroheme synthesis genes and two genes annotated as *norCB* (*Suden_1983-*

297   *1984*).

298   The *norCB* genes, which encode the small and large subunits of nitrous oxide-

299   forming nitric oxide reductase (cNOR), a member of the heme-copper oxidase (HCO)

300   superfamily (16), are usually clustered together with additional genes required for

301   enzyme assembly and activation (76). These latter genes, *norD* and *norQ*, are missing

302   from the *S. denitrificans* genome. While there is evidence for functional cNOR in

303   bacteria that lack the *norQ/cbbQ* gene, there are presently no experimental reports that

304   demonstrate functionality of cNOR in bacteria that also lack the *norD* gene. Attempts to

305   test this for *Hydrogenobacter thermophilus* using *Pseudomonas aeruginosa* or

306   *Escherichia coli* as an expression hosts were inconclusive (66). The *S. denitrificans*

307   *norCB* genes are closest related by sequence similarity to the *norCB* genes in the

308   genomes of *H. thermophilus, Methylococcus capsulatus* as well as *Sulfurovum* sp.

309   NBC37-1 and *Nitratiruptor* sp. SB155-2, the latter of which are two newly sequenced

310   marine epsilonproteobacteria (41). All of these genomes also lack the *norD* gene.

311   Even though the functionality of cNOR is questionable, *S. denitrificans* is a

312   complete denitrifier and must be able to reduce NO. Attempts to find other inventory

313   implicated in NO reduction were successful and yielded additional candidate systems.

314   One of them, NADH:flavorubredoxin-NOR, also known as the NorVW complex (17,

315   19), was also not complete and thus likely non-functional because a NorW-encoding gene

316   was not identified. Interestingly, the genome encodes NorV in the unusual form of two

317   genes: one encoding a rubredoxin (*Suden_1582*), which is succeeded by a flavodoxin

318   gene (*Suden_1581*). Although both cNOR and NADH:flavorubredoxin-NOR may not

319   have catalytic activity, it is possible that their NO-binding capacity has a function in NO

320   sequestration and detoxification.

321   Interestingly, the *S. denitrificans* genome also encodes a previously unidentified

322   member of the HCO superfamily that is also a candidate for catalyzing nitric oxide

323   reduction. These HCO genes follow a set of pseudogenes normally involved in nitrate

324   reduction (*Suden_0100-0102*). Based on structural modeling and genome analysis it is

325   expected that this new HCO family is a novel non-electrogenic quinone-oxidizing nitric

326   oxide reductase, gNOR (J. Hemp, M. G. Klotz, L. Y. Stein, R. B. Gennis, unpubl. data).

327   The gNOR family, encoded by the *norGHJ* genes (*Suden_0103-0105*), is unique within

328   the HCO superfamily in that it exhibits a novel active-site metal ligation, with one of the

329   three conserved histidine ligands being replaced with an aspartate. This ligation pattern

330   strongly suggests that the active-site metal is an iron. Structural modeling of members of

331   the gNOR family has identified three conserved acidic residues, which form a charged

332   pocket within the active site, a feature shared with the cNOR family (49). Besides *S.*

333   *denitrificans,* gNOR also appears to be present in *Sulfurovum* sp. NBC37-1 and

334   *Persephonella marina* strain EX-H1, whereas it is missing in *Nitratiruptor* sp. SB155-2.

335   Since phylogenetic analysis demonstrates that nitric oxide reductase activity has evolved

336   multiple times independently within the heme-copper superfamily, these shared features

337   between the distantly related gNOR and cNOR families are interesting examples of

338   convergent evolution (J. Hemp, M. G. Klotz, L. Y. Stein, R. B. Gennis, unpubl. data).

339   Nitrous oxide reduction is carried out by nitrous oxide reductase encoded by an

340   unusual *nos* gene cluster (*Suden_1298-1308)* similar to one previously identified in *W.*

341   *succinogenes* (63). As in *Wolinella*, the NosZ in *S. denitrificans* (*Suden_1298)* contains a

342   C-terminal extension of about 200 residues that carries a monoheme cytochrome c

343   binding motif (CXGCH), suggesting it, too, functions as a cytochrome c nitrous oxide

344   reductase (cNOS; Fig. 2). This feature is also shared by NosZ of *Sulfurovum* sp. NBC37-

345   1 and *Nitratiruptor* sp. SB155-2, which form a cluster with NosZ from *S. denitrificans*

346   (Fig. S3)*,* possibly allowing the design of primers to screen for the presence of

347   denitrifying epsilonproteobacteria in the environment. It has been hypothesized that the

348   *nos* gene cluster in *W. succinogenes* codes for proteins involved in an electron transport

349   chain from menaquinol to cytochrome c nitrous oxide reductase (63), and it is likely that

350   the same holds true for *S. denitrificans,* as well as *Sulfurovum* sp. NBC37-1 and

351   *Nitratiruptor* sp. SB155-2. In addition, *S. denitrificans* has an almost identical copy of

352   *nosZ* next to a c553-type monoheme cytochrome c (*Suden_1770, 1769*), but its function

353   is at this point unknown. Interestingly, *S. denitrificans* also has a gene coding for a large

354   subunit of a ferredoxin-nitrite reductase (*nirB*; *Suden_1241*), which could be involved in

355   nitrite assimilation or detoxification (8). However, no gene coding for the small subunit

356   was identified, raising questions about its function.

357   Additional electron acceptors are suggested by this organism's gene complement

358   (Fig. 2). Like some other delta- and epsilonprotreobacteria, the *S. denitrificans* genome

359   contains a cluster of four genes (*Suden_0081-0084*) that encode the FixNOQP proteins,

360   which constitute a proton-pumping cbb_3- type cytochrome c oxidase., suggesting an

361   ability to use oxygen as a terminal electron acceptor. This is somewhat unexpected, since

362   *S. denitrificans* was originally described as an obligate denitrifier and is quite sensitive to

363   oxygen (S. Sievert, unpubl. data). *Cbb3*-type cytochrome c oxidase complexes have

364   extremely high affinities for oxygen (48), which might allow this organism to use oxygen

365   as an electron acceptor under extremely low oxygen tensions, or alternatively, to

366   scavenge oxygen to prevent poisoning. Another possibility might be the involvement of

367   *cbb3*-type cytochrome c oxidase in the catalytic reduction of NO (14). It is interesting,

368   that, in line with other epsilonproteobacteria, the genome does not contain genes

369   encoding FixGHIJ, which are present in all other bacteria that express a cbb_3-type

370   cytochrome c oxidase and are involved in assembly and maturation of the cbb_3-type

371   cytochrome c oxidase complex (33, 44). The reasons for the absence of *fixGHIJ* in

372   epsilonproteobacteria remain unkown, and it needs to be tested experimentally what the

373   actual role of cbb_3-type cytochrome c oxidase in *S. denitrificans* is.

374   Additionally, *S. denitrificans, Nitratiruptor* sp. SB155-2, and *Sulfurovum* sp.

375   NBC37-1 may be able to use sulfur compounds as terminal electron acceptors. Genes

376   *Suden_0498-0500* encode an Fe/S-protein, a NrfD-related membrane anchor (8 helices),

377   and an unusually long molybdopterin-containing oxidoreductase that contains a twin-

378   arginine translocation pathway signal, respectively; homologs of these three genes are

379   also encoded adjacent to each other in the two hydrothermal vent epsilonproteobacteria.

380   Together these proteins could form a periplasm-facing membrane-bound complex that is

381 most likely involved in the reduction of sulfur compounds like elemental sulfur,

382 polysulfide, thiosulfate or tetrathionate. However, only a limited number of these

383 molybdopterin-containing oxidoreductases have been characterized and it is presently not

384 possible to infer substrate specificity from the primary structure. Further experiments are

385 needed to confirm the substrate used by this complex.

