



LAWRENCE  
LIVERMORE  
NATIONAL  
LABORATORY

# **The Genome Sequence of the Obligately Chemolithoautotrophic, Facultatively Anaerobic Bacterium *Thiobacillus denitrificans***

*H. R. Beller, P. S. G. Chain, T. E. Letain, A.  
Chakicherla, F. W. Larimer, P. M. Richardson,  
M. Coleman, A. P. Wood, and D. P. Kelly*

**Published February 2006**

Published in the *Journal of Bacteriology*

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

This work was performed under the auspices of the U.S. Department of Energy by University of California, Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.

**PREPRINT**

To be published in the *Journal of Bacteriology*

The genome sequence of the obligately chemolithoautotrophic,  
facultatively anaerobic bacterium *Thiobacillus denitrificans*

Harry R. Beller,<sup>1\*</sup> P. S. G. Chain,<sup>1,3</sup> T. E. Letain,<sup>1</sup> A. Chakicherla,<sup>1</sup> F. W.  
Larimer,<sup>2,3</sup> P. M. Richardson,<sup>3</sup> M. Coleman,<sup>1</sup> A. P. Wood,<sup>4</sup> and D. P. Kelly<sup>5</sup>

<sup>1</sup>Lawrence Livermore National Laboratory, Livermore, CA 94551

<sup>2</sup>Oak Ridge National Laboratory, Oak Ridge, TN 37831

<sup>3</sup>U.S. DOE Joint Genome Institute Production Genomics Facility, Walnut Creek, CA 94598

<sup>4</sup>Department of Microbiology, King's College London, London, SE1 9RT, UK

<sup>5</sup>Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, UK

**Running title:** Genome sequence of *Thiobacillus denitrificans*

**Corresponding author:**

Lawrence Livermore National Laboratory

P.O. Box 808, L-542

Livermore, CA 94551-0808

E-mail [beller2@llnl.gov](mailto:beller2@llnl.gov)

Phone (925) 422-0081

Fax (925) 423-7998

1 **Abstract**

2 The complete genome sequence of *Thiobacillus denitrificans* ATCC 25259 is the first to  
3 become available for an obligately chemolithoautotrophic, sulfur-compound-oxidizing,  $\beta$ -  
4 proteobacterium. Analysis of the 2,909,809-base pair genome will facilitate our molecular and  
5 biochemical understanding of the unusual metabolic repertoire of this bacterium, including its  
6 ability to couple denitrification to sulfur-compound oxidation, to catalyze anaerobic, nitrate-  
7 dependent oxidation of Fe(II) and U(IV), and to oxidize mineral electron donors. Notable  
8 genomic features include (1) one to two percent of the genes encoding *c*-type cytochromes,  
9 which is greater than for almost all bacterial and archaeal species sequenced to date, (2) genes  
10 encoding two [NiFe]hydrogenases, which is particularly significant because no information on  
11 hydrogenases has previously been reported for *T. denitrificans* and hydrogen oxidation appears  
12 to be critical for anaerobic U(IV) oxidation by this species, (3) a diverse complement of more  
13 than 50 genes associated with sulfur-compound oxidation (including *sox* genes, *dsr* genes, and  
14 genes associated with the AMP-dependent oxidation of sulfite to sulfate), some of which occur  
15 in multiple (up to eight) copies, (4) a relatively large number of genes associated with inorganic  
16 ion transport and heavy metal resistance, and (5) a paucity of genes encoding organic-compound  
17 transporters, commensurate with obligate chemolithoautotrophy. Ultimately, the genome  
18 sequence of *T. denitrificans* will enable elucidation of the mechanisms of aerobic and anaerobic  
19 sulfur-compound oxidation by  $\beta$ -proteobacteria, and will help reveal the molecular basis of this  
20 organism's role in major biogeochemical cycles (i.e., sulfur, nitrogen, and carbon) and  
21 groundwater restoration.

## INTRODUCTION

1  
2       *Thiobacillus denitrificans*, first isolated by Beijerinck over a century ago (4), was one of  
3 the first non-filamentous bacteria ever described to be capable of growth on inorganic sulfur  
4 compounds as sole energy sources (47, 49). Characterized by its ability to conserve energy from  
5 the oxidation of inorganic sulfur compounds under either aerobic or denitrifying conditions, *T.*  
6 *denitrificans* is the best studied of the very few obligate chemolithoautotrophic species known to  
7 couple denitrification to sulfur-compound oxidation (*Thiomicrospira denitrificans* and  
8 *Thioalkalivibrio thiocyanodenitrificans* also have this ability; 76, 85). Despite many years of  
9 work on the biochemistry of inorganic sulfur compound oxidation by *Thiobacillus thioparus* and  
10 *T. denitrificans*, the mechanisms of oxidation and how they are coupled to energy conservation  
11 are still not well understood in these  $\beta$ -proteobacteria, relative to the advances made with  
12 facultatively chemolithotrophic  $\alpha$ -proteobacterial genera, such as *Paracoccus* and *Starkeya* (28,  
13 39, 45, 50). The availability of the complete genome sequence should enable elucidation of the  
14 sulfur-oxidation pathway(s) and lead to specifically focused biochemical investigations to  
15 resolve these knowledge gaps.

16       Recent studies have revealed that, in addition to sulfur-compound oxidation, *T.*  
17 *denitrificans* has broader oxidative capabilities that may not contribute to energy conservation,  
18 including anaerobic, nitrate-dependent oxidation of certain metals, such as iron (78). The  
19 metabolic repertoire of this widely distributed bacterium can influence the carbon, nitrogen,  
20 sulfur, and iron cycles in many soil, aquifer, and sediment environments and is particularly  
21 relevant to *in situ* bioremediation of contaminated groundwater. Environmentally relevant  
22 capabilities of *T. denitrificans* include intrinsic biodegradation of nitrate, one of the most  
23 problematic groundwater contaminants worldwide, by anaerobic, nitrate-dependent oxidation of

1 minerals such as FeS and pyrite (FeS<sub>2</sub>) (e.g., 6, 56, 78). *T. denitrificans* is the first and only  
2 autotrophic bacterium reported to carry out anaerobic (nitrate-dependent) oxidation of U(IV)  
3 oxide minerals (5), which could partially counteract efforts to remediate uranium-contaminated  
4 aquifers by *in situ* reductive immobilization [i.e., microbially mediated conversion of relatively  
5 soluble U(VI) species to poorly soluble U(IV) minerals]. The intriguing mechanism by which  
6 this species can oxidize mineral electron donors that cannot be taken into the cell is currently  
7 unknown but its elucidation will be facilitated by availability of the genome.

8           In this article, we present the complete genome sequence of *T. denitrificans* strain ATCC  
9 25259, the first obligately chemolithoautotrophic, sulfur-oxidizing, β-proteobacterium to be  
10 sequenced. We describe some general features of the *T. denitrificans* genome, including recent  
11 gene acquisition, as well as genetic components involved in sulfur-compound oxidation,  
12 hydrogen metabolism, aerobic respiration, denitrification, autotrophy, central carbon  
13 metabolism, and heavy metal resistance.

## 14                                   **MATERIALS AND METHODS**

15           **Organism source.** *T. denitrificans* strain ATCC 25259 was obtained from the American  
16 Type Culture Collection (ATCC). This strain was originally isolated by B. F. Taylor in the  
17 1960s (83, 84) and was deposited with the ATCC in 1969, where it was freeze-dried for storage  
18 and distribution. For this study, unless noted otherwise, *T. denitrificans* was grown with  
19 thiosulfate under denitrifying conditions, as described elsewhere (5).

20           **Sequencing, coding sequence prediction, and annotation.** Genomic DNA was isolated  
21 from *T. denitrificans* and the complete genome was sequenced as described previously (15).  
22 Briefly, small-insert (2-3 kb), medium-insert (6.5-8.5 kb), and large-insert (35-45 kb) libraries  
23 were generated by random mechanical shearing of genomic DNA. In the initial random

1 sequencing phase, approximately nine-fold sequence coverage was achieved. The sequences  
2 from all libraries were assembled together and viewed using the Phred/Phrap/Consed suite (P.  
3 Green, University of Washington; 22, 23, 31). Physical gaps were closed by PCR and  
4 sequencing. Open reading frames likely to encode proteins (CDS, or coding sequences) were  
5 identified and annotated by automated and manual curation as previously described (15).

6 **Comparative genomics.** The Integrated Microbial Genomes (IMG) system of the Joint  
7 Genome Institute (<http://img.jgi.doe.gov/>) was used for identification of orthologs and to identify  
8 CDS unique to *T. denitrificans* based on BLASTP; the cutoff values applied were  $E < 10^{-5}$  and  
9 30% identity, and  $E < 10^{-2}$  and 20% identity, respectively.

10 **Generation of phylogenetic trees and other analyses.** Phylogenetic trees were  
11 generated by identifying potential homologs of translated *T. denitrificans* CDS using BLASTP  
12 (1) searches against the non-redundant (nr) GenBank database from the National Center for  
13 Biotechnology Information. Typically, trees included the top 50 BLASTP matches. However,  
14 sequences were excluded if their BLASTP E values fell below a cutoff of  $10^{-5}$ . For the  
15 phylogenetic trees presented in this article, the lowest ranking sequences shared 24 to 54%  
16 identity and 43 to 71% similarity with the query sequences. Sequences were aligned and  
17 alignments refined using ClustalX (38) version 1.8 and manual adjustments. Phylogenetic trees  
18 were generated by using the *protdist* program and the *neighbor* program of the PHYLIP package  
19 (24, 25) to calculate distances (using the Jones-Taylor-Thorton matrix) and for clustering  
20 (Neighbor-Joining method), respectively. Membrane-spanning domains of proteins were  
21 identified using TmHMM (52, 75). The SignalP program (7) was used to identify putative signal  
22 peptides.

1           **Reverse transcription, quantitative PCR analysis.** For 18 selected genes, mRNA  
2 levels under exposure conditions of interest were determined by reverse transcription (RT),  
3 quantitative PCR analysis. The target genes included Tbd0210, 0561, 0562, 0822, 0823, 0872,  
4 0873, 0874, 1365, 1408, 1926, 2282, 2283, 2326, 2327, 2480, 2488, 2658. Cultures of *T.*  
5 *denitrificans* were grown anaerobically with thiosulfate and nitrate and were then anaerobically  
6 harvested by centrifugation and washed using techniques and conditions described elsewhere (5).  
7 Washed cells (~6.5 mg protein) were resuspended in stoppered serum bottles under strictly  
8 anaerobic conditions in 5 or 10 mL of bicarbonate-buffered medium. Cells were resuspended  
9 either with thiosulfate (20 mM) and nitrate (20 mM) or FeCO<sub>3</sub> (3.5 mmol/L) and nitrate (3.5  
10 mM). Chemical analyses (ion chromatography; spectrophotometric determination of ferrozine-  
11 iron complexes) of the resuspended cultures were performed to confirm that nitrate reduction and  
12 either thiosulfate or Fe(II) oxidation were occurring when the cultures were harvested for RNA.  
13 RNA extraction was carried out with a MasterPure Complete DNA and RNA Purification Kit  
14 (EpiCentre) using a modified protocol. Total RNA was reverse transcribed and amplified using  
15 a QuantiTect SYBR Green RT-PCR kit (Qiagen) with gene-specific primers. Each gene-specific  
16 PCR was performed in triplicate using 25- $\mu$ L reactions containing ~20 ng of template on a Prism  
17 7000 cycler (ABI). Calibration curves were performed with genomic DNA serially diluted over  
18 a range of four to five orders of magnitude. The PCR conditions were optimized to be performed  
19 as follows for all transcripts: 30-35 cycles at 50°C for 30 minutes; 95°C for 15 minutes; 94°C for  
20 15 seconds; 58°C for 30 seconds; 72°C for 30 seconds.

21           **Nucleotide sequence accession number.** The annotated genome sequence has been  
22 deposited in the GenBank/EMBL sequence database under accession no. CP000116.



## RESULTS AND DISCUSSION

1  
2       **General genome features.** The genome of *T. denitrificans* strain ATCC 25259 consists  
3 of a single circular chromosome of 2,909,809 bp in length with an average G+C content of  
4 66.1% (Fig. 1 and Table 1). GC skew analysis does not reveal the origin of replication.  
5 Nucleotide position 1 of the chromosome is assigned to the predicted origin of replication,  
6 flanked on one side by the *dnaA* (Tbd0001), *dnaN* (Tbd0002), and *gyrB* (Tbd0003) genes and on  
7 the other by *rpmH* (Tbd2827) and *rnpA* (Tbd2826). Two copies of the 16S ribosomal RNA  
8 operon are located in regions of the genome separated by more than 1380 kb. Few other  
9 repeated sequences were discovered in the genome, with the exception of a duplicated IS4-like  
10 insertion sequence. Thus, the numbers of repeated elements, IS elements, and transposons are  
11 low compared with those typically found in bacterial genomes sequenced to date.

12       A total of 2827 protein-encoding genes were identified, with an average length of 952 bp,  
13 accounting for 92.5% of the sequence. Among the predicted genes, 2183 (77.2%) have been  
14 assigned a putative function, with the remainder designated as encoding a protein with unknown  
15 function, a conserved hypothetical protein, or a hypothetical protein. Of these, 89 predicted  
16 genes have no match using BLAST vs. the non-redundant (nr) database with an E value cutoff of  
17  $10^{-6}$  or better. When distributed into biological role categories based on Clusters of Orthologous  
18 Groups of proteins (COGs) (82), the largest number of predicted proteins fell into the categories  
19 of Energy Production (6.6%), Cell Envelope Biogenesis (6.0%), and Inorganic Ion Transport  
20 (5.7%).

21       When searched against the KEGG database of complete genomes, more than half (1624)  
22 of the predicted proteins revealed top BLAST hits to one of the twelve  $\beta$ -proteobacteria available  
23 in this database, including *Azoarcus* sp. EbN1 (626 hits), the obligate chemolithoautotrophic

1 bacterium *Nitrosomonas europaea* (346 hits), *Chromobacterium violaceum* (236 hits), *Ralstonia*  
2 *solanacearum* GMI1000 (189 hits), *Burkholderia pseudomallei* (118 hits), and *Bordetella*  
3 *bronchiseptica* (61 hits). Aside from  $\beta$ -proteobacteria, the organisms most frequently associated  
4 with the top BLAST hits were *Methylococcus capsulatus* (107 hits), *Pseudomonas aeruginosa*  
5 (75 hits), an environmentally versatile  $\gamma$ -proteobacterium, *Pseudomonas putida* (42 hits), and  
6 *Geobacter sulfurreducens* (42 hits), a  $\delta$ -proteobacterium best known for its versatility in  
7 dissimilatory metal reduction.

8       **Recent gene acquisition.** There is ample evidence of horizontal transfer in the *T.*  
9 *denitrificans* genome, inferred from local base composition to have been imported from an  
10 evolutionarily distant source. At least 13 regions, up to 25 kb in size, have been identified [based  
11 on anomalies in observed G+C content and trinucleotide signature (40)] as likely recent  
12 integration events into the *T. denitrificans* genome that have not had time to drift toward the  
13 genome average (Table 2). Almost all of these putative regions of horizontal gene transfer carry  
14 phage integrases or other phage remnants and a few of these regions are also flanked by tRNA  
15 genes, which are used as integration sites for many bacteriophages. Most of these regions are  
16 found to harbor many hypothetical or conserved hypothetical genes, several carry restriction  
17 modification systems, which are known to be associated with mobile elements and indeed act as  
18 selfish mobile genetic elements themselves (51), whereas another region carries a cluster of  
19 genes encoding part of a type IV pilus. Although these likely recent insertions in the *T.*  
20 *denitrificans* chromosome are consistent with the concept of a fluid genome, it remains to be  
21 shown what role these regions may play in the metabolic or defense repertoire of *T. denitrificans*.

22       **Taxonomy of *T. denitrificans* strain ATCC 25259.** Following the reclassification of  
23 numerous  $\alpha$ -,  $\beta$ -, and  $\gamma$ -proteobacterial species previously classified as “*Thiobacillus*”, only three

1  $\beta$ -proteobacteria are now securely placed in that genus: *T. thioparus*, *T. denitrificans*, and *T.*  
2 *aquaesulis* (47, 49). Of these, the last is a moderate thermophile and only *T. denitrificans* is  
3 capable of strictly chemolithotrophic anaerobic growth with denitrification using inorganic  
4 sulfur-compound oxidation as the sole source of energy (46, 49). Apart from this feature, *T.*  
5 *denitrificans* and *T. thioparus* are physiologically very similar mesophilic obligate  
6 chemolithoautotrophs. The G+C content of different strains range between 62-67% for *T.*  
7 *thioparus* and 63-68% for *T. denitrificans* (46, 79, 88). DNA hybridization distinguished the two  
8 species, which show only 22-29% cross-hybridization (46). Comparison of 16S rRNA gene  
9 sequences in the GenBank nr database shows that identities between different strains of *T.*  
10 *thioparus* range between 95.2-99.6%, and that the type strains of *T. thioparus* (ATCC 8158) and  
11 *T. denitrificans* (NCIMB 9548) show 96.3% identity to each other (D.P. Kelly, unpublished  
12 data). Consequently, the ability to denitrify is a key physiological criterion in distinguishing  
13 between the species, given the similarity of their 16S rRNA gene sequences. *T. denitrificans*  
14 strain ATCC 25259 was originally isolated from Texas soil (84; ATCC catalog) and has been  
15 shown to have physiological properties characteristic of this species. *T. denitrificans* ATCC  
16 25259 shows 97.6% identity to the type strain *T. denitrificans* NCIMB 9548 (with respect to 16S  
17 rRNA genes).

18 **Sulfur-compound oxidation.** A number of enzymes involved in inorganic sulfur-  
19 compound oxidation and energy conservation have previously been studied in *T. denitrificans*  
20 and the closely related  $\beta$ -proteobacterium, *T. thioparus*, some of which have been purified and  
21 characterized biochemically. The *T. denitrificans* enzymes include APS (adenylylsulfate)  
22 reductase, an AMP-independent sulfite oxidase, a siroheme sulfite reductase, and APS:phosphate  
23 adenylyltransferase (or APAT) (2, 12, 13, 29, 69-71). Biochemical studies, as well as the high

1 growth yields and the entry into the electron transport chain of electrons from sulfur-compound  
2 oxidation at the level of quinone/cytochrome *b*, indicate highly efficient energy-conserving  
3 mechanisms of sulfur-compound oxidation in *T. denitrificans* (42-45). The mechanisms of  
4 inorganic-sulfur oxidation by *T. denitrificans*, which oxidizes polythionates as well as  
5 thiosulfate, sulfide and, for some strains, thiocyanate, appear more similar to those of  $\gamma$ -  
6 proteobacteria such as *Halothiobacillus* than to the well-defined thiosulfate-oxidizing,  
7 multienzyme system of  $\alpha$ -proteobacterial chemolithotrophs (45, 50, 68).

8         As would be expected from previous biochemical work with *T. denitrificans*, its genome  
9 contains a diverse complement of genes encoding enzymes that catalyze inorganic sulfur-  
10 compound oxidation and energy conservation (by both substrate-level and electron transport-  
11 linked phosphorylation). The importance of sulfur-compound oxidation to *T. denitrificans* is  
12 underscored by the occurrence of multiple oxidation pathways for certain sulfur compounds and  
13 multiple copies of a number of genes associated with sulfur-compound oxidation (see overview  
14 in Fig. 2). The unusual ability of this bacterium to oxidize inorganic sulfur compounds under  
15 both aerobic and denitrifying conditions raises the question of whether different sulfur-oxidizing  
16 enzymes are involved under aerobic vs. anaerobic conditions. Although such questions cannot  
17 be resolved by genome analysis alone, they can be addressed by transcriptional studies using  
18 whole-genome oligonucleotide microarrays, which are now enabled by the availability of the *T.*  
19 *denitrificans* genome. Included among the genes that are likely to be critical to sulfur-compound  
20 oxidation by *T. denitrificans* are *sox* (sulfur-oxidation) genes, *dsr* (dissimilatory sulfite  
21 reductase) genes, and genes associated with the AMP-dependent oxidation of sulfite to sulfate  
22 (Fig. 2). These genes serve as the focus of this section.

1 Little molecular genetic work on the sulfur-oxidation (*sox*) systems of *T. denitrificans* has  
2 been conducted to date. The model for genes encoding sulfur-oxidation enzymes has been  
3 derived primarily from three  $\alpha$ -proteobacteria: *Paracoccus pantotrophus*, *Starkeya novella*, and  
4 *Pseudaminobacter salicylatoxidans*. In *T. denitrificans*, genes showing various levels of  
5 sequence identity to *sox* genes of these  $\alpha$ -proteobacteria have been detected, but gene clusters of  
6 the length found in facultatively chemolithotrophic, aerobic, thiosulfate-oxidizing bacteria do not  
7 occur. Thus, extensive *sox* clusters have been observed in *P. pantotrophus*, *S. novella*, and *Psb.*  
8 *salicylatoxidans*: *soxRSVWXYZABCDEFGH*, *soxFDCBZYAXWV* and *soxGTRSVWXYZABCD*,  
9 respectively (GenBank X79242, AF139113, AJ404005; 27, 28, 91). In contrast, the largest  
10 cluster in *T. denitrificans* consists of *soxXYZAB* (Tbd0567-Tbd0563), using the *P. pantotrophus*  
11 naming scheme. Encoded polypeptide sequence identities compared to the *soxX,Y,Z,A,B* of *P.*  
12 *pantotrophus*, *Chlorobium tepidum*, *Psb. salicylatoxidans*, and *S. novella* were in the range of  
13 28-55%. This indicated that while the genes coding for these Sox functions had been putatively  
14 identified, they differed significantly from those of the reference organisms. Interestingly, the  
15 translated *soxXYZA* genes showed higher identities to those of the green sulfur bacterium *C.*  
16 *tepidum* (42-55%) than to those of *Paracoccus*, *Starkeya* and *Pseudaminobacter* (28-38%).  
17 Additional copies of *soxXA* were identified outside of the *soxXYZAB* cluster in *T. denitrificans*  
18 (Tbd0917-0918). A noteworthy difference between SoxA encoded in the *T. denitrificans*  
19 genome and those of *P. pantotrophus* and *Psb. salicylatoxidans* is that the latter are diheme  
20 cytochromes (28) whereas SoxA copies in *T. denitrificans* (Tbd0564, and putatively Tbd0918)  
21 are monoheme cytochromes.

22 For *soxB*, the *T. denitrificans* sequence can be compared to another  $\beta$ -proteobacterium as  
23 well as to  $\alpha$ -proteobacteria, because the sequence from *T. thioparus* is available (59; GenBank

1 AJ294326; partial sequence – 344 translated amino acids). In fact, of the *sox* genes, only *soxB*  
2 has thus far been the subject of intercomparison across the  $\alpha$ -,  $\beta$ - and  $\gamma$ -proteobacterial groups  
3 (59); this comparison revealed a distant phylogenetic relationship of the *soxB* sequence of *T.*  
4 *thioparus* to those of  $\alpha$ - and  $\gamma$ -proteobacteria. Consistent with phylogenetic relationships based  
5 on 16S rRNA, the predicted SoxB sequence of *T. denitrificans* (Tbd0563) is much more similar  
6 to that of *T. thioparus* (88% identity) than to those of *P. pantotrophus*, *S. novella*, and *Psb.*  
7 *salicylatoxidans* (48-50% identity). The encoded SoxB sequences for *T. thioparus* and *P.*  
8 *pantotrophus* (GenBank CAC82470 and CAA55824) showed 50% identity to each other.