386

387 **Carbon fixation and central carbon metabolism.** Genes encoding the enzymes

388 of the reductive citric acid cycle are apparent (Fig. 3), which is consistent with prior

389 biochemical and genetic analyses of this organism (24). Of particular note are genes

390 encoding the enzymes necessary for the cycle to operate in the reductive direction:

391 pyruvate:acceptor oxidoreductase (*Suden_0096 – 0099,* based on similarity to

392 biochemically characterized orthologs in *H. pylori* (23)), pyruvate carboxylase

393 (*Suden_0622, Suden_1259*, based on biochemically characterized orthologs from *C.*

394 *jejuni* (72), but see below), 2-oxoglutarate:ferredoxin oxidoreductase (*Suden_1052 – 55,*

395 as for pyruvate:acceptor oxidoreductase, based on (23)), and ATP-dependent citrate lyase

396 (*Suden_0570, Suden_0571*)(24). Also noteworthy is the presence of two copies of

397 succinate dehydrogenase/fumarate reductase, one of which has a subunit that would

398 anchor it to the membrane (*Suden_1028 - 1030*), while the other lacks this subunit and

399 may be cytoplasmic (*Suden_0037, Suden_ 0038*). The membrane-bound form is unusual

400 in that it has a cysteine-rich, type-E membrane anchor. It is similar to SdhABE from *W.*

401 *succinogenes* which has been characterized as a membrane-bound fumarate-reducing

402 complex with subunits SdhAB facing the periplasm. In contrast to *W. succinogenes*, the

403 *S. denitrificans* gene encoding subunit A is about 43 residues shorter and lacks the TAT

404   signal peptide present in its *W. succinogenes* homolog, something that appears to be

405   shared with *Sulfurovum* sp. NBC37-1 and *Nitratiruptor* sp. SB155-2, both of which are

406   also chemolithoautotrophic epsilonproteobacteria using the reductive citric acid cycle for

407   carbon fixation. Thus, it appears that in these organisms the membrane-bound fumarate-

408   reducing complex faces into the cytoplasm. At present the exact function of the two

409   fumarate reductases is unknown, although it is likely that the membrane-bound one, due

410   to its potential for additional energy generation, might be involved in the reductive citric

411   acid cycle for autotrophic carbon fixation. The intriguing possibility that *S. denitrificans*

412   might also be able to carry out fumarate respiration has to await further experimentation.

413   However, *S. denitrificans* does not contain a *frdCAB* operon typical for

414   menaquinol:fumarate reductase sustaining fumarate respiration in other

415   epsilonproteobacteria.

416   The acetyl-CoA and oxaloacetate produced by the reductive citric acid cycle

417   could be funneled to central carbon metabolism: acetyl-CoA could be converted to

418   pyruvate via pyruvate:acceptor oxidoreductase (see above) and oxaloacetate could be

419   used to form phosphoenolpyruvate via phosphoenolpyruvate carboxykinase

420   (*Suden_1696*). Acetyl-CoA could also be directed to fatty acid synthesis (acetyl-CoA

421   carboxylase; *Suden_1174; Suden_1608*). Genes are present that suggest an ability to

422   supplement autotrophic growth with acetate assimilation in all three sulfur-oxidizing

423   epsilonproteobacteria. Two possible systems for converting acetate to acetyl-CoA are

424   present: acetyl-CoA ligase (*Suden_1451*), as well as phosphate acetyltransferase

425   (*Suden_0055*) and acetate kinase (*Suden_0056*), and are also present in *Sulfurovum* sp.

426   NBC37-1, while acetate kinase is absent in *Nitratiruptor* sp. SB155-2. Perhaps the two

427   systems have different affinities for acetate, as has been demonstrated in methanogens

428   (60), and are differentially expressed depending on environmental concentrations of this

429   organic acid.

430   In order for *S. denitrificans* to grow autotrophically using the reductive citric acid

431   cycle, there must be a means of carboxylating pyruvate to form oxaloacetate. In some

432   organisms this is accomplished by the tandem activities of phosphoenolpyruvate

433   synthetase and phosphoenolpyruvate carboxylase, while others use pyruvate carboxylase

434   (55). In contrast to the two autotrophic epsilonproteobacteria *Sulfurovum* NBC37-1 and

435   *Nitratiruptor* SB155-2, which have genes encoding phosphoenolpyruvate synthase as

436   well as pyruvate kinase (41), the latter of which is usually involved in ATP synthesis

437   during glycolysis, *S. denitrificans* does not appear to have any genes that might encode an

438   enzyme that could interconvert phosphoenolpyruvate and pyruvate. Instead, it may use

439   pyruvate carboxylase, as genes encoding both the biotin carboxylase subunit

440   (*Suden_0622*) and biotin carboxyl carrier subunit (*Suden_1259)* of this enzyme are

441   present. The biotin carboxyl carrier subunit gene (*Suden_1259*) occurs in an apparent

442   operon with other genes homologous to the subunits of sodium-transporting oxaloacetate

443   decarboxylase (*Suden_1258 – 60*). *Suden_1259*, which encodes the α subunit of this

444   complex, has a high level of sequence similarity with pyruvate carboxylase genes from

445   various *Campylobacter* species (including one from *C. jejuni* which has been

446   biochemically characterized) (72), while the β and γ subunits (*Suden_1258* and *1260*),

447   which are absent from the heterotrophic epsilonproteobacteria, are similar to those found

448   from *Sulfurovum* sp. NBC37-1, *Nitratiruptor* sp. SB155-2, many gammaproteobacteria,

449   many *Chlorobia*, and *Desulfotalea psychrophila*, a deltaproteobacterial sulfate reducer.

450 Heterotrophic organisms that have this complex ferment citrate. After cleaving citrate to

451 acetate and oxaloacetate, they use the oxaloacetate decarboxylase complex to couple the

452 exothermic decarboxylation of this organic acid to the extrusion of sodium ions. For these

453 organisms, the other genes necessary for citrate fermentation (e.g., citrate transporter) are

454 encoded nearby (9). This is not the case in *S. denitrificans*. An alternative function for

455 *Suden_1259* is suggested by phylogenetic analysis, which places it within a clade with

456 the biochemically characterized pyruvate carboxylase from *C. jejuni* (Fig. 4) and separate

457 from biochemically characterized oxaloacetate decarboxylase genes from *Klebsiella*

458 *pneumoniae* and *Vibrio cholerae* (5). Other members of this clade include the genes from

459 the *Chlorobia, Sulfurovum* sp. NBC37-1, and *Nitratiruptor* sp. SB155-2, which also use

460 the reductive citric acid cycle for carbon fixation and are not known to ferment citrate. It

461 is possible to operate the oxaloacetate decarboxylase complex as a pyruvate carboxylase

462 by imposing a sodium gradient across the membrane (10). It is tempting to speculate that

463 in the autotrophic epsilonproteobacteria, the *Chlorobia*, and possibly *D. psychrophila,*

464 this complex functions as a pyruvate carboxylase. Interestingly, the sequenced

465 autotrophic epsilonproteobacteria and *Desulfotalea psychrophila* are all marine

466 organisms, and the *Chlorobia* evolved in the marine environment (1). Only five of the ten

467 sequenced *Chlorobia* have the sodium-transporting oxaloacetate decarboxylase/pyruvate

468 carboxylase complex, while the other five, including *C. tepidum,* have the alpha subunit

469 (on which Fig. 4 is based), but not the three-subunit pump. With the exception of *C.*

470 *limicola*, which was isolated from a mineral hot spring, all *Chlorobia* containing the

471 sodium pump have a requirement for sodium. Thus, it is likely that these organisms have

472 found a way to couple pyruvate carboxylation, which is energetically unfavorable, to a

473   sodium gradient, something that was previously proposed, but never shown in an

474   organism (10). Whether the complex encoded by *Suden_1258-1260* functions as a

475   pyruvate carboxylase or an oxaloacetate decarboxylase is a key point begging

476   clarification, which will be nontrivial, given that a genetic system has not been developed

477   in this organism.

478   Carbon fixed by the reductive citric acid cycle can be shunted through

479   gluconeogenesis, as all genes necessary for this pathway are present. The presence of

480   genes encoding citrate synthase (*Suden_2100*) and ATP-dependent (irreversible)

481   phosphofructokinase (*Suden_0549*) are enigmatic as their roles are unclear in this

482   obligate autotroph.

483   Genes are apparent whose products could utilize the carbon skeletons synthesized

484   by central carbon metabolism for ammonia assimilation (see below) and amino acid,

485   nucleotide, fatty acid, and phospholipids synthesis. Cysteine biosynthesis is notable, in

486   that the reduction of sulfate proceeds via adenosine 5'-phosphosulfate (APS) rather than

487   3'-phosphoadenylylsulfate (PAPS) in a pathway that was until recently only known from

488   plants (42). Genes encoding assimilatory sulfate reduction co-occur in an apparent operon

489   (*Suden_0154-0160)*. Most likely this operon-like structure is turned on or off depending

490   on whether *S. denitrificans* is inhabiting an environment with a high concentration of

491   reduced inorganic sulfur compounds. Interestingly, sulfate assimilation in *Sulfurovum* sp.