9 Other *sox* genes found in the contiguous cluster of *P. pantotrophus* were remote from  
10 each other in the *T. denitrificans* genome. These included tentatively identified copies of *soxH*  
11 (Tbd1041, Tbd1103), *soxE* (Tbd2027, Tbd2034), *soxF* (Tbd2035), and *soxW* (Tbd2117).  
12 Tbd2349 corresponded to the *soxC* (sulfite dehydrogenase) of *Pseudaminobacter* and  
13 *Paracoccus*, and the *sorA* of *Starkeya*; thus, the encoded protein may be a sulfite dehydrogenase,  
14 catalyzing AMP-independent oxidation of sulfite to sulfate. BLAST probing of the genome with  
15 the nucleotide and encoded amino acid sequences of the *soxD* of *P. pantotrophus* and *S. novella*  
16 produced no hits, so a gene corresponding to the  $\alpha$ -proteobacterial *soxD* appeared to be absent.  
17 As the proteins encoded by *soxCD* in *Paracoccus* are believed to catalyze the oxidation of  
18 sulfane-sulfur to the oxidation level of sulfite (28), an alternative system is likely present in *T.*  
19 *denitrificans*.

20 Coding functions have been ascribed to many of the  $\alpha$ -proteobacterial *sox* genes  
21 discussed above (27, 28, 63-65, 67, 68). *soxX* and *soxA* encode SoxXA, a heterodimeric *c*-type  
22 cytochrome. *soxY* and *soxZ* encode SoxYZ, the “thiosulfate-binding” Enzyme A of *P.*  
23 *pantotrophus*. *soxC* and *soxD* encode the molybdoprotein SoxC and diheme *c*-type cytochrome

1 SoxD of an  $\alpha_2\beta_2$ -heterodimeric “sulfur dehydrogenase”; SoxCD has also been shown to function  
2 as a sulfite dehydrogenase. *soxB* encodes SoxB, a sulfur-thiol esterase, identified in *P.*  
3 *pantotrophus* as Enzyme B, a protein with a dinuclear manganese center that catalyzes  
4 thiosulfate cleavage and sulfate production. *soxE* encodes a *c*-type cytochrome. *soxF* encodes a  
5 sulfide dehydrogenase/flavocytochrome-*c* oxidoreductase protein. *soxRS* are reported to have a  
6 regulatory function. *soxGH* are currently of uncertain function. Finally, the SoxFGH proteins  
7 (which are all periplasmic) were reportedly not required for lithotrophic growth of *P.*  
8 *pantotrophus* on thiosulfate, although they were induced by thiosulfate. The proteins SoxXA,  
9 SoxYZ, SoxB, and SoxCD can be reconstituted into a system catalyzing thiosulfate-, sulfite-,  
10 sulfur-, and hydrogen sulfide-dependent cytochrome *c* reduction in *P. pantotrophus* and *P.*  
11 *versutus*, although this multienzyme system may not operate in the facultative chemolithotroph  
12 *S. novella*. The *soxB* gene has also been identified in *T. thioparus*, some  $\gamma$ -proteobacteria  
13 (including *Halothiobacillus* and *Thiomicrospira*), and some phototrophic sulfur bacteria (59).

14 In *T. denitrificans*, central roles for the putative SoxXA, SoxB, and SoxYZ gene products  
15 are suggested by the clustering of the genes encoding them, but the low polypeptide sequence  
16 identity of these to the corresponding sequences in *P. pantotrophus* could mean that their  
17 biochemical functions might differ considerably from those of *P. pantotrophus*. The thiosulfate-  
18 oxidizing multienzyme system of *P. pantotrophus* (and of other  $\alpha$ -proteobacteria) is located in  
19 the periplasm, but there is considerable evidence from *T. denitrificans* (and other obligately  
20 chemolithotrophic sulfur oxidizers) that at least some reactions of thiosulfate, sulfite, and sulfide  
21 oxidation require membrane-associated processes (43). Deduction of the functions of the putative  
22 *T. denitrificans* *sox* genes solely by reference to the roles of the *sox* complexes in *Paracoccus*  
23 species must clearly be done with caution, because even if they were acquired primordially by

1 lateral gene transfer (59), they could encode significantly modified enzyme functions in extant  
2  $\alpha$ - and  $\beta$ -proteobacteria.

3 In a few bacteria that oxidize inorganic sulfur compounds, namely *T. denitrificans*,  
4 *Allochromatium vinosum*, and *C. tepidum*, a siroheme-containing sulfite reductase has been  
5 proposed to catalyze the oxidation of certain inorganic sulfur species (e.g., hydrogen sulfide or  
6 sulfane-sulfur derived from thiosulfate) to sulfite (20, 71, 86). Thus, dissimilatory sulfite  
7 reductase, which is encoded by *dsr* genes and named for its catalytic role in sulfate-reducing  
8 bacteria and archaea, is apparently used in the reverse direction for dissimilatory oxidation of  
9 sulfur compounds. Siroheme-containing sulfite reductase (with an  $\alpha_2\beta_2$  structure encoded by  
10 *dsrAB*) was previously purified from *T. denitrificans* strain DSM 807 (86). A gene cluster  
11 *dsrABEFHCMKLJOPNR* (Tbd2485-2472) occurs in *T. denitrificans* ATCC 25259 that is very  
12 similar in terms of gene sequence and organization to a *dsr* cluster in *A. vinosum* that was studied  
13 by Dahl et al. (17). A notable difference in gene organization is that *dsrS* in *T. denitrificans*  
14 (Tbd2558) is not part of the *dsr* cluster, as it is in *A. vinosum*, and that a *cysG* (siroheme  
15 synthase)-like gene is located adjacent to *dsrR* in *T. denitrificans* (Tbd2471). For the translations  
16 of most of these *dsr* genes, the degree of identity with the corresponding *A. vinosum* sequences is  
17 >55%. This is the only major cluster of *dsr* genes in the finished *T. denitrificans* sequence,  
18 which contrasts with the claim by Dahl et al. (17) that there are two *dsr* gene clusters (based on  
19 their examination of shotgun clone sequences in GenBank).

20 A noteworthy finding with regard to *dsr* genes in *T. denitrificans* was that two genes,  
21 *dsrC* and *dsrA*, are replicated multiple times in the genome, typically with no more than one or  
22 two adjacent genes that are putatively associated with sulfur-compound oxidation. Eight  
23 putative copies of *dsrC* were identified (Fig. 3): Tbd2480 (located in the large *dsr* gene cluster),



1 Tbd2488 (located next to Tbd2489, which encodes a rhodanese-related sulfurtransferase),  
2 Tbd2658, Tbd2326 and 2327 (which are adjacent to each other), Tbd1365 (located near a *dsrA*  
3 copy, Tbd1369), Tbd1408 (located next to Tbd1407, which encodes sulfide:quinone  
4 oxidoreductase), and Tbd1926. The degree of sequence identity among the corresponding  
5 predicted DsrC copies ranges from 26 to 88%. DsrC is a soluble, cytoplasmic protein whose  
6 function is not currently known (62). A cysteine residue at the C-terminus of DsrC that is highly  
7 conserved in a range of bacteria including *A. vinosum* and various sulfate-reducing bacteria (62)  
8 is present in only two of the eight copies encoded in the *T. denitrificans* genome, including the  
9 copy located in the *dsr* cluster (Tbd2480) and Tbd2658. In fact, overall, the translated DsrC  
10 from Tbd2480 was more similar to DsrC in *A. vinosum* than to the other seven copies in *T.*  
11 *denitrificans* (Fig. 3). Inasmuch as Tbd2480 is the most likely *dsrC* copy in *T. denitrificans* to  
12 code for functional DsrC, it is noteworthy that, like *dsrC* in *A. vinosum* (17, 62) and unlike all  
13 other *dsrC* copies in *T. denitrificans*, Tbd2480 appears to be constitutively expressed (based on  
14 quantitative, RT, real-time PCR results for *T. denitrificans* carrying out thiosulfate or Fe(II)  
15 oxidation under denitrifying conditions; Table 3). As shown in Table 3, Tbd2480 is relatively  
16 highly expressed when oxidizing either thiosulfate or Fe(II), whereas none of the other seven  
17 putative *dsrC* copies is highly expressed under both conditions.

18 Three putative copies of *dsrA* were identified in the *T. denitrificans* genome: Tbd2485  
19 (located in the large *dsr* gene cluster), Tbd1309, and Tbd1369 (located near a *dsrC* copy). The  
20 degree of sequence identity among the corresponding predicted DsrA copies was high (from 78  
21 to 83%). As *dsrA* codes for the  $\alpha$  subunits of the  $\alpha_2\beta_2$ -structured siroheme sulfite reductase, it is  
22 curious that two copies of *dsrA* in the *T. denitrificans* genome are not located near copies of  
23 *dsrB*, which codes for the  $\beta$  subunits; indeed, only one *dsrB* copy was identified in the genome

1 (Tbd2484). Thus, the nature of *dsr* gene duplication in *T. denitrificans* differs from that  
2 observed in *C. tepidum*, whose genome includes two copies of the *dsrCABL* cluster (17, 20).

3 *T. denitrificans* ATCC 25259 contains genes encoding all the enzymes necessary to  
4 catalyze the AMP-dependent oxidation of sulfite to sulfate (Fig. 2). All but one of these  
5 enzymes is used in the reverse direction by sulfate-reducing bacteria for the activation of sulfate  
6 to APS and the subsequent AMP-yielding reduction of APS to sulfite. APS reductase, an  $\alpha\beta$ -  
7 heterodimeric iron-sulfur flavoenzyme, is encoded by Tbd0872-0873 and catalyzes the AMP-  
8 dependent oxidation of sulfite to APS. Another pair of genes, Tbd2282-2283, also putatively  
9 encode APS reductase, but quantitative, RT, real-time PCR results suggest that the Tbd0872-  
10 0873 genes are much more highly expressed during thiosulfate oxidation (Table 3). ATP  
11 sulfurylase, which catalyzes an ATP-yielding substrate level phosphorylation that converts APS  
12 to sulfate, is encoded by Tbd0874 and also by Tbd0210 (with the former being more highly  
13 expressed; Table 3). APS:phosphate adenylyltransferase (or APAT), which is encoded by  
14 Tbd0601 and catalyzes an alternative substrate-level phosphorylation that converts APS to  
15 sulfate (yielding ADP rather than ATP), is not reversible (although it was formerly misnamed as  
16 ADP sulfurylase) (13).

17 The *T. denitrificans* genome includes other genes that are likely to play a role in sulfur-  
18 compound oxidation, including sulfide-quinone oxidoreductases (Tbd1407, Tbd2225) and  
19 rhodanese (thiosulfate-sulfurtransferase; Tbd1650, Tbd2399, Tbd2489). Genes for  
20 dimethylsulfoxide (DMSO) reductase (Tbd0570-0572) and tetrathionate reductase (Tbd1739-  
21 1741) are also present in the *T. denitrificans* genome, but these may have a role as anaerobic  
22 electron acceptors rather than in sulfur-compound oxidation. No genes for the  $\alpha$ - and  $\beta$ -subunits  
23 of DMSO dehydrogenase were detected. The type strain of *T. denitrificans* (NCIMB 9548) and

1 some strains of *T. thioparus* can oxidize and grow on thiocyanate as a sole electron donor (47),  
2 and three genes of *T. thioparus* encode the thiocyanate hydrolase enzyme that initiates  
3 thiocyanate degradation (GenBank AB007989, *scnBAC*). In contrast, *T. denitrificans* strain  
4 ATCC 25259 does not oxidize thiocyanate, and its genome lacks the genes encoding thiocyanate  
5 hydrolase.