492   NBC37-1 and *Nitratiruptor* SB155-2 appears to proceed via PAPS.

493

494   **Nitrogen assimilation.** The *S. denitrificans* genome contains *nirC* (*focA*) and

495   *nirB* genes encoding the formate-nitrite transporter (*Suden_0716*) and the large subunit of

496   NAD(P)H-dependent ammonia-forming siroheme nitrite reductase (*Suden_1241*),

497   respectively, along with the inventory for siroheme synthesis (*Suden_1977*, *cysG*,

498   siroheme synthase; *Suden_1988*, *cobA-cysG*, uroporphyrinogen III methylase); however,

499   it lacks the *nirD* gene, which encodes the small subunit of siroheme nitrite

500   reductase. Because the genome also lacks *nrfHA* genes, which encode respiratory nitrite

501   ammonification capacity in many delta- and epsilonproteobacteria (62), it appears that

502   *Sulfurimonas* is solely dependent on ammonia uptake from the environment. The genome

503   contains, indeed, two genes encoding different ammonia permeases (noted above) (22,

504   43), one AmtB-like (*Suden_0641*) and one Rhesus factor-like (*Suden_0643*) permease,

505   which are clustered together with the gene encoding nitrogen regulatory protein PII

506   (*glnK*, *Suden_0642*). Whereas AmtB proteins function as ammonia gas uptake channels,

507   the substrate for Rh-like protein channels is still debated and includes ammonia as well as

508   CO_2 (43). The genome contains also all the additional genes needed for 2-oxoglutarate-

509   sensing and regulation of nitrogen assimilation (reviewed in (34)).

510

511   **Chemotaxis and other regulatory and signaling proteins.** Close relatives of *S.*

512   *denitrificans* are motile , while this particular strain is nonmotile, probably due to the

513   interruption of a flagellar biosynthetic operon by a transposon (see Genome Structure,

514   above). Based on the presence of all of the genes necessary to encode the flagellar

515   apparatus, none of which display any evidence of degeneration, an abundant sensory

516   apparatus necessary to detect the presence of chemoattractants or repellants, and

517   communicate this information to the flagellar motor, as well as the sequence identity of

518   this transposon with a duplicate in the genome (see above), it is likely that nonmotility is

519 a recently acquired property. Interestingly, many of the genes encoding the chemotaxis

520 components are in a large cluster with multiple kinases and response regulators (Fig. 5),

521 as in *Nitratiruptor* sp. SB155-2 (41), suggesting interconnectivity between the

522 chemotaxis and other signal transduction systems. Perhaps the original enrichment and

523 isolation procedure for this strain (in a chemostat) might have selected for a non-motile

524 strain.

525   The *S. denitrificans* genome encodes a relative abundance of signaling proteins.

526 Particularly well-represented in these genomes are genes encoding proteins with EAL

527 and GGDEF domains (based on hits to PFAMS, 16 and 38 genes, respectively), which

528 likely function in the synthesis and hydrolysis of the intracellular signaling compound

529 cyclic diguanylate (53). Further, six proteins with PAS/PAC-domains are encoded, which

530 may function as redox sensors (75). The genomic repertoire of signaling and regulatory

531 genes was compared with two other free-living, obligate chemolithoautotrophs for which

532 these data are available (Table 2). Some features are similar to *Thiomicrospira*

533 *crunogena*: both of these species have a relative abundance of signal transduction

534 proteins compared to *Nitrosococcus oceani* (31), which may be a response to more

535 spatially (sediments; *S. denitrificans)* or temporally (hydrothermal vents; *T. crunogena*)

536 heterogeneous habitats, compared to the open ocean (*N. oceani*). Both have a large

537 number of genes encoding methyl-accepting chemotaxis proteins (MCPs; Table 2).

538 Unlike *T. crunogena,* none of the MCPs from *S. denitrificans* are predicted to have

539 PAS/PAC domains that could bind redox-sensitive cofactors (57), but a gene encoding a

540 protein with a PAS/PAC domain is present in the large cluster of chemotaxis genes

541 described above (Fig. 5), suggesting potential communication between sensing cellular or

542 environmental redox conditions and the chemotactic apparatus. Another similarity

543 between *T. crunogena* and *S. denitrificans* is an abundance of genes predicted to be

544 involved with cyclic nucleotide signal transduction, and many of these are predicted to

545 have EAL and/or GGDEF domains, indicating a role for cyclic diguanylate in

546 intracellular signaling in this organism. Many of these predicted proteins also have

547 PAS/PAC domains, as in *T. crunogena* (57).

548   Unique among the three species compared here, *S. denitrificans* has a relative

549 abundance of signal transduction histidine kinases and an expanded complement of

550 winged helix family two component transcriptional regulators (Table 2). Both *T.*

551 *crunogena* and *N. oceani* can use a rather limited variety of electron donors and

552 acceptors, compared to *S. denitrificans*. Perhaps this expansion in histidine kinases and

553 transcriptional regulators coordinate expression of the complexes necessary for the

554 oxidation of multiple electron donors (e.g., H_2, reduced sulfur compounds, formate), and

555 reduction of multiple electron acceptors (O_2 and NO_x).

556

557   **Oxidative stress.** *S. denitrificans* has reasonably elaborate defenses against

558 oxidative stress, on par with what has been observed for pathogenic

559 epsilonproteobacterium *H. pylori* (73). Initially this was considered surprising to us,

560 because this nonpathogenic species does not have to elude the oxidative arsenal of a host

561 immune system, nor, as a microaerophile capable of growth via denitrification, does it

562 grow in the presence of high concentrations of oxygen. However, given the presence of

563 enzymes with labile iron-sulfur clusters with irreplaceable roles in central carbon

564 metabolism in this organism (e.g., pyruvate:acceptor oxidoreductase; *Suden_0096 –*

565 *0099;* 2-oxoglutarate:acceptor oxidoreductase; *Suden_1053 - 1055*), perhaps added

566 defenses are a necessary part of survival.

567 Several genes are present whose products could prevent the buildup of

568 intracellular Fe^{2+} that can spur hydroxyl radical generation via the Fenton reaction. A

569 gene is present that encodes the Fur protein (*Suden_1272*), which regulates iron uptake

570 (14). A gene encoding iron-binding ferretin (*Suden_1760*) is also present.

571 Many enzymes to detoxify reactive oxygen and nitrogen species are encoded in

572 this genome. An iron/manganese superoxide dismutase (*Suden_1129*) is present in this

573 species, but not the other sulfur-oxidizing epsilonproteobacteria (41). This enzyme and

574 ruberythrin (*Suden_0739*) could convert superoxide to hydrogen peroxide, and this

575 superoxide could in turn be dispatched by catalase (*Suden_1323*), peroxiredoxins

576 *(Suden_0132, Suden_0630, Suden_1778, Suden_1803)*, or cytochrome c peroxidase

577 *(Suden_0214; Suden_1585).* Peroxiredoxins, particularly alkylhydroperoxide reductases,

578 might be specifically targeted towards low levels of H_2O_2 or organic peroxides (58, 73).

579 To dispense with nitric oxide that escapes the periplasmic and membrane-associated

580 respiratory nitrogen reduction complexes, a truncated hemoglobin gene is present

581 *(Suden_0993)*, which may convert nitric oxide to nitrate (46). Thioredoxins

582 (*Suden_0342; Suden_0501; Suden_1867; Suden_2099)* and thioredoxin reductase

583 *(Suden_1869)* could funnel electrons to these oxidative stress proteins, as glutathione

584 does not appear to play this role in this organism, since genes encoding glutathione

585 synthetase or gamma-glutamate-cysteine ligase are absent.

586 Genes are apparent whose products could enable a cell to cope with damage

587 inflicted by any reactive oxygen or nitrogen species (ROS and RNS) that escape cellular

588 defenses. Endonuclease III (*Suden_0516*) and IV *(Suden_1835)* and MutS *(Suden_0755)*

589 could repair oxidative DNA damage. Methionine sulphoxide reductase *(Suden_0012)* and

590 alkylhydroxide reductase *(Suden_1778)* could contend with any methionine residues or

591 lipids that had been oxidized by interaction with ROS or RNS, while nitroreductases

592 (*Suden_0519; Suden_1158)* could prevent oxidized cellular nitrogenous and other

593 compounds from generating peroxide.

594

595   **Conclusions**. *S. denitrificans* has several unique features which differentiate it

596 from the other epsilonproteobacteria that have been sequenced to date. It has a larger

597 genome than most of the others, which likely provides the sensory, regulatory, and

598 metabolic versatility necessary for survival in a habitat more heterogeneous than found in

599 a metazoan host. For example, the numerous genes whose products have redox sensory

600 domains likely function to position these cells in the redoxcline to enable them to obtain

601 the electron donors and acceptors needed for growth. Furthermore, these cells are quite

602 versatile with respect to electron donors and acceptors, as the genome data suggest a

603 capability of using donors and acceptors beyond those based on cultivation studies.