6 **Hydrogen metabolism.** Analysis of the genome of *T. denitrificans* ATCC 25259 has  
7 revealed the presence of genes encoding two [NiFe]hydrogenases. Hydrogenases are  
8 metalloenzymes that catalyze the reversible oxidation of H<sub>2</sub> to protons and are vital components  
9 of the energy metabolism of many microbes. Notably, hydrogenases have not previously been  
10 reported in *T. denitrificans*, and the sequenced strain does not appear to be able to grow on  
11 hydrogen as a sole electron donor under denitrifying conditions (H. Beller, unpublished data);  
12 however, hydrogen oxidation appears to be required for nitrate-dependent U(IV) oxidation by *T.*  
13 *denitrificans* (5).

14 One of the hydrogenases encoded in the *T. denitrificans* genome is putatively a  
15 cytoplasmic, heterotetrameric, Group 3b [NiFe]hydrogenase [following the classification system  
16 described by Vignais et al. (87)]. The four-gene cluster (Tbd1260-1263) does not appear to be  
17 near any accessory or maturation genes encoding proteins necessary for assembly of the  
18 functional holoenzyme. Although Group 3b hydrogenases have primarily been found in  
19 hyperthermophilic archaea (87), BLASTP analysis revealed that a group of similar predicted  
20 proteins (43 to 53% identity for the four subunits) occur in *Azotobacter vinelandii*, a mesophilic,  
21  $\delta$ -proteobacterium; gene organization in *T. denitrificans* and *A. vinelandii* was also similar.  
22 Group 3b hydrogenases in the hyperthermophile *Pyrococcus furiosus* are among the better  
23 characterized and are thought to play a role in disposing of excess reductant generated during

1 fermentation (whereby NADPH can serve as the physiological electron donor for H<sub>2</sub>  
2 evolution)(87). The role of a Group 3b hydrogenase in *T. denitrificans*, if indeed this  
3 hydrogenase is expressed in functional form, is not currently known.

4         The other hydrogenase encoded in the *T. denitrificans* genome is putatively a  
5 periplasmic, Group 1 [NiFe]hydrogenase [following the classification system described by  
6 Vignais et al. (87)]. Although the sequences of the small (HynS) and large (HynL) subunits of  
7 this hydrogenase are similar to those in many bacteria and archaea, they occur in an unusual gene  
8 cluster (*hynS-isp1-isp2-hynL*; Tbd1378-1375) that has only been observed in four other microbes  
9 to date, none of which is a mesophilic, chemolithoautotrophic bacterium like *T. denitrificans*: the  
10 phototrophic sulfur bacteria *Thiocapsa roseopersicina* (66) and *A. vinosum* (16), the  
11 hyperthermophilic bacterium *Aquifex aeolicus* (GenBank NP 213658.1-213655.1), and the  
12 archaeon *Acidianus ambivalens* (53). Typically, *hynS* and *hynL* are adjacent to one another, but  
13 they are separated by two intervening genes in these species. Not only is the organization of the  
14 *hynS-isp1-isp2-hynL* gene cluster identical in *T. denitrificans* and these other four species, the  
15 sequence similarity of the four predicted proteins is also relatively high among these species  
16 (Fig. 4). The degree of sequence identity for HynS, Isp1, Isp2, and HynL in *T. denitrificans*  
17 compared to *T. roseopersicina*, *A. aeolicus*, and *A. vinosum* ranges from 26 to 68%. A  
18 phylogenetic tree of Isp1 in *T. denitrificans* ATCC 25259 and the small complement of related  
19 proteins from the GenBank nr database (BLASTP E values < 10<sup>-5</sup>) reveals similarity not only to  
20 Isp1 copies of other species but also to NarI (the  $\gamma$  subunit of cytoplasmic nitrate reductase) and  
21 Hmc5 (part of the high-molecular-weight, transmembrane, electron transport protein complex  
22 found in *Desulfovibrio vulgaris*)(Fig. 4B). Alignment of the predicted amino acid sequence of  
23 Isp1 from *T. denitrificans* with other sequences represented in Fig. 4B shows that four histidine

1 residues (located in two of the five predicted transmembrane helices in Isp1, NarI, and Hmc5)  
2 are highly conserved. Similar observations have been made for other Isp1 sequences (16, 53, 66)  
3 and Berks et al. (8) elucidated how the conserved histidines (and the two *b*-hemes that they  
4 putatively bind, one in each half of the membrane bilayer) play a role in mediating transport of  
5 electrons across the cytoplasmic membrane as part of an electron-carrying arm of a redox loop in  
6 the integral membrane proteins NarI or HyaC (a *b*-type cytochrome associated with periplasmic  
7 [NiFe]hydrogenases). Based on these and other observations, it is likely that Isp1 serves two  
8 functions: (1) to anchor HynSL on the periplasmic side of the cytoplasmic membrane, where H<sub>2</sub>  
9 can be converted to 2H<sup>+</sup> and 2 e<sup>-</sup> by HynSL, and (2) to mediate transmembrane electron transfer  
10 from HynSL to the quinone pool of electron transport chains and thereby participate in a  
11 chemiosmotic mechanism of energy conservation. Notably, in serving these functions, Isp1 may  
12 be substituting for the *b*-type cytochrome HupC/HoxZ/HupZ/HyaC that is typically found in  
13 Group 1 [NiFe]hydrogenases but which is apparently absent in the *T. denitrificans* genome. A  
14 phylogenetic tree of Isp2 in *T. denitrificans* (Fig. 4C) reveals a similarity not only to Isp2  
15 sequences of other species but also to various iron-sulfur-containing proteins such as  
16 heterodisulfide reductases, DsrK (from other sulfur-compound oxidizers *A. vinosum* and *C.*  
17 *tepidum*), and Hmc6 from *D. vulgaris*; such similarities have been noted previously for other  
18 Isp2 copies (16, 66).

19         The two adjacent gene clusters (Tbd1380-1374 and Tbd1381-1386; on opposite DNA  
20 strands) that code for the Group 1 [NiFe]hydrogenase of *T. denitrificans* include a number of  
21 genes putatively involved in biosynthesis and maturation of the hydrogenase. This clustering of  
22 maturation genes along with the *hynS-isp1-isp2-hynL* cluster in *T. denitrificans* distinguishes its  
23 gene organization from that of *T. roseopersicina* and *A. aeolicus*, which do not have accessory

1 genes in the immediate vicinity of the *hynS-isp1-isp2-hynL* cluster. In the Tbd1380-1374 gene  
2 cluster, Tbd1374 codes for a homolog of *UreJ* and is probably involved in Ni transport; *hynD*  
3 (Tbd1380) codes for a putative maturation protease. Most or all of the genes in the *hypCABDFE*  
4 gene cluster (Tbd1381-1386) encode proteins that are putatively involved with insertion of Ni,  
5 Fe, CO, and CN in the active site of the hydrogenase (87).

6 **Aerobic respiration.** The genome of *T. denitrificans*, a facultative anaerobe, encodes all  
7 the necessary machinery for aerobic respiration, including NADH:ubiquinone oxidoreductase  
8 (complex I; Tbd1142-1155), succinate dehydrogenase (complex II; Tbd1182-1185), and  
9 cytochrome *bc<sub>1</sub>*-type ubiquinol oxidoreductase (complex III; Tbd1831-1833) (9). These  
10 components can provide the reducing equivalents needed for terminal respiration with nitrate or  
11 oxygen, the latter in conjunction with one of three terminal cytochrome *c* oxidases encoded by  
12 two gene clusters on the *T. denitrificans* genome. The first cluster encodes a cytochrome *aa<sub>3</sub>*-  
13 type cytochrome *c* oxidase (Tbd0325, 0326, 0328, 0330) and a *cbb<sub>3</sub>*-type cytochrome *c* oxidase  
14 (Tbd0338-0341), whereas the second cluster (Tbd0640-0643) encodes a second *cbb<sub>3</sub>*-type  
15 cytochrome *c* oxidase. The presence of both *aa<sub>3</sub>*- and *cbb<sub>3</sub>*-type oxidases, in addition to the  
16 denitrification machinery, allows *T. denitrificans* to survive under a wide range of redox  
17 conditions; presumably, the *aa<sub>3</sub>* oxidase operates under high oxygen tension, *cbb<sub>3</sub>* oxidases  
18 operate under microaerophilic conditions, and the denitrification complex operates under  
19 anaerobic conditions (58, 61).

20 **Denitrification.** *T. denitrificans* has all necessary genes encoding the four essential  
21 enzymes that catalyze denitrification (reduction of nitrate to nitrogen gas): nitrate reductase,  
22 nitrite reductase, nitric oxide reductase, and nitrous oxide reductase (60, 94). The membrane-  
23 bound, dissimilatory nitrate reductase is encoded in a *narKK<sub>2</sub>GHJI* cluster (Tbd1401-1406),

1 whereas the NarXL two-component regulatory system is encoded on the reverse DNA strand  
2 (Tbd1400-1399). This gene organization is similar to that described for *Pseudomonas*  
3 *aeruginosa* PAO1 and *P. fluorescens* C7R12 (60). A *nir* operon including the CDS for  
4 cytochrome *cd*<sub>1</sub>-nitrite reductase (*nirS*; Tbd0077) is present in *T. denitrificans*, and in fact, this  
5 protein was purified from a different strain of *T. denitrificans* (strain DSM 807; 36). *T.*  
6 *denitrificans* ATCC 25259 contains two *nor* gene clusters that include the *norCB* structural  
7 genes encoding nitric oxide reductase, a membrane-anchored protein complex. Quantitative, RT,  
8 real-time PCR results carried out with thiosulfate-oxidizing, denitrifying cells (Table 3) indicated  
9 that the *norCB* genes Tbd0562-0561 had >50-fold higher expression levels than *norCB* genes  
10 Tbd0822-0823, strongly indicating that the former genes are of greater functional importance.  
11 The predicted NorC amino acid sequence for the less-expressed copy (Tbd0822) contains >120  
12 more residues at the C-terminus than the highly expressed copy (Tbd0562) or a similar copy in  
13 *Azoarcus* sp. EbN1 (GenBank YP\_157125); these additional amino acid residues for Tbd0822  
14 include a second CXXCH heme-binding motif. The enzyme catalyzing the final step of  
15 denitrification, nitrous oxide reductase, is associated with the structural gene *nosZ* (Tbd1389),  
16 and, like NirS, was purified from *T. denitrificans* strain DSM 807 (36). Organization of *nos*  
17 genes in *T. denitrificans* is similar, but not identical, to that described for *Ralstonia*  
18 *metallidurans* CH34 (60). As has been observed for some other  $\beta$ -proteobacteria (60), *nosR*  
19 (Tbd1390) is located downstream from *nosZ* in *T. denitrificans*.

20 **c-Type cytochromes.** *T. denitrificans* strain ATCC 25259 has a higher number of *c*-type  
21 cytochromes than most bacteria with finished genomes, based on analysis of the characteristic  
22 CXXCH heme-binding motif throughout the genome. Fifty-six genes, or approximately 2% of  
23 the total CDS, contained at least one CXXCH motif (some encoded proteins, such as DnaJ and

1 ribosomal protein L31, are not truly *c*-type cytochromes; Table 4). In the context of bacteria  
2 reported to have relatively high numbers of *c*-type cytochromes, the number in *T. denitrificans* is  
3 less than that cited for *Geobacter sulfurreducens* (111, or 3.2% of the total CDS) (54) but more  
4 than for *Shewanella oneidensis* (39, or 0.8% of the total CDS)(34) and *Pseudomonas aeruginosa*  
5 (35, or 0.6% of the total CDS)(77). Overall, *c*-type cytochromes in *T. denitrificans* range in  
6 predicted molecular mass from 9.5 to 138 kDa and in number of heme groups from one to three  
7 (with only one triheme *c*-type cytochrome and the vast majority as monoheme proteins)(Table  
8 4). Although the functions of some *c*-type cytochromes in *T. denitrificans* can be confidently  
9 predicted, such as NirS and NorC, some are of less certain or unknown function (Table 4). The  
10 ability of *T. denitrificans* to catalyze anaerobic, nitrate-dependent oxidation of metals with high  
11 reduction potentials, such as Fe(II) or U(IV) (5, 78), may be mediated by *c*-type cytochromes, as  
12 few other electron carriers in a bacterial cell would have sufficiently high reduction potentials to  
13 accept electrons from compounds such as uraninite, or UO<sub>2</sub> (the UO<sub>2</sub><sup>2+</sup>/UO<sub>2</sub> couple has an E<sub>0</sub>'  
14 value of +0.26V; 5).

15 **Autotrophy.** The genome of this obligate chemolithoautotroph encodes both Form I and  
16 Form II ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCO) enzymes for CO<sub>2</sub> fixation  
17 (21, 35). The Form I genes (*cbbL* and *cbbS*) occur in an operon with *cbbQ* and *cbbO* genes  
18 (*cbbLSQO*; Tbd2624-2621). The Form II gene, *cbbM*, is in a separate operon that also includes  
19 *cbbQ* and *cbbO* genes; the operon has the form *cbbMQO* (Tbd2638-2636). Both operons are  
20 preceded by divergently transcribed *cbbR* genes encoding LysR-type transcriptional regulators  
21 (Tbd2625 and 2639). The *cbbQ* and *cbbO* genes encode proteins involved in the post-  
22 translational activation of RuBisCO (33). The two copies of each gene are quite distinct: the two  
23 CbbQ proteins have 71% identical residues, whereas the two CbbO proteins only share 34%



1 identity. The two RuBisCO operons, although separate, fall within a 33-kb region that also  
2 includes an operon of ten genes (Tbd2641-2650) encoding the carboxysome shell proteins (14).  
3 The carboxysome operon does not begin with *cbbLS*, unlike other known examples (14). These  
4 three operons, all located on the reverse strand, are oriented to be transcribed in the same  
5 direction. The carboxysome operon includes the gene (Tbd2649) for carbonic anhydrase (CA)  
6 epsilon (74). The carboxysome operon is also preceded by a divergently transcribed *cbbR* gene  
7 (Tbd2651) encoding a LysR-type transcriptional regulator. In addition to the epsilon-type CA, a  
8 eukaryotic-type CA (Tbd2167) is encoded elsewhere in the genome. Since there is no evidence  
9 that the carboxysome operon is functional (14, 72), this alternative CA may be the primary  
10 source of this activity.

11 Inasmuch as *T. denitrificans* can grow under both aerobic and denitrifying conditions,  
12 and Form I and Form II RuBisCO in this species were shown to have markedly different abilities  
13 to discriminate between CO<sub>2</sub> and O<sub>2</sub> (35), it is possible that Form I and II RuBisCO are  
14 differentially expressed in *T. denitrificans* based on the concentration of O<sub>2</sub>. Molecular oxygen  
15 competes with CO<sub>2</sub> for the active site of RuBisCO and thereby decreases its efficiency for  
16 carbon fixation. Hence, Form I, which has a higher CO<sub>2</sub>/O<sub>2</sub> specificity, should be more highly  
17 expressed under aerobic conditions, whereas Form II should be more highly expressed under  
18 anaerobic conditions.

19 Genes for all enzymes to complete the Calvin-Benson-Bassham (CBB) cycle are present.  
20 Transketolase, NAD-dependent glyceraldehyde 3-phosphate dehydrogenase (E.C.1.2.1.12),  
21 phosphoglycerate kinase, pyruvate kinase, and fructose 1,6-bisphosphate aldolase (Tbd0159-  
22 0163) are encoded by an operon, whereas fructose 1,6-/sedoheptulose-1,7- bisphosphatase  
23 (Tbd2577), ribose 5-phosphate isomerase (Tbd2364), and phosphoribulokinase (Tbd2447) are

1 encoded by isolated genes. Genes encoding ribulose 5-phosphate 3-epimerase and  
2 phosphoglycolate phosphatase (Tbd2230-2229) are adjacent but may not be co-transcribed.

3         **Central intermediary metabolism.** Organic storage materials have not been reported in  
4 *T. denitrificans*, but the presence of genes encoding glycogen synthase (Tbd2057), maltooligosyl  
5 trehalose synthase (Tbd1174), and various glucan branching enzymes (e.g., an  $\alpha$ -1,4-D-glucan  
6 branching enzyme; Tbd1173, Tbd2058) and glucano- and glycosyl-transferases suggests that the  
7 bacterium synthesizes a polyglucose storage product (cf. *Halothiobacillus neapolitanus*; 10). The  
8 gene for the key enzyme of the Embden-Meyerhof-Parnas (EMP) pathway necessary for it to  
9 effect gluconeogenesis (fructose 1,6-bisphosphatase) is present in the genome (Tbd2577). The  
10 synthesis of such a storage product would provide a rationale for the presence of genes for  
11 glucokinase (Tbd2062, Tbd2216), which could thus assist in the endogenous catabolism of  
12 stored polyglucose. Several genes encoding enzymes for polysaccharide hydrolysis also appear  
13 to be present:  $\beta$ -galactosidase (Tbd2429), several glycosidases (Tbd0727, 0923, 1172) and  
14 glycosyl transferases (e.g., Tbd0289, 0293, 0294, 0301, 2139), an  $\alpha$ -mannosidase (Tbd2060),  
15 and an  $\alpha$ -arabinofuranosidase (Tbd1789).

16         Genes encoding all the enzymes of the EMP pathway necessary for it to function in the  
17 forward direction for the conversion of glucose to pyruvate are present in the genome, including  
18 the gene for phosphofructokinase (Tbd1502), which is unique to the degradative EMP pathway.  
19 Genes for all the enzymes of the oxidative pentose phosphate (OPP) pathway for the oxidation of  
20 glucose to carbon dioxide are also present in the genome. Some of the enzymes of the EMP and  
21 OPP pathways are also common to the CBB reductive pentose phosphate cycle for carbon  
22 dioxide fixation, and ribose 5-phosphate for nucleic acid synthesis can be produced from glucose  
23 6-phosphate by glucose 6-phosphate dehydrogenase, phosphogluconolactonase, and 6-

1 phosphogluconate dehydrogenase (encoded by Tbd2121-2123). In addition, genes encoding  
2 alcohol dehydrogenase (EC 1.1.1.1; Tbd1767) and other short-chain alcohol dehydrogenases  
3 (Tbd0924, 1469, 1549, 1699, 1886, 2701, 2756), lactate dehydrogenase (Tbd1998), and  
4 phosphoketolase (EC 4.1.2.9) (Tbd0049, 0831) suggest that *T. denitrificans* might be able to  
5 produce ethanol and lactate (by homo- or hetero-fermentative metabolism) from endogenous  
6 glucose under anoxic conditions. This would parallel the heterolactic fermentation of stored  
7 polyglucose carried out anaerobically by *Halothiobacillus neapolitanus* (11). Glucose 6-  
8 phosphatase and glycerol phosphatase are absent from the *T. denitrificans* genome, reflecting the  
9 lack of need to produce free glucose or glycerol.

10 Genes for some enzymes of the Entner-Doudoroff (ED) pathway are present, but notably  
11 absent is the gene encoding 2-keto-3-deoxy-6-phosphogluconate (KDPG) aldolase, meaning that  
12 *T. denitrificans* cannot express a functional ED pathway. We failed to detect KDPG aldolase by  
13 BLAST searching of the genome using database polypeptide sequences from *Zymomonas*,  
14 *Escherichia coli*, *Gluconobacter*, and *Neisseria*. The gene putatively encoding 6-  
15 phosphogluconate (6-PG) dehydratase (Tbd2730), whose product is KDPG, may be a false  
16 identification, as a TBLASTN search of the *T. denitrificans* genome with the 6-PG dehydratase  
17 protein sequence of *Xanthomonas* (GenBank NP\_642389) failed to detect any matching  
18 sequences, although direct BLAST2(P) comparison of the polypeptide encoded by Tbd2730 did  
19 show low identity (27–29%) to the 6-PG dehydratases of *Xanthomonas* and *Helicobacter*  
20 (GenBank NP\_223740). The highest identities of Tbd2730 indicated by BLAST analyses were to  
21 dihydroxyacid dehydratases, and comparisons by BLASTP of the polypeptide sequences of the  
22 dihydroxyacid dehydratase of *Sinorhizobium* (GenBank AL591792) with the 6-PG dehydratase

1 of *Xanthomonas* also showed 29% identity. Assigning putative function must thus be done with  
2 caution.

3 Genes for all the enzymes of the Krebs tricarboxylic acid cycle were identified in the  
4 genome. The E1 and E2 subunits of 2-oxoglutarate dehydrogenase are encoded by Tbd1188 and  
5 Tbd1189, whereas the E3 subunit (common also to pyruvate dehydrogenase) is encoded by  
6 Tbd0652. Genes encoding the E1, E2 and E3 subunits of pyruvate dehydrogenase are Tbd0652,  
7 0654, and 0655, whereas Tbd1847 putatively encodes all three subunits, with highest identity to  
8 the *pdhA*, *pdhB*, and *pdhL* genes of *Ralstonia eutropha*. It was surprising to find genes for the E1  
9 and E2 subunits of 2-oxoglutarate dehydrogenase, as *T. denitrificans* strains (including ATCC  
10 25259) do not express an active 2-oxoglutarate dehydrogenase enzyme when growing  
11 autotrophically (57, 83). This inability is shared with other obligate chemolithotrophs and  
12 methanotrophs, including *Nitrosomonas europaea* and *Methylococcus capsulatus*, and has been  
13 proposed as a contributory factor in the obligate growth modes of these bacteria (15, 41, 48, 73,  
14 89, 90, 92). The E3 component of the pyruvate and 2-oxoglutarate dehydrogenases is also  
15 known to be controlled by multiple regulatory mechanisms (93), so it may be that failure of  
16 obligate chemolithotrophs to express active 2-oxoglutarate dehydrogenase results from  
17 regulation at the E3 gene expression level.

18 While identifying genes encoding some Krebs cycle enzymes, it was found that a gene  
19 (Tbd2119) showed significant identity to both fumarase (fumarate hydratase; EC 4.2.1.2) and  
20 aspartate ammonia-lyase (EC 4.3.1.1). BLASTP analysis showed the translated sequence of  
21 Tbd2119 to share 55% sequence identity with the aspartate ammonia-lyase of *Geobacter*  
22 *sulfurreducens* (GenBank NP\_951538.1), 55% identity with the fumarase sequence of  
23 *Pelobacter propionicus* (GenBank ZP\_00677090.1), and 41% with the fumarase C of *E. coli*. We

1 suggest that the role of the Tbd2119 gene product is as a fumarase in the Krebs cycle. This  
2 similarity of genes for fumarase and aspartate ammonia-lyase is common across the database  
3 sequences for numerous organisms, possibly indicating multifunctional roles for the genes or  
4 their encoded proteins.

5 Genes encoding isocitrate lyase and malate synthase have not been detected using  
6 BLASTP searches with the translated polypeptide sequences of the genes from *E. coli*, meaning  
7 that the glyoxylate cycle cannot be present and hence exogenous acetate could not be  
8 metabolized as a sole source of carbon by that route.