604 Although *S. denitrificans* has been isolated from coastal marine sediments, its genome

605 shares many features with two recently described autotrophic deep-sea hydrothermal vent

606 epsilonproteobacteria, including the potential to utilize a variety of redox substrates

607 (hydrogen gas, reduced sulfur compounds, oxygen, and nitrate), its responses to oxidative

608 stress and high metal content, and a genome size intermediate between the two. This

609 suggests that while these habitats appear at first strikingly different, they require similar

610 adaptations on the scale of the microbes. Several additional features, which are absent in

611   their hydrothermal vent relatives, are present in *S. denitrificans* that may be particularly

612   valuable in the sediment habitat. Their formate dehydrogenase complex would enable *S.*

613   *denitrificans* to utilize a major by-product of fermentation that would co-occur with it

614   should sediment organic carbon loads be high. The presence of the additional oxidative

615   stress protein (superoxide dismutase) may enable *S. denitrificans* to cope with diurnal

616   shifts in sediment oxygen concentration, and several additional RND-family efflux

617   pumps relative to hydrothermal vent epsilonproteobacteria suggest that survival in marine

618   sediments requires a degree of versatility and defense against environmental insult

619   beyond what is necessary at moderate temperatures at hydrothermal vents.

620

621

622

623

624   **ACKNOWLEDGMENTS**

625

REFERENCES

638

639   1.   **Alexander, B., J. H. Andersen, R. P. Cox, and J. F. Imhoff.** 2002. Phylogeny

640   of green sulfur bacteria on the basis of gene sequences of 16S rRNA and of the

641   Fenna-Matthews-Olson protein. Arch Microbiol **178:**131-140.

642   2.   **Bedmar, E. J., E. F. Robles, and M. J. Delgado.** 2004. The complete

643   denitrification pathway of the symbiotic, nitrogen-fixing bacterium

644   *Bradyrhizobium japonicum*. Biochem Soc Trans **33:**141-144.

645   3.   **Brettar, I., M. Labrenz, S. Flavier, J. Botel, H. Kuosa, R. Christen, and M. G.**

646   **Hofle.** 2006. Identification of a *Thiomicrospira denitrificans*-Like

647   Epsilonproteobacterium as a Catalyst for Autotrophic Denitrification in the

648   Central Baltic Sea. Appl Environ Microbiol **72:**1364-1372.

649   4.   **Brugna-Guiral, M., P. Tron, W. Nitschke, K. O. Stetter, B. Burlat, B.**

650   **Guigliarelli, M. Bruschi, and M. T. Giudici-Orticoni.** 2003. [NiFe]

651   hydrogenases from the hyperthermophilic bacterium *Aquifex aeolicus*: properties,

652   function, and phylogenetics. Extremophiles **7:**145-157.

653   5.   **Buckel, W.** 2001. Sodium ion-translocating decarboxylases. Biochim Biophys

654   Acta **1505:**15-27.

655   6.   **Campbell, B. J., A. S. Engel, M. L. Porter, and K. Takai.** 2006. The versatile

656   epsilon-proteobacteria: key players in sulphidic habitats. Nature Rev Microbiol

657   **4:**458-468.

658   7.   **Chain, P., J. Lamerdin, F. Larimer, W. Regala, V. Lao, M. Land, L. Hauser,**

659   **A. Hooper, M. Klotz, J. Norton, L. Sayavedra-Soto, D. Arciero, N. Hommes,**

660   **M. Whittaker, and D. Arp.** 2003. Complete genome sequence of the ammonia-

661   oxidizing bacterium and obligate chemolithoautotroph *Nitrosomonas europaea*. J

662   Bacteriol **185:**2759-2773.

663   8.   **Cole, J.** 1996. Nitrate reduction to ammonia by enteric bacteria: redundancy, or a

664   strategy for survival during oxygen starvation? . FEMS Microbiol Let **136:**1-11.

665   9.   **Dimroth, P.** 2004. Molecular basis for bacterial growth on citrate or malonate. *In*

666   A. Bock, R. Curtiss, J. B. Kaper, F. C. Neidhardt, T. Nystrom, K. E. Rudd, and C.

667   L. Squires (ed.), EcoSal--*Escherichia coli* and *Salmonella:* cellular and molecular

668   biology, vol. 2004. ASM Press, Washington, DC.

669 10. **Dimroth, P., and W. Hilpert.** 1984. Carboxylation of pyruvate and acetyl

670   coenzyme A by reversal of the Na+ pumps oxaloacetate decarboxylase and

671   methylmalonyl-CoA decarboxylase. Biochem **23:**5360-5366.

672 11. **Engel, A. S., N. Lee, M. L. Porter, L. A. Stren, P. C. Bennett, and M.**

673   **Wagner.** 2003. Filamentous "Epsilonproteobacteria" Dominate Microbial Mats

674   from Sulfidic Cave Springs. Appl Environ Microbiol **69:**5503-5511.

675 12. **Ewing, B., and P. Green.** 1998. Basecalling of automated sequencer traces using

676   phred. II. Error probabilities. Genome Res. **8:**186–194.

677 13. **Ewing, B. L., M. Hillier, P. Wendl, and P. Green.** 1998. Basecalling of

678   automated sequencer traces using phred. I. Accuracy assessment. Genome

679   Res.**8:**175–185.

680 14. **Fang, F. C.** 2004. Antimicrobial reactive oxygen and nitrogen species: Concepts

681   and controversies. Nature Microbiol Rev **2:**820-832.

682   15.   **Friedrich, C. G., D. Rother, F. Bardischewsky, A. Quentmeier, and J.**

683   **Fischer.** 2001. Oxidation of Reduced Inorganic Sulfur Compounds by Bacteria:

684   Emergence of a Common Mechanism? Appl. Environ. Microbiol. **67:**2873-2882.

685   16.   **Garcia-Horsman, J. A., B. Barquera, J. Rumbley, J. Ma, and R. B. Gennis.**

686   1994. The superfamily of heme-copper respiratory oxidases. J Bacteriol

687   **176:**5587-5600.

688   17.   **Gardiner, A. M., and P. R. Gardiner.** 2002. Flavohemoglobin detoxifies nitric

689   oxide in aerobic, but not anaerobic, *Escherichia coli.* Evidence for a novel

690   inducible anaerobic nitric oxide-scavenging activity. J Biol Chem **277:**8166-8171.

691   18.   **Gevertz, D., A. J. Telang, G. Voordouw, and G. E. Jenneman.** 2000. Isolation

692   and Characterization of Strains CVO and FWKO B, Two Novel Nitrate-

693   Reducing, Sulfide-Oxidizing Bacteria Isolated from Oil Field Brine. Appl.

694   Environ. Microbiol. **66:**2491-2501.

695   19.   **Gomes, C. M., A. Giuffre, E. Forte, J. B. Vicente, L. M. Saraiva, M. Brunori,**

696   **and M. Teixeira.** 2002. A Novel Type of Nitric-oxide Reductase - *Escherichia*

697   *coli* flavorubredoxin. J Biol Chem **277:**25273-25276.

698   20.   **Gordon, D., C. Abajian, and P. Green.** 1998. Consed: a graphical tool for

699   sequence finishing. Genome Res. **8:**195–202

700   21.   **Hensen, D., D. Sperling, H. G. Trüper, D. C. Brune, and C. Dahl.** 2006.

701   Thiosulphate oxidation in the phototrophic sulphur bacterium *Allochromatium*

702   *vinosum*. Mol Microbiol **62:**794-810.

703   22.   **Huang, C. H., and J. Peng.** 2005. Evolutionary conservation and diversification

704   of Rh family genes and proteins. Proc Natl Acad Sci **102:**15112-17.

705   23.   **Hughes, N. J., C. L. Clayton, P. A. Chalk, and D. J. Kelly.** 1998. *Helicobacter*

706   *pylori porCDAB* and *oorDABC* Genes Encode Distinct Pyruvate:Flavodoxin and

707   2-Oxoglutarate:Acceptor Oxidoreductases Which Mediate Electron Transport to

708   NADP. J Bacteriol **180:**1119-1128.

709   24.   **Hügler, M., C. O. Wirsen, G. Fuchs, C. D. Taylor, and S. M. Sievert.** 2005.

710   Evidence for autotrophic CO2 fixation via the reductive tricarboxylic acid cycle

711   by members of the epsilon subdivision of proteobacteria. J Bacteriol **187:**3020-

712   3027.

713   25.   **Inagaki, F., K. Takai, H. Kobayashi, K. H. Nealson, and K. Horikoshi.** 2003.

714   *Sulfurimonas autotrophica* gen. nov., sp. nov., a novel sulfur-oxidizing epsilon-

715   proteobacterium isolated from hydrothermal sediments in the Mid-Okinawa

716   Trough. Int J Syst Evol Microbiol **53:**1801-1805.

717   26.   **Jormakka, M., B. Byrne, and S. Iwata.** 2003. Formate dehydrogenase - a

718   versatile enzyme in changing environments. Curr Opin Struct Biol **13:**418-423.

719   27.   **Kappler, U., and C. Dahl.** 2001. Enzymology and molecular biology of

720   prokaryotic sulfite oxidation. FEMS Microbiol Let **203:**1-9.

721   28.   **Kelly, D. P., J. K. Shergill, W. P. Lu, and A. P. Wood.** 1997. Oxidative

722   metabolism of inorganic sulfur compounds by bacteria. Antonie Van

723   Leeuwenhoek Int J Gen Molec Microbiol **71:**95-107.

724   29.   **Kholodii, G. Y., S. Z. Mindlin, I. A. Bass, O. V. Yurieva, S. V. Minakhina,**

725   **and V. G. Nikiforov.** 1995. Four genes, two ends, and a res region are involved

726   in transposition of Tn5053: a paradigm for a novel family of transposons carrying

727   either a mer operon or an integron. Mol Microbiol **17:**1189-1200.

728   30.   **Klappenbach, J. A., J. M. Dunbar, and T. M. Schmidt.** 2000. rRNA operon

729   copy number reflects ecological strategies of bacteria. Appl Environ Microbiol

730   **66:**1328-1333.

731   31.   **Klotz, M. G., D. J. Arp, P. S. G. Chain, A. F. El-Sheikh, L. J. Hauser, N. G.**

732   **Hommes, F. W. Larimer, S. A. Malfatti, J. M. Norton, A. T. Poret-Peterson,**

733   **L. M. Vergez, and B. B. Ward.** 2006. Complete genome sequence of the marine,

734   chemolithoautotrophic, ammonia-oxidizing bacterium *Nitrosococcus oceani*

735   ATCC 19707. Appl Environ Microbiol **72:**6299-6315.

736   32.   **Kodama, Y., and K. Watanabe.** 2003. Isolation and characterization of a sulfur-

737   oxidizing chemolithotroph growing on crude oil under anaerobic conditions. Appl

738   Environ Microbiol **69:**107-112.

739   33.   **Kulajta, C., J. O. Thumfart, S. Haid, F. Daldal, and H. G. Koch.** 2006. Multi-

740   step assembly pathway of the cbb3-type cytochrome c oxidase complex. J Mol

741   Biol **355:**989-1004.

742   34.   **Leigh, J. A., and J. A. Dodsworth.** 2007. Nitrogen regulation in bacteria and

743   archaea. Annu Rev Microbiol **61:**349-377.

744   35.   **Mehlman, I. J., and A. Romero.** 1982. Improved Growth-Medium for

745   Campylobacter Species. Appl Environ Microbiol **43:**615-618.

746   36.   **Meyer, B., J. F. Imhoff, and J. Kuever.** 2007. Molecular analysis of the

747   distribution and phylogeny of the *soxB* gene among sulfur-oxidizing bacteria --

748   evolution of the Sox sulfur oxidation enzyme system. Environ Microbiol **9:**2957-

749   2977.

750   37.   **Meyer, F., A. Goesmann, A. C. McHardy, D. Bartels, T. Bekel, J. Clausen, J.**

751   **Kalinowski, B. Linke, O. Rupp, R. Giegerich, and A. Puhler.** 2003. GenDB--

752   An open source genome annotation system for prokaryotic genomes. Nucleic

753   Acids Res **31:**2187-2195.

754   38.   **Moreno-Vivian, C., P. Cabello, M. Martinez-Luque, R. Blasco, and F.**

755   **Castillo.** 1999. Prokaryotic nitrate reduction: Molecular properties and functional

756   distinction among bacterial nitrate reductases. J Bacteriol **181:**6573-6584.

757   39.   **Moyer, C. L., F. C. Dobbs, and D. M. Karl.** 1995. Phylogenetic diversity of the

758   bacterial community from a microbial mat at an active, hydrothermal vent system,

759   Loihi seamount, Hawaii. Appl Environ Microbiol **61:**1555-1562.

760   40.   **Muyzer, G., A. Teske, C.O. Wirsen, H.W. Jannasch.** 1995. Phylogenetic

761   relationships of *Thiomicrospira* species and their identification in deeop-sea

762   hydrothermal vent samples by denaturing gradient gel electrophoresis of 16S

763   rDNA fragments. Arch Microbiol **164:**165-172.

764   41.   **Nakagawa, S., Y. Takaki, S. Shimamura, A. Reysenbach, K. Takai, and K.**

765   **Horikoshi.** 2007. Deep-sea vent epsilonproteobacterial genomes provide insights

766   into the emergence of pathogens. Proc Natl Acad Sci **104:**12146-12150.

767   42.   **Neumann, S., A. Wynen, G. Trüper, and C. Dahl.** 2000. Characterization of the

768   *cys* gene locus from *Allochromatium vinosum* indicates an unusual sulfate

769   assimilation pathway. Mol Biol Reports **27:**27-33.

770   43.   **Peng, J., and C. H. Huang.** 2006. Rh proteins vs Amt proteins: an organismal

771   and phylogenetic perspective on CO2 and NH3 gas channels. Transfus Clin Biol

772   **13:**85-94.

773   44.   **Pitcher, R. S., and N. J. Watmough.** 2004. The bacterial cytochrome cbb3

774   oxidases. Biochim Biophys Acta **1655:**388-399.

775   45.   **Polz, M. F., and C. M. Cavanaugh.** 1995. Dominance of one bacterial

776   phylotype at a Mid-Atlantic Ridge hydrothermal vent site. Proc. Natl. Acad. Sci.

777   USA **92:**7232-7236.

778   46.   **Poole, R. K.** 2005. Nitric oxide and nitrosative stress tolerance in bacteria.

779   Biochem Soc Trans **33:**176-180.

780   47.   **Pott, A. S., and C. Dahl.** 1998. Sirohaem-sulfite reductase and other proteins

781   encoded in the dsr locus of *Chromatium vinosum* are involved in the oxidation of

782   intracellular sulfur. Microbiol   **144:**1881-1894.

783   48.   **Preisig, O., R. Zufferey, L. Thony-Meyer, C. Appleby, and H. Hennecke.**

784   1996. A high-affinity *cbb3*-type cytochrome oxidase terminates the symbiosis-

785   specific respiratory chain of *Bradyrhizobium japonicum*. J Bacteriol **178:**1532-

786   1538.

787   49.   **Reimann, J., U. Flock, H. Lepp, A. Honigmann, and P. Adelroth.** 2007. A

788   pathway for protons in nitric oxide reductase from *Paracoccus denitrificans*.

789   Biochim Biophys Acta **1767:**362-73.

790   50.   **Ren, Q., K. Chen, and I. T. Paulsen.** 2007. TransportDB: a comprehensive

791   database resource for cytoplasmic membrane transport systems and outer

792   membrane channels. Nucl Acids Res **35:**D274-9.

793   51.   **Ren, Q., K. H. Kang, and I. T. Paulsen.** 2004. TransportDB: a relational

794   database of cellular membrane transport systems. Nucl Acids Res **32:**D284-D288.

795   52.   **Richardson, D. J., and N. J. Watmough.** 1999. Inorganic nitrogen metabolism

796   in bacteria. Curr Opin Chem Biol **3:**207-219.

797   53.   **Romling, U., M. Gomelsky, and M. Y. Galperin.** 2005. C-di-GMP: the dawning

798   of a novel bacterial signalling system. Molec Microbiol **57:**629-639.

799   54.   **Sachs, G., J. A. Kraut, Y. Wen, J. Feng, and D. R. Scott.** 2006. Urea transport

800   in bacteria: Acid acclimation by gastric *Helicobacter* spp. J Membr Biol **212:**71-

801   82.

802   55.   **Sauer, U., and B. J. Eikmanns.** 2005. The PEP-pyruvate-oxaloacetate node as

803   the switch point for carbon flux distribution in bacteria. FEMS Microbiol Rev

804   **29:**765-794.

805   56.   **Schutz, M., I. Maldener, C. Griesbeck, and G. Hauska.** 1999. Sulfide-quinone

806   reductase from *Rhodobacter capsulatus*: Requirement for growth, periplasmic

807   localization, and extension of gene sequence analysis. J Bacteriol **181:**6516-6523.