9 **Transport systems for organic nutrients.** Permease systems for inorganic ions and  
10 numerous ABC transporter components are encoded within the genome, but relatively few  
11 specific systems have been detected for the uptake of sugars or organic acids. Those detected  
12 include various components constituting a tripartite ATP-independent periplasmic (TRAP)-type  
13 C<sub>4</sub>-dicarboxylate transporter system (Tbd0466, 0467, 0468, 2151, 2164), and phosphotransferase  
14 system components including Enzyme I (Tbd2414), HPr (Tbd2413), two Enzyme IIA subunits  
15 specific for mannose/fructose (Tbd2412, Tbd0531), and Hpr(ser) kinase/phosphorylase  
16 (Tbd0530). The absence from the *T. denitrificans* genome of a gene encoding Enzyme IIC (the  
17 sugar permease component) suggests that the role of this PTS system may be regulatory rather  
18 than for sugar transport (30). The genome of *T. denitrificans* also encodes two sodium:solute  
19 symporter family proteins that may be involved in acetate uptake (Tbd0088, 0212). A  
20 sodium:solute symporter protein, ActP, has been shown to be involved in acetate uptake in *E.*  
21 *coli* (GenBank P32705). A functional transport system for acetate in *T. denitrificans* was  
22 indicated by the uptake of <sup>14</sup>C-labeled acetate into bacteria growing chemolithotrophically  
23 [acetate provided 6-11% of the cell-carbon of strain ATCC 25259 (84)]. As well as being used

1 for lipid biosynthesis,  $^{14}\text{C}$ -acetate was incorporated only into the protein-amino acids glutamate,  
2 proline, and arginine, as reported for other obligate chemolithotrophs and methanotrophs that  
3 lack 2-oxoglutarate dehydrogenase (19, 41, 73, 92). Incorporation of acetate-carbon into  
4 glutamate by *T. denitrificans* was unaffected by exogenous glutamic acid, which was presumably  
5 not taken up significantly by the bacteria. The possibility remains to be tested that  
6 chemolithoorganotrophic growth by *T. denitrificans* might be possible if the organism were  
7 presented with compounds such as acetate, or a suitable  $\text{C}_4$ -dicarboxylic acid, in the presence of  
8 thiosulfate and nitrate as the energy source. This would be akin to the chemolithoorganotrophic  
9 growth of *Nitrosomonas europaea* on fructose or pyruvate and ammonia (15, 37).

10 **Inorganic ion transport and heavy metal resistance.** A number of genes were  
11 identified in *T. denitrificans* that putatively code for transporters that can mediate either the  
12 uptake or efflux of a range of inorganic ions. In all, at least 18 complete ABC (ATP Binding  
13 Cassette) transporters predominantly for inorganic molecules are present in the genome allowing  
14 for the uptake of  $\text{Fe}^{3+}$ , thiosulfate, nitrate, nitrite, and many other ions (18). Numerous non-ABC  
15 type transporters allowing for uptake of other ions, such as various sulfur-containing compounds  
16 and bicarbonate, are also present. The genome of *T. denitrificans* also contains a surprisingly  
17 large number of metal resistance systems, particularly considering its relatively small genome  
18 size (Table 5). In total, the *T. denitrificans* genome encodes as many as 17 possible metal  
19 resistance systems [described by Nies (55)] including five heavy metal efflux (HME) systems  
20 from the resistance-nodulation-cell division (RND) family of transporters; three cation diffusion  
21 facilitators (CDF); three CPx-type ATPases (heavy-metal specific P-type ATPases); and five  
22 additional gene clusters encoding possible resistance systems specific for metals such as  $\text{Ni}^{2+}$   
23 (*nreB*),  $\text{Pb}^{2+}$  (*pbrT*),  $\text{Hg}^{2+}$  (3), chromium (as chromate; *chrA*), and  $\text{Cu}^{2+}/\text{Ag}^{2+}$  (26, 32). Although

1 the *T. denitrificans* genome has fewer systems than the model metal-resistant bacterium  
2 *Ralstonia metallidurans* (55), it has more than most other bacteria characterized to date (Table  
3 5). Notably, *T. denitrificans* (or species with >98% 16S rDNA sequence similarity) was found to  
4 be prevalent in an inactive uranium mine with relatively high concentrations of uranium, nickel,  
5 cobalt, and zinc (80).

6 Several of the efflux-mediated heavy metal resistance systems, along with various  
7 systems involved in metal uptake and storage, are found in large gene clusters on the *T.*  
8 *denitrificans* genome (Fig. 1). The largest of these metal transport clusters, Tbd0704-0726,  
9 encodes proteins allowing for high-affinity Fe<sup>3+</sup> acquisition, Fe<sup>3+</sup> storage, Pb<sup>2+</sup> resistance, and  
10 heavy metal resistance. Genes involved in high-affinity Fe<sup>3+</sup> acquisition include homologs of a  
11 portion of the *Vibrio parahaemolyticus* polyhydroxycarboxylate-type siderophore biosynthesis,  
12 secretion, and uptake gene cluster *pvuApvsABCDE psuA* (Tbd0722-0717, Tbd0715) (81), and  
13 *tonBexbBD* (Tbd0713-0711), which allows for active transport across the outer membrane of  
14 Fe<sup>3+</sup>-bound siderophore. Also found are genes encoding the Fe<sup>3+</sup> storage proteins bacterioferritin  
15 and bacterioferritin-associated ferredoxin (Tbd0704-0705), Pb<sup>2+</sup> resistance (Tbd0723), and CDF  
16 family heavy metal resistance (Tbd0726). A second large gene cluster (encompassing genes on  
17 both the forward and reverse strands) encodes proteins primarily associated with metal resistance  
18 (Tbd1324-1341). This cluster includes a multi-copper oxidase (Tbd1324) (32), a periplasmic  
19 Cu<sup>2+</sup>-binding protein (Tbd1326) (26), two HME-RND systems (Tbd1327-1329, Tbd1333-1335),  
20 and a Hg<sup>2+</sup> resistance system *merRTPA* (Tbd1338-1341) (3).

21 **Future prospects.** The availability of the *T. denitrificans* genome has fostered new  
22 insights into this bacterium and will effectively focus further investigations into its biochemistry  
23 and physiology. Among the new insights reported here are the following genomic

1 characteristics: an unusually large number of genes encoding *c*-type cytochromes, a relatively  
2 large complement of genes associated with inorganic ion transport and heavy metal resistance,  
3 and the presence of genes encoding two [NiFe]hydrogenases, which is particularly significant  
4 because no physiological, biochemical, or genetic information on hydrogenases has previously  
5 been reported for this species. The genome also provides much more information on the genetic  
6 basis of sulfur-compound oxidation in chemolithotrophic  $\beta$ -proteobacteria (particularly with  
7 respect to *sox* and *dsr* genes). Much more work will be required to understand the genetic and  
8 biochemical basis of unusual and enigmatic metabolic capabilities of this bacterium, including  
9 the coupling of denitrification with sulfur-compound oxidation, the use of mineral electron  
10 donors, and the anaerobic, nitrate-dependent oxidation of metals. Whole-genome transcriptional  
11 studies with cDNA microarrays are currently underway to address such questions.

## 12 ACKNOWLEDGMENTS

13 We thank A. Lapidus and C. Detter (DOE Joint Genome Institute), M. Land (Oak Ridge  
14 National Laboratory), and M. Shin, T. Legler, F. Bourguet, W. Marrs, and A. Krefft (Lawrence  
15 Livermore National Laboratory) for their valuable contributions.

16 This work was performed under the auspices of the U.S. Department of Energy by the  
17 University of California, Lawrence Livermore National Laboratory under contract No. W-7405-  
18 Eng-48.

## 19 REFERENCES

- 20 1. **Altschul, S. F., T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D.**  
21 **J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein  
22 database search programs. *Nucleic Acids Res.* **25**:3389-3402.



- 1 2. **Aminuddin, M., and D. J. D. Nicholas.** 1974. An AMP-independent sulphite oxidase  
2 from *Thiobacillus denitrificans*: purification and properties. J. Gen. Microbiol. **82**:103-  
3 113.
- 4 3. **Barkay, T., S. M. Miller, and A. O. Summers.** 2003. Bacterial mercury resistance from  
5 atoms to ecosystems. FEMS Microbiol. Rev. 27:355-384. FEMS Microbiol. Rev. 27:355-  
6 384.
- 7 4. **Beijerinck, M. W.** 1904. Phénomènes de réduction produits par les microbes  
8 (Conférence avec démonstrations faite – Delft, le 16 avril 1903). Archs. Neerl. Sci. Ser. 2  
9 **9**:131-157.
- 10 5. **Beller, H. R.** 2005. Anaerobic, nitrate-dependent oxidation of U(IV) oxide minerals by  
11 the chemolithoautotrophic bacterium *Thiobacillus denitrificans*. Appl. Environ.  
12 Microbiol. **71**:2170-2174.
- 13 6. **Beller, H. R., V. Madrid, G. B. Hudson, W. W. McNab, and T. Carlsen.** 2004.  
14 Biogeochemistry and natural attenuation of nitrate in groundwater at an explosives test  
15 facility. Appl. Geochem. **19**:1483-1494.
- 16 7. **Bendtsen, J. D., H. Nielsen, G. von Heijne, and S. Brunak.** 2004. Improved prediction  
17 of signal peptides: SignalP 3.0. J. Mol. Biol. **340**:783-795.
- 18 8. **Berks, B. C., M. D. Page, D. J. Richardson, A. Reilly, A. Cavill, F. Outen, and S. J.**  
19 **Ferguson.** 1995. Sequence analysis of subunits of the membrane-bound nitrate reductase  
20 from a denitrifying bacterium: the integral membrane subunit provides a prototype for the  
21 dihaem electron-carrying arm of a redox loop. Mol. Microbiol. **15**:319-331.
- 22 9. **Berlyn, M. K. B., K. B. Low, and K. E. Rudd.** 1996. Linkage map of *Escherichia coli*  
23 K12, edition 9, p. 1715-1902. In F.C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C.

- 1 Lin, et al. (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2<sup>nd</sup> ed.,  
2 vol. 2, ASM Press, Washington, DC.
- 3 10. **Beudeker, R. F., J. W. M. Kerver, and J. G. Kuenen.** 1981a. Occurrence, structure  
4 and function of intracellular polyglucose in the obligate chemolithotroph *Thiobacillus*  
5 *neapolitanus*. Arch. Microbiol. **129**: 221-226.
- 6 11. **Beudeker, R. F., W. de Boer, and J. G. Kuenen.** 1981b. Heterolactic fermentation of  
7 intracellular polyglucose by the obligate chemolithotroph *Thiobacillus neapolitanus*  
8 under anaerobic conditions. FEMS Microbiol. Lett. **12**: 337-342.
- 9 12. **Bowen, T. J., F. C. Happold, and B. F. Taylor.** 1966. Studies on the adenosine-5'-  
10 phosphosulfate reductase from *Thiobacillus denitrificans*. Biochim. Biophys. Acta  
11 **118**:566-576.
- 12 13. **Brüser, T., T. Selmer, and C. Dahl.** 2000. "ADP sulfurylase" from *Thiobacillus*  
13 *denitrificans* is an adenylylsulfate:phosphate adenylyltransferase and belongs to a new  
14 family of nucleotidyltransferases. J. Biol. Chem. **275**:1691-1698.
- 15 14. **Cannon, G. C., S. H. Baker, F. Soyer, D. R. Johnson, C. E. Bradburne, J. L.**  
16 **Mehlman, P. S. Davies, Q. L. Jiang, S. Heinhorst, J. M. Shively.** 2003. Organization  
17 of carboxysome genes in the thiobacilli. Curr. Microbiol. **46**:115-119.
- 18 15. **Chain, P., J. Lamerdin, F. Larimer, W. Regala, V. Lao, M. Land, L. Hauser, A.**  
19 **Hooper, M. Klotz, J. Norton, L. Sayavedra-Soto, D. Arciero, N. Hommes, M.**  
20 **Whittaker, and D. Arp.** 2003. Complete genome sequence of the ammonia-oxidizing  
21 bacterium and obligate chemolithoautotroph *Nitrosomonas europaea*. J. Bacteriol.  
22 **185**:2759-73.