808   57.   **Scott, K. M., S. M. Sievert, F. N. Abril, L. A. Ball, C. J. Barrett, R. A. Blake,**

809   **A. J. Boller, P. S. Chain, J. A. Clark, C. R. Davis, C. Detter, K. F. Do, K. P.**

810   **Dobrinski, B. I. Faza, K. A. Fitzpatrick, S. K. Freyermuth, T. L. Harmer, L.**

811   **J. Hauser, M. Hügler, C. A. Kerfeld, M. G. Klotz, W. W. Kong, M. Land, A.**

812   **Lapidus, F. W. Larimer, D. L. Longo, S. Lucas, S. A. Malfatti, S. E. Massey,**

813   **D. D. Martin, Z. McCuddin, F. Meyer, J. L. Moore, L. H. Ocampo, J. H.**

814   **Paul, I. T. Paulsen, D. K. Reep, Q. Ren, R. L. Ross, P. Y. Sato, P. Thomas, L.**

815   **E. Tinkham, and G. T. Zeruth.** 2006. The genome of deep-sea vent

816   chemolithoautotroph *Thiomicrospira crunogena*. PLoS Biology **4:**1-17.

817 58. **Seaver, L. C., and J. A. Imlay.** 2001. Alkyl hydroperoxide reductase is the

818   primary scavenger of endogenous hydrogen peroxide in *Escherichia coli*. J

819   Bacteriol **183:**7173-7181.

820 59. **Shahak, Y., M. Schütz, M. Bronstein, G. Hauska, and E. Padan.** 1999.

821   Sulfide-dependent anoxygenic photosynthesis in prokaryotes: sulfide:quinone

822   reductase (SQR), the intial step, p. 211-228. *In* G. A. Peshek, W. Löffelhardt, and

823   C. Schmetterer (ed.), The phototrophic prokaryotes. Kluwer Academic/Plenum

824   Publishers, New York.

825 60. **Shieh, J. S., and W. B. Whitman.** 1987. Pathway of acetate assimilation in

826   autotrophic and heterotrophic methanococci. J. Bacteriol. **169:**5327-5329.

827 61. **Silver, S., and L. T. Phung.** 2005. Genes and enzymes involved in bacterial

828   oxidation and reduction of inorganic arsenic. Appl Environ Microbiol **71:**599-

829   608.

830 62. **Simon, J.** 2002. Enzymology and bioenergetics of respiratory nitrite

831   ammonification. FEMS Microbiol Rev **26**.

832 63. **Simon, J., O. Einsle, P. H. Kroneck, and W. G. Zumft.** 2004. The

833   unprecedented nos gene cluster of *Wolinella succinogenes* encodes a novel

834   respiratory electron transfer pathway to cytochrome c nitrous oxide reductase.

835   FEBS Letters **569:**7-12.

836 64. **Simon, J., M. Sänger, C. Schuster, and R. Gross.** 2003. Electron transport to

837   periplasmic nitrate reductase (NapA) of *Wolinella succinogenes* is independent of

838   a NapC protein. Molec Microbiol **49:**69-79.

839 65. **Stevenson, B. S., and T. M. Schmidt.** 2004. Life history implications of rRNA

840   gene copy number in *Escherichia coli*. Appl Environ Microbiol **70:**6670-6677.

841 66. **Suzuki, M., H. Arai, M. Ishii, and Y. Igarashi.** 2006. Gene structure and

842   expression profile of cytochrome bc nitric oxide reductase from *Hydrogenobacter*

843   *thermophilus* TK-6. Biosci Biotech Biochem **70:**1666-1671.

844 67. **Takai, K., B. J. Campbell, S. C. Cary, M. Suzuki, H. Oida, T. Nunoura, H.**

845   **Hirayama, S. Nakagawa, Y. Suzuki, F. Inagaki, and K. Horikoshi.** 2005.

846   Enzymatic and genetic characterization of carbon and energy metabolisms by

847   deep-sea hydrothermal chemolithoautotrophic isolated of *Epsilonproteobacteria*.

848   Appl Environ Microbiol **71:**7310-7320.

849 68. **Takai, K., M. Suzuki, S. Nakagawa, M. Miyazaki, Y. Suzuki, F. Inagaki, and**

850   **K. Horikoshi.** 2006. *Sulfurimonas paralvinellae* sp nov., a novel mesophilic,

851   hydrogen- and sulfur-oxidizing chemolithoautotroph within the Epsilonproteo-

852   bacteria isolated from a deep-sea hydrothermal vent polychaete nest,

853   reclassification of *Thiomicrospira denitrificans* as *Sulfurimonas denitrificans*

854   comb. nov and emended description of the genus Sulfurimonas. Int J Syst Evol

855   Microbiol **56:**1725-1733.

856 69. **Testerman, T. L., P. B. Conn, H. L. T. Mobley, and D. J. McGee.** 2006.

857   Nutritional requirements and antibiotic resistance patterns of *Helicobacter* species

858   in chemically defined media. J Clin Microbiol **44:**1650-1658.

859 70. **Timmer-Ten Hoor, A.** 1975. A new type of thiosulphate oxidizing, nitrate

860   reducing microorganism: *Thiomicrospira denitrificans* sp. nov. . Netherland J Sea

861   Res **9:**344-350.

862   71.   **Urakawa, H., N. Dubilier, Y. Fujiwara, D. Cunningham, S. Kojima, and D.**

863   **Stahl.** 2005. Hydrothermal vent gastropods from the same family (Provannidae)

864   harbor gamma- and epsilon-proteobacterial endosymbionts. Environ Microbiol

865   **7:**750-754.

866   72.   **Velayudhan, J., and D. J. Kelly.** 2002. Analysis of gluconeogenic and

867   anaplerotic enzymes in *Campylobacter jejuni*: an essential role for

868   phosphoenolpyruvate carboxykinase. Microbiol **148:**685-694.

869   73.   **Wang, G., P. Alamuri, and R. J. Maier.** 2006. The diverse antioxidant systems

870   of *Helicobacter pylori*. Molec Microbiol **61:**847-860.

871   74.   **Watanabe, K., K. Watanabe, Y. Kodama, K. Syutsubo, and S. Harayama.**

872   2000. Molecular characterization of bacterial populations in petroleum-

873   contaminated groundwater discharged from underground crude oil storage

874   cavities. . Appl Environ Microbiol **66:** 4803-4809.

875   75.   **Zhulin, I., B. Taylor, and R. Dixon.** 1997. PAS domain S-boxes in Archaea,

876   bacteria and sensors for oxygen and redox. Tr Biochem Sci **22:**331-333.

877   76.   **Zumft, W. G.** 2005. Nitric oxide reductases of prokaryotes with emphasis on the

878   respiratory, heme-copper oxidase types. J Inorg Biochem **99:**194-215.

879

FIGURE LEGENDS

882

883   FIG. 1. Map of the *Sulfurimonas denitrificans* DSM1251 genome. The two outer

884 rings include protein-encoding genes, which are color-coded based on their membership

885 in COG categories. Ring 3 depicts the deviation from the average G+C (%), while the

886 innermost ring is the GC skew (=[G-C]/[G+C]). R1, R2, R3, and R4 are ribosomal RNA

887 operons (with their orientation indicated with arrows), and the two regions marked T are

888 identical large transposons. The G+C and GC skew rings were calculated with a sliding

889 window of 10,000 bp with a window step of 100.

890

891

892   FIG. 2. Model for electron transport in *Sulfurimonas denitrificans.* Abbreviations:

893 I—NADH dehydrogenase; II—succinate dehydrogenase/fumarate reductase; III—bc_1

894 complex; IV—cbb_3 cytochrome c oxidase; Cyt_n c—cytochrome c; CCP—cytochrome c

895 peroxidase; FDH—formate dehydrogenase; HYD—hydrogenase; MK—menaquinone;

896 MKH_2—menaquinol; Mo—molybdodenum containing cofactor; NAP—periplasmic

897 nitrate reductase; NIR—nitrite reductase; NOR—nitric oxide reductase; NOS—nitrous

898 oxide reductase; PM—plasma membrane; PSR—polysulfide reductase; SOX—sulfur

899 oxidation system. Candidate monoheme cytochromes *c* listed as "Cyt_x *c*" are

900 *Suden_0904*, *Suden_0741*, *Suden_0578* (all COG2863) as well as *Suden_0865*,

901 *Suden_1329* and *Suden_1112* (no COG assignment).

902

903

904   FIG. 3. Central carbon metabolism in *Sulfurimonas denitrificans.* Abbreviations:

905   2-OG—2-oxoglutarate; acCoA—acetyl-Coenzyme A; ACL—ATP-citrate lyase;

906   APFK—ATP-dependent phosphofructokinase; CS—citrate synthase; F6P—fructose 6-

907   phosphate; FBP—fructose 1,6-bisphosphate; FBPP—fructose 1,6 bisphosphate

908   phosphatase; FR—fumarate reductase; fum—fumarate; isocit—isocitrate; mal—malate;

909   MQ—menaquinone; OAA—oxaloacetate; PC—pyruvate carboxylase; PEP—

910   phosphoenolpyruvate; pyr—pyruvate; suc—succinate; suCoA—succinyl-Coenzyme A.

911

912   FIG. 4. Phylogenetic relationships of alpha-subunits of oxaloacetate

913   decarboxylase (OAD), pyruvate carboxylase (PVC), oxoglutarate carboxylase (OGC),

914   and type III pyruvate carboxylase to the product of *Suden_1259* of *Sulfurimonas*

915   *denitrificans*. Sequences were aligned using the program package MacVector. Neighbor-

916   joining and Parsimony trees based on the predicted amino acid sequences were calculated

917   using PAUP 4.0b10. Bootstrap values (1,000 replicates) for the major nodes are given for

918   the neighbor-joining (first value) and parsimony analyses (second value).

919

920   FIG. 5. A large gene cluster from the *Sulfurimonas denitrificans* genome that

921   includes many of the genes for chemotaxis signal transduction.

922

923

924

925

926

927

928   TABLE 1. Comparative genome features for epsilonproteobacteria[‡]

929

	Size	%		rRNA	#
Species	(Mbp)	coding	%GC	operons	CDS
Sulfurimonas denitrificans DSM1251	2.20	93.8	34.5	4	2104
Sulfurovum sp. NBC37-1	2.56	90.1	43.8	3	2466
Nitratiruptor sp. SB155-2	1.88	95.1	39.7	3	1857
Campylobacter fetus 82-40	1.80	90.0	33.3	3	1719
Campylobacter jejuni NCTC 11168	1.64	95.4	30.6	3	1629
Campylobacter jejuni RM1221	1.78	91.8	30.3	3	1838
Helicobacter hepaticus ATCC 51449	1.80	93.4	35.9	1	1875
Helicobacter acinonychis Sheeba	1.55	89.0	38.2	2	1618
*Helicobacter pylori*26695	1.67	90.0	38.9	2*	1576
Helicobacter pylori J99	1.64	90.7	39.2	2*	1491
Helicobacter pylori HPAG1	1.59	91.0	39.1	2*	1544
Wolinella succinogenes DSM1740	2.11	94.5	48.5	3	2043

930

931   [‡]Data for all taxa, except for *Sulfurivom* sp. NBC37-1 and *Nitratiruptor* sp.

932   SB155-2 were collated from the Integrated Microbial Genomes webpage and had been

933   generated using consistent methodology. For *Sulfurivom* sp. NBC37-1 and *Nitratiruptor*

934   sp. SB155-2, data were collected from (41), for which slightly different methodologies

935   were used to identify coding sequences (CDS).

936   *In *H. pylori*, the 16S gene is not collocated with the 23S and 5S genes in an

937   operon. Additionally, an orphan 5S sequence is found in strain 26695

938

939

940 TABLE 2. Regulatory and signaling proteins of *Sulfurimonas denitrificans* and other

941 obligate chemolithoautotrophs*

942

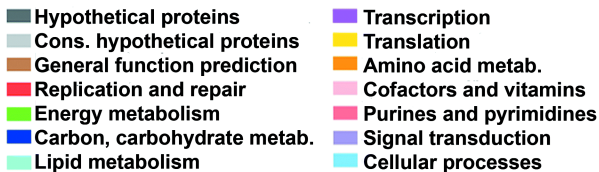| Number:			
S. denitrificans	*T. crunogena*	*N. oceani*	Functional Description
56	72	104	Transcription/Elongation/Termination Factors
19	4	2	Two component transcriptional regulator, winged helix family
146	128	75	Signal Transduction proteins
28	27	8	Chemotaxis Signal Transduction proteins
13	14	1	Methyl-accepting chemotaxis proteins
118	101	67	Non-Chemotaxis Signal Transduction
36	17	18	Signal Transduction Histidine Kinase
42	49	16	Cyclic nucleotide signal transduction
202	200	179	Total

943   *Sulfurimonas denitrificans* DSM1251 is compared to gammaproteobacteria

944   *Thiomicrospira crunogena* XCL-2 and *Nitrosococcus oceani* ATCC 19707.

945

946

947

Hypothetical proteins	Transcription
Cons. hypothetical proteins	Translation
General function prediction	Amino acid metab.
Replication and repair	Cofactors and vitamins
Energy metabolism	Purines and pyrimidines
Carbon, carbohydrate metab.	Signal transduction
Lipid metabolism	Cellular processes

Chl. vibrioformis
Chl. luteolum
Chl. limicola
Chl. tepidum
Ptc. aestuarii
Chl. phaeobacteroides
C. jejuni
C. fetus
Nitratiruptor
Desulfotalea
Sulfurovum
Corynebacterium
Sulfurimonas
100/100
99/100
100/100
100/100
Geobacter
Legionella
Type III PVC
P. aeruginosa
P. putida
100/97
PVC
Debaryomyces
100/100
80/-
100/100
Thiobacillus
63/62
100/100
Pyrococcus
100/100
Helicobacter
100/100
Thermococcus
PVC
OAD
Methanocaldococcus
100/100
Wolinella
Methanosarcina
PVC
Ruthia
Methanothermobacter
OGC
Shewanella
Dehalococcoides
Methylococcus
V. cholerae 2
Aquifex
V. cholerae 1
Hydrogenobacter
Aquifex
Hydrogenobacter
Salmonella
Klebsiella

0.1 changes

Tmden_0963

winged-helix transcription regulator

histidine kinase

histidine kinase

diguanylate cyclase/phosphodiesterase
with PAS/PAC domain

response regulator receiver protein

cheW

methyl-accepting chemotaxis protein

hypothetical

cheA

cheR

histidine kinase

cheD

cheB

histidine kinase

conserved hypothetical protein

methyl-accepting chemotaxis protein

response regulator receiver histidine kinase

methyl-accepting chemotaxis protein

diguanylate cyclase/phosphodiesterase
with a response regulator receiver

histidine kinase

histidine kinase

response regulator receiver protein

methyl-accepting chemotaxis protein

methyl-accepting chemotaxis protein

Tmden_0990

1002406

1012406

1022406

1032406

1042406

1

Further details of genome structure. In addition to the large transposon

2

3 interrupting one of the flagellar biosynthetic operons (Fig. S1), another transposase gene

4 (*Tmden_1713*) is located near a tRNA^{Thr} gene, adjacent to a hypothetical protein gene

5 (*Tmden_1712*) whose 3'-end is 82% identical, at the nucleotide level, with two

6 transposase genes located downstream (*Tmden_1724* and *Tmden_1725*). These two

7 genes, and regions 5' and 3' of each (totaling 1302 bp apiece), are 100% identical to each

8 other. This region also includes genes encoding a recombinase and phage integrase

9 (*Tmden_1723;* Fig. S2); the presence of the phage integrase gene, identical repeats, and

10 juxtaposition to a tRNA gene suggest that this portion of the genome may be a remnant of

11 a degraded or partially excised prophage. Other potential tranposase genes are present

12 (*Tmden_0961; Tmden_1698; Tmden_1708)*, but have insufficient sequence similarity to

13 known proteins for deducing their function convincingly.

14 Twelve phage integrase genes are present. As expected for phage genes, six of

15 these are near tRNA genes (*Tmden_0248, Tmden_0779, Tmden_0800, Tmden_1618,*

16 *Tmden_1723; Tmden_1743)* which are common insertion sites for lysogenic phages (1),

17 and three are flanked by clusters of genes encoding hypothetical proteins *(Tmden_1247;*

18 *Tmden_1618; Tmden_1633)*, which is consistent with the observation that many phage

19 genes are unique and uncharacterized (2). Two phage integrase genes *(Tmden_0938;*

20 *Tmden_0959)* flank genes encoding a type I restriction modification system gene cluster

21 and are part of a larger region (bp 977850 – 1002764; 5 o'clock, Fig. 1), that have a

22 negative G+C anomaly (31.1%). The remaining two are included in the large transposons

23 described above (*Tmden_0693* and *Tmden_1590).*

24   Another negative G+C anomaly, which also has a GC skew anomaly, is visible on

25   the genome map at approximately 1 o'clock (bps 192095 – 210856; 30.5% G+C; Fig. 1).

26   Although this region contains several genes encoding hypothetical proteins, as one might

27   expect were it derived from a phage, it does not appear to include any transposase or

28   integrase genes, nor does it include any repeated sequences that might suggest recent

29   gene rearrangement in this region.