- 1 16. **Dahl, C., G. Rakhely, A. S. Pott-Sperling, B. Fodor, M. Takacs, A. Toth, M.**  
2 **Kraeling, K. Gyorfi, A. Kovacs, J. Tusz, and K. L. Kovacs.** 1999. Genes involved in  
3 hydrogen and sulfur metabolism in phototrophic sulfur bacteria. *FEMS Microbiol. Lett.*  
4 **180**:317-324.
- 5 17. **Dahl, C., S. Engels, A. S. Pott-Sperling, A. Schulte, J. Sander, Y. Lübbe, O. Deuster,**  
6 **and D. C. Brune.** 2005. Novel genes of the *dsr* gene cluster and evidence for close  
7 interaction of Dsr proteins during sulfur oxidation in the phototrophic sulfur bacterium  
8 *Allochromatium vinosum*. *J. Bacteriol.* **187**:1392-1404.
- 9 18. **Dassa, E., and P. Bouige.** 2001. The ABC of ABCs: a phylogenetic and functional  
10 classification of ABC systems in living organisms. *Res. Microbiol.* **152**: 211-229.
- 11 19. **Eccleston M., and D. P. Kelly.** 1973. Assimilation and toxicity of some exogenous C-1  
12 compounds, alcohols, sugars and acetate in the methane-oxidizing bacterium  
13 *Methylococcus capsulatus*. *J Gen. Microbiol.* **75**:211-221.
- 14 20. **Eisen, J. A., K. E. Nelson, I. T. Paulsen, J. F. Heidelberg, M. Wu, R. J. Dodson, R.**  
15 **Deboy, M. L. Gwinn, W. C. Nelson, D. H. Haft, E. K. Hickey, J. D. Peterson, A. S.**  
16 **Durkin, J. L. Kolonay, F. Yang, I. Holt, L. A. Umayam, T. Mason, M. Brenner, T. P.**  
17 **Shea, D. Parksey, W. C. Nierman, T. V. Feldblyum, C. L. Hansen, M. B. Craven, D.**  
18 **Radune, J. Vamathevan, H. Khouri, O. White, T. M. Gruber, K. A. Ketchum, J. C.**  
19 **Venter, H. Tettelin, D. A. Bryant, and C. M. Fraser.** 2002. The complete genome  
20 sequence of *Chlorobium tepidum* TLS, a photosynthetic, anaerobic, green-sulfur  
21 bacterium. *Proc. Natl. Acad. Sci. USA* **99**:9509-9514.

- 1 21. **English, R. S., C. A. Williams, S. C. Lorbach, and J. M. Shively.** 1992. Two forms of  
2 ribulose 1,5-bisphosphate carboxylase/oxygenase from *Thiobacillus denitrificans*. FEMS  
3 Microbiol. Lett. **73**:111-9.
- 4 22. **Ewing, B., and P. Green.** 1998. Basecalling of automated sequencer traces using phred.  
5 II. Error probabilities. Genome Research **8**:186-194.
- 6 23. **Ewing, B., L. Hillier, M. Wendl, and P. Green.** 1998. Basecalling of automated  
7 sequencer traces using phred. I. Accuracy assessment. Genome Research **8**:175-185.
- 8 24. **Felsenstein, J.** 1989. PHYLIP - Phylogeny Inference Package (version 3.2). Cladistics  
9 **5**:164-166.
- 10 25. **Felsenstein, J.** 2002. PHYLIP (Phylogeny Inference Package) version 3.6a3. Distributed  
11 by the author. Department of Genome Sciences, University of Washington, Seattle.
- 12 26. **Franke, S., G. Grass, C. Rensing, and D. H. Nies.** 2003. Molecular analysis of the  
13 copper-transporting efflux system CusCFBA of *Escherichia coli*. J. Bacteriol. **185**: 3804-  
14 3812.
- 15 27. **Friedrich, C. G., A. Quentmeier, F. Bardichewsky, D. Rother, R. Kraft, S. Kostka,**  
16 **and H. Prinz.** 2000. Novel genes coding for lithotrophic sulfur oxidation of *Paracoccus*  
17 *pantotrophus* GB17. J Bacteriol. **182**:4677-4687.
- 18 28. **Friedrich, C. G., D. Rother, F. Bardichewsky, A. Quentmeier, and J. Fischer.** 2001.  
19 Oxidation of reduced inorganic sulfur compounds by bacteria: emergence of a common  
20 mechanism? Appl. Environ. Microbiol. **67**:2873-2882.
- 21 29. **Fritz, G., T. Büchert, H. Huber, K. O. Stetter, and P. M. H. Kroneck.** 2000.  
22 Adenylylsulfate reductases from archaea and bacteria are 1:1  $\alpha\beta$ -heterodimeric iron-

- 1 sulfur flavoenzymes – high similarity of molecular properties emphasizes their central  
2 role in sulfur metabolism. FEBS Letters **473**:63-66.
- 3 30. **Gonzalez, C. F., A. J. Stonestrom, G. L. Lorca, and M. H. Saier.** 2005. Biochemical  
4 characterization of phosphoryl transfer involving HPr of the phosphoenolpyruvate-  
5 dependent phosphotransferase system in *Treponema denticola*, an organism that lacks  
6 PTS permeases. Biochemistry **44**:598-608.
- 7 31. **Gordon, D., C. Abajian, and P. Green.** 1998. Consed: a graphical tool for sequence  
8 finishing. Genome Research **8**:195-202.
- 9 32. **Grass, G., and C. Rensing.** 2001. CueO is a multi-copper oxidase that confers copper  
10 tolerance in *Escherichia coli*. Biochem. Biophys. Res. Commun. **286**:902-908.
- 11 33. **Hayashi, N. R., H. Arai, T. Kodama, and Y. Igarashi.** 1997. The novel genes, *cbbQ*  
12 and *cbbO*, located downstream from the RubisCO genes of *Pseudomonas*  
13 *hydrogenothermophila*, affect the conformational states and activity of RubisCO.  
14 Biochem. Biophys. Res. Commun. **241**: 565–569.
- 15 34. **Heidelberg, J. F., I. T. Paulsen, K. E. Nelson, E. J. Gaidos, W. C. Nelson, T. D. Read,**  
16 **J. A. Eisen, R. Seshadri, N. Ward, B. Methe, R. A. Clayton, T. Meyer, A. Tsapin, J.**  
17 **Scott, M. Beanan, L. Brinkac, S. Daugherty, R. T. DeBoy, R. J. Dodson, A. S.**  
18 **Durkin, D. H. Haft, J. F. Kolonay, R. Madupu, J. D. Peterson, L. A. Umayam, O.**  
19 **White, A. M. Wolf, J. Vamathevan, J. Weidman, M. Impraim, K. Lee, K. Berry, C.**  
20 **Lee, J. Mueller, H. Khouri, J. Gill, T. R. Utterback, L. A. McDonald, T. V.**  
21 **Feldblyum, H. O. Smith, J. C. Venter, K. H. Nealson, and C. M. Fraser.** 2002.  
22 Genome sequence of the dissimilatory metal ion-reducing bacterium *Shewanella*  
23 *oneidensis*. Nature Biotechnol. **20**: 1118-1123.

- 1 35. **Hernandez, J. M., S. H. Baker, S. C. Lorbach, J. M. Shively, and F. R. Tabita.** 1996.  
2 Deduced amino acid sequence, functional expression, and unique enzymatic properties of  
3 the Form I and Form II ribulose biphosphate carboxylase/oxygenase from the  
4 chemoautotrophic bacterium *Thiobacillus denitrificans*. J. Bacteriol. **178**: 347-356.
- 5 36. **Hole, U. H., K.-U. Vollack, W. G. Zumft, E. Eisenmann, R. A. Siddiqui, B.**  
6 **Friedrich, P. M. H. Kroneck.** 1996. Characterization of the membranous denitrification  
7 enzymes nitrite reductase (cytochrome *cd<sub>1</sub>*) and copper-containing nitrous oxide  
8 reductase from *Thiobacillus denitrificans*. Arch. Microbiol. **165**:55-61.
- 9 37. **Hommes, N. G., L. A. Sayavedra-Soto, and D. J. Arp.** 2003. Chemolithoorganotrophic  
10 growth of *Nitrosomonas europaea* on fructose. J. Bacteriol. **185**: 6809-6814.
- 11 38. **Jeanmougin, F., J. D. Thompson, M. Gouy, D. G. Higgins, and T. J. Gibson.** 1998.  
12 Multiple sequence alignment with Clustal X. Trends Biochem Sci. **23**:403-5.
- 13 39. **Kappler, U., C. G. Friedrich, H. G. Trüper, and C. Dahl.** 2001. Evidence for two  
14 pathways of thiosulfate oxidation in *Starkeya novella* (formerly *Thiobacillus novellus*).  
15 Arch. Microbiol. **175**:102-111.
- 16 40. **Karlin S., J. Mrazek, and A. M. Campbell.** 1997. Compositional biases of bacterial  
17 genomes and evolutionary implications. J. Bacteriol. **179**:3899-3913.
- 18 41. **Kelly, D. P.** 1967. The incorporation of acetate by the chemoautotroph *Thiobacillus*  
19 *neapolitanus* strain C. Arch. Mikrobiol. **58**: 99-116.
- 20 42. **Kelly, D. P.** 1982. Biochemistry of the chemolithotrophic oxidation of inorganic sulphur.  
21 Phil. Trans. Roy. Soc. London **B298**:499-528.

- 1 43. **Kelly, D. P.** 1989. Physiology and biochemistry of unicellular sulfur bacteria, p. 193-217.  
2 *In* H. G. Schlegel and B. Bowien (ed.), *Biology of autotrophic bacteria*, Science Tech  
3 Publishers, Madison, Wisconsin.
- 4 44. **Kelly, D. P.** 1990. Energetics of chemolithotrophic bacteria, p. 479-503. *In* T. A.  
5 Krulwich (ed.), *Bacterial energetics*, Academic Press, San Diego.
- 6 45. **Kelly, D. P.** 1999. Thermodynamic aspects of energy conservation by chemolithotrophic  
7 sulfur bacteria in relation to the sulfur oxidation pathways. *Arch. Microbiol.* **171**:219-  
8 229.
- 9 46. **Kelly, D. P., and A. P. Harrison.** 1989. Genus *Thiobacillus*, p. 1842-1858. *In* J. T.  
10 Staley, M. P. Bryant, N. Pfennig, and J. G. Holt (ed.), *Bergey's Manual of Systematic*  
11 *Bacteriology*, 1<sup>st</sup> edition, vol. 3. Williams & Wilkins, Baltimore.
- 12 47. **Kelly, D. P., and A. P. Wood.** 2000. Confirmation of *Thiobacillus denitrificans* as a  
13 species of the genus *Thiobacillus*, in the beta-subclass of the *Proteobacteria*, with strain  
14 NCIMB 9548 as the type strain. *Int. J. Syst. Evol. Microbiol.* **50**:547-550.
- 15 48. **Kelly, D. P., and A. P. Wood.** 2000. The chemolithotrophic prokaryotes. *In* *The*  
16 *prokaryotes*, an evolving electronic resource for the microbiological community  
17 (PKIIIIE), M. Dworkin, Falkow S, Rosenberg E, Schleifer K-H, and E. Stackebrandt.  
18 Springer-Verlag, New York: (<http://www.prokaryotes.com>).
- 19 49. **Kelly, D. P., A. P. Wood, and E. Stackebrandt.** 2005. Genus II. *Thiobacillus*  
20 Beijerinck, p. 764-769. *In* *Bergey's Manual of Systematic Bacteriology*, 2<sup>nd</sup> edition, vol.  
21 2, part C, Springer, New York.
- 22 50. **Kelly, D. P., J. K. Shergill, W-P. Lu, and A. P. Wood.** 1997. Oxidative metabolism of  
23 inorganic sulfur compounds by bacteria. *Antonie van Leeuwenhoek* **71**:95-107.