30   **Restriction-modification systems.** *S. denitrificans* has numerous restriction-

31   modification (RM) systems encoded in its genome. Eleven DNA methyltransferase genes

32   are present, and encode methyltransferases similar to those found in Type I

33   (*Tmden_0697, Tmden_0942, Tmden_1594*), Type II (*Tmden_0121; Tmden_0129;*

34   *Tmden_0130; Tmden_0478; Tmden_0537; Tmden_1565; Tmden_1839; Tmden_1855*)

35   and Type III (*Tmden_1355*) restriction-modification systems. For 6 of these

36   methyltransferases, genes encoding restriction enzymes are nearby (Type 1:

37   *Tmden_0700; Tmden_0948; Tmden_1597;* Type II: *Tmden_128; Tmden_1854;* Type III:

38   *Tmden_1350)* and for two of them the genes appear to encode fused methylase/restriction

39   enzymes (*Tmden_0478, Tmden_0537*)(5). Based on genome sequence data, such large

40   numbers of RM systems are not unusual for epsilonproteobacteria: *Helicobacter pylori*

41   has 24 RM systems (3), *C, jejuni* has 10, and *W. succinogenes* has 5 (4, 5). If active,

42   perhaps in *S. denitrificans* these systems provide a robust defense against the introduction

43   of phage and other 'non-native' DNA into the genome.

44

REFERENCES

1. **Canchaya, C., G. Fournous, and H. Brussow.** 2004. The impact of prophages on bacterial chromosomes. Molec Microbiol **53:**9-18.

2. **Hambly, E., and C. A. Suttle.** 2005. The viriosphere, diversity, and genetic exchange within phage communities. Curr Opin Microbiol **8:**444-450.

3. **Kong, H., L. Lin, N. Porter, S. Stickel, D. Byrd, J. Posfai, and R. j. Roberts.** 2000. Functional analysis of putative restriction-modification system genes in the *Helicobacter pylori* J99 genome. Nucl Acids Res **28:**3216-3223.

4. **Roberts, R. J., T. Vincze, J. Posfai, and D. Macelis.** 2007. REBASE--enzymes and genes for DNA restriction and modification. Nucl Acids Res **35:**D269-D270

5. **Roberts, R. J., T. Vincze, J. Posfai, and D. Macelis.** 2005. REBASE Restriction enzymes and DNA methyltransferases. Nucl Acids Res **33:**D230-232.

58
59 **FIGURE LEGENDS.**
60
61   FIG. S1. Two large identical transposons from the *Sulfurimonas denitrificans*

62 genome. Numbers indicate the position of the regions, in nucleotides, with respect to the

63 origin of replication, and the arrows indicate the presence of the inverted repeat

64 sequences at each end: > = TGTCATTTACAA; < = TTGTAAATGACA.

65

66   FIG. S2. Map of a region from the *Sulfurimonas denitrificans* genome that

67 includes a small repeated region. The duplicate copies of this repeat include the two

68 adjacent transposase genes (shaded grey), while a third region with a high level of

69 identity (82%) is included within a hypothetical gene upstream (also shaded grey).

70 Numbers indicate the position of the regions, in nucleotides, with respect to the origin of

71 replication.

72

73
74   FIG. S3. Phylogenetic relationships of NosZ from different bacteria and the

75 archaeon *Pyrobaculum calidifontis*. Suden_1298 is part of a novel nos cluster previously

76 identified in *Wolinella succinogenes* (63). All epsilonproteobacterial sequences have a C-

77 terminal extension and contain a heme *c*-binding motif. The sequences from

78 *Dechloromonas aromatica* and *Magentospirillum magnetotacticum* also have a

79 (somewhat shorter) C-terminal extension, but are lacking a heme *c*-binding motif (63)

80 Sequences were aligned using the program package MacVector. Neighbor-joining and

81 Parsimony trees based on the predicted amino acid sequences were calculated using

82   PAUP 4.0b10. Bootstrap values (1,000 replicates) for the major nodes are given for the

83   neighbor-joining (first value) and parsimony analyses (second value).
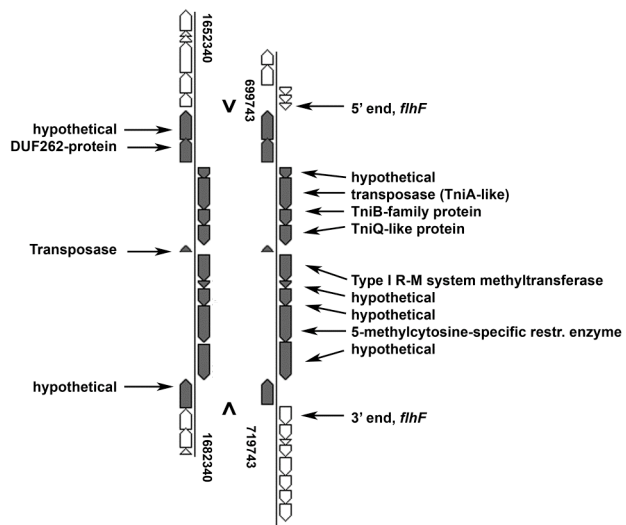
84

84

85

86

87

FIG. S1

88

89

90

91

92

93
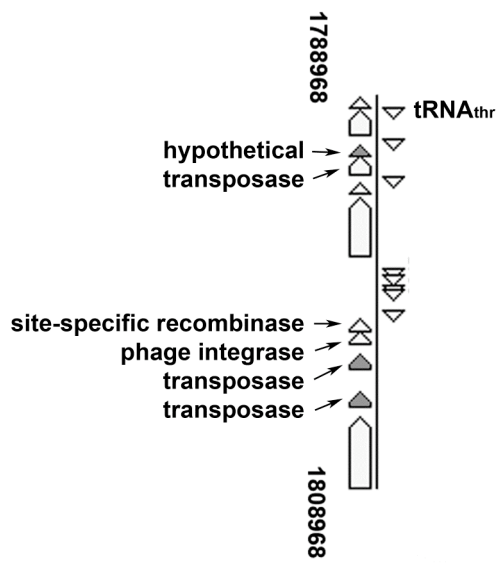
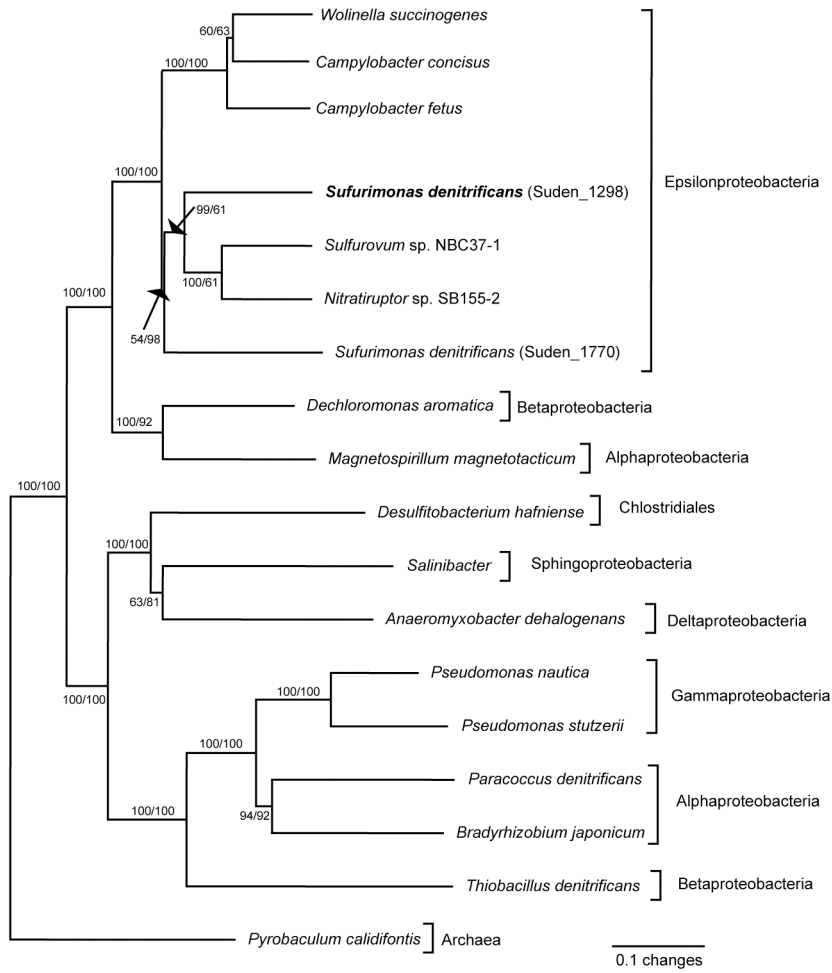94   FIG. S1

95

96

97

98

99

100

101

102

103

104

105   FIG. S2

106

107

108

FIG. S3