- 1 51. **Kobayashi, I.** 2001. Behavior of restriction-modification systems as selfish mobile  
2 elements and their impact on genome evolution. *Nucleic Acids Res.* **29**:3742-56.
- 3 52. **Krogh, A., B. Larsson, G. von Heijne, and E. L. L. Sonnhammer.** 2001. Predicting  
4 transmembrane protein topology with a hidden Markov model: Application to complete  
5 genomes. *J. Mol. Biol.* **305**:567-580.
- 6 53. **Laska, S., F. Lottspeich, and A. Kletzin.** 2003. Membrane-bound hydrogenase and  
7 sulfur reductase of the hyperthermophilic and acidophilic archaeon *Acidianus*  
8 *ambivalens*. *Microbiology* **149**:2357-2371.
- 9 54. **Methé, B. A., K. E. Nelson, J. A. Eisen, I. T. Paulsen, W. Nelson, J. F. Heidelberg, D.**  
10 **Wu, M. Wu, N. Ward, M. J. Beanan, R. J. Dodson, R. Madupu, L. M. Brinkac, S. C.**  
11 **Daugherty, R. T. DeBoy, A. S. Durkin, M. Gwinn, J. F. Kolonay, S. A. Sullivan, D.**  
12 **H. Haft, J. Selengut, T. M. Davidsen, N. Zafar, O. White, B. Tran, C. Romero, H. A.**  
13 **Forberger, J. Weidman, H. Khouri, T. V. Feldblyum, T. R. Utterback, S. E. Van**  
14 **Aken, D. R. Lovley, and C. M. Fraser.** 2003. Genome of *Geobacter sulfurreducens*:  
15 metal reduction in subsurface environments. *Science* **302**: 1967-1969.
- 16 55. **Nies, D. H.** 2003. Efflux-mediated heavy metal resistance in prokaryotes. *FEMS*  
17 *Microbiol. Rev.* **27**:313-339.
- 18 56. **Pauwels, H., W. Kloppmann, J.-C. Foucher, A. Martelat, and V. Fritsche.** 1998.  
19 Field tracer test for denitrification in a pyrite-bearing schist aquifer. *Appl. Geochem.* **13**:  
20 767-778.
- 21 57. **Peeters, T. L., M. S. Liu, and M. I. H. Aleem.** 1970. The tricarboxylic acid cycle in  
22 *Thiobacillus denitrificans* and *Thiobacillus-A2*. *J. Gen. Microbiol.* **64**: 29-35.



- 1 58. **Pereira, M. M., and M. Teixeira.** 2004. Proton pathways, ligand binding and dynamics  
2 of the catalytic site in haem-copper oxygen reductases: a comparison between the three  
3 families. *Biochim. Biophys. Acta* **1655**:340–346.
- 4 59. **Petri, R., L. Podgorsek, and J. F. Imhoff.** 2001. Phylogeny and distribution of the *soxB*  
5 gene among thiosulfate-oxidizing bacteria. *FEMS Microbiol. Lett.* **197**:171-178.
- 6 60. **Philippot, L.** 2002. Denitrifying genes in bacterial and Archaeal genomes. *Biochim.*  
7 *Biophys. Acta* **1577**:355-376.
- 8 61. **Pitcher, R. S., and N. J. Watmough.** 2004. The bacterial cytochrome *cbb<sub>3</sub>* oxidases.  
9 *Biochim. Biophys. Acta* **1655**:388-399.
- 10 62. **Pott, A. S., and C. Dahl.** 1998. Sirohaem sulfite reductase and other proteins encoded by  
11 genes at the *dsr* locus of *Chromatium vinosum* are involved in the oxidation of  
12 intracellular sulfur. *Microbiology, UK* **144**:1881-1894.
- 13 63. **Quentmeier, A., P. Hellwig, F. Bardichewsky, G. Grolle, R. Kraft, and C. G.**  
14 **Friedrich.** 2003. Sulfur oxidation in *Paracoccus pantotrophus*: interaction of the sulfur-  
15 binding protein SoxYZ with the dimanganese SoxB protein. *Biochem. Biophys. Res.*  
16 *Commun.* **312**:1011-1018.
- 17 64. **Quentmeier, A., P. Hellwig, F. Bardichewsky, R. Wichmann, C. G. Friedrich.** 2004.  
18 Sulfide dehydrogenase activity of the monomeric flavoprotein SoxF of *Paracoccus*  
19 *pantotrophus*. *Biochemistry* **43**:14696-14703.
- 20 65. **Quentmeier, A., R. Kraft, S. Kostka, R. Klockenkaper, and C. G. Friedrich.** 2000.  
21 Characterization of a new type of sulfite dehydrogenase from *Paracoccus pantotrophus*  
22 GB17. *Arch. Microbiol.* **173**:117-125.

- 1 66. **Rakhely, G., A. Colbeau, J. Garin, P. M. Vignais, and K. L. Kovacs.** 1998. Unusual  
2 organization of the genes coding for HydSL, the stable [NiFe]hydrogenase in the  
3 photosynthetic bacterium *Thiocapsa roseopersicina*. J. Bacteriol. **180**:1460-1465.
- 4 67. **Rother, D., and C. G. Friedrich.** 2002. The cytochrome complex SoxXA of *Paracoccus*  
5 *pantotrophus* is produced in *Escherichia coli* and functional in the reconstituted sulfur-  
6 oxidizing enzyme system. Biochim. Biophys. Acta **1598**:65-73.
- 7 68. **Rother, D., H. J. Henrich, A. Quentmeier, F. Bardichewsky, and C. G. Friedrich.**  
8 **2001.** Novel genes of the *sox* gene cluster, mutagenesis of the flavoprotein SoxF, and  
9 evidence for a general sulfur-oxidizing system in *Paracoccus pantotrophus* GB17. J.  
10 Bacteriol. **183**:4499-4508.
- 11 69. **Sawnhey, V., and D. J. D. Nicholas.** 1977. Sulphite-dependent nitrate reductase and  
12 NADH-dependent nitrate reductase from *Thiobacillus denitrificans*. J. Gen. Microbiol.  
13 **100**:49-58.
- 14 70. **Sawnhey, V., and D. J. D. Nicholas.** 1978. Sulphide-linked nitrite reductase from  
15 *Thiobacillus denitrificans* with cytochrome oxidase activity: purification and properties.  
16 J. Gen. Microbiol. **106**:119-128.
- 17 71. **Schedel, M. and H. G. Trüper.** 1979. Purification of *Thiobacillus-denitrificans*  
18 siroheme sulfite reductase and investigation of some molecular and catalytic properties.  
19 Biochim. Biophys. Acta **568**:454-466.
- 20 72. **Shively, J. M., G. L. Decker, and J. W. Greenwalt.** 1970. Comparative ultrastructure of  
21 the thiobacilli. J. Bacteriol. **96**:2138–2143.
- 22 73. **Smith, A. J., and D. S. Hoare.** 1977. Specialist phototrophs, lithotrophs, and  
23 methylotrophs: a unity among a diversity of prokaryotes? Bacteriol. Rev. **41**: 419-448.

- 1 74. **So, A. K.-C., G. S. Espie, E. B. Williams, J. M. Shively, S. Heinhorst, and G. C.**  
2 **Cannon.** 2004. A novel evolutionary lineage of carbonic anhydrase ( $\epsilon$  class) is a  
3 component of the carboxysome shell. *J. Bacteriol.* **186**: 623-630.
- 4 75. **Sonnhammer, E. L. L., G. von Heijne, and A. Krogh.** 1998. A hidden Markov model  
5 for predicting transmembrane helices in protein sequences, p. 175-182. *In* J. Glasgow, T.  
6 Littlejohn, F. Major, R. Lathrop, D. Sankoff, and C. Sensen (ed.), Proceedings of the  
7 Sixth International Conference on Intelligent Systems for Molecular Biology, AAAI  
8 Press, Menlo Park, CA.
- 9 76. **Sorokin, D. Y., T. P. Tourova, A. N. Antipov, G. Muyzer, and J. G. Kuenen.** 2004.  
10 Anaerobic growth of the haloalkaliphilic denitrifying sulfur-oxidizing bacterium  
11 *Thi alkalivibrio thiocyanodenitrificans* sp. nov. with thiocyanate. *Microbiology*, UK  
12 **150**:2435-2442.
- 13 77. **Stover, C. K., X. Q. Pham, A. L. Erwin, S. D. Mizoguchi, P. Warrenner, M. J. Hickey,**  
14 **F. S. L. Brinkman, W. O. Hufnagle, D. J. Kowalik, M. Lagrou, R. L. Garber, L.**  
15 **Goltry, E. Tolentino, S. Westbrook-Wadman, Y. Yuan, L. L. Brody, S. N. Coulter,**  
16 **K. R. Folger, A. Kas, K. Larbig, R. Lim, K. Smith, D. Spencer, G. K.-S. Wong, Z.**  
17 **Wu, I. T. Paulsen, J. Reizer, M. H. Saier, R. E. W. Hancock, S. Lory, and M. V.**  
18 **Olson.** 2000. Complete genome sequence of *Pseudomonas aeruginosa* PA01, an  
19 opportunistic pathogen. *Nature* **406**:959-964.
- 20 78. **Straub, K. L., M. Benz, B. Schink, and F. Widdel.** 1996. Anaerobic, nitrate-dependent  
21 microbial oxidation of ferrous iron. *Appl. Environ. Microbiol.* **62**:1458-1460.

- 1 79. **Stubner S., T. Wind, and R. Conrad.** 1998. Sulfur oxidation in rice field soil: activity,  
2 enumeration, isolation and characterization of thiosulfate-oxidizing bacteria. *Syst. Appl.*  
3 *Microbiol.* **21**: 569-578.
- 4 80. **Suzuki, Y., S. D. Kelly, K. M. Kemner, and J. F. Banfield.** 2003. Microbial  
5 populations stimulated for hexavalent uranium reduction in uranium mine sediment.  
6 *Appl. Environ. Microbiol.* **69**:1337-1346.
- 7 81. **Tanabe, T., T. Funahashi, H. Nakao, S. Miyoshi, S. Shinoda, and S. Yamamoto.**  
8 2003. Identification and characterization of genes required for biosynthesis and transport  
9 of the siderophore vibrioferrin in *Vibrio parahaemolyticus*. *J. Bacteriol.* **185**: 6938-6949.
- 10 82. **Tatusov, R. L., M. Y. Galperin, D. A. Natale, and E. V. Koonin.** 2000. The COG  
11 database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic*  
12 *Acids Res.* **28**:33-36.
- 13 83. **Taylor, B. F., and D. S. Hoare.** 1971. *Thiobacillus denitrificans* as an obligate  
14 autotroph. Cell suspension and enzymic studies. *Arch. Mikrobiol.* **80**: 262-276.
- 15 84. **Taylor, B. F., D. S. Hoare, and S. L. Hoare.** 1971. *Thiobacillus denitrificans* as an  
16 obligate autotroph. Isolation and growth studies. *Arch. Mikrobiol.* **78**: 193-204.
- 17 85. **Timmer-Ten Hoor, A.** 1975. A new type of thiosulfate oxidizing, nitrate reducing  
18 microorganism: *Thiomicrospira denitrificans* sp. nov. *Netherlands J. of Sea Research*  
19 **9**:344-351.
- 20 86. **Trüper, H. G.** 1994. Reverse siroheme sulfite reductase from *Thiobacillus denitrificans*.  
21 *Meth. Enzymol.* **243**: 422-426.
- 22 87. **Vignais, P. M., B. Billoud, and J. Meyer.** 2001. Classification and phylogeny of  
23 hydrogenases. *FEMS Microbiol. Rev.* **25**: 455-501.

- 1 88. **Vlasceanu, L., R. Popa, and B. K. Kinkle.** 1999. Characterization of *Thiobacillus*  
2 *thioparus* LV43 and its distribution in a chemoautotrophically based groundwater  
3 ecosystem. *Appl. Environ. Microbiol.* **63**: 3123-3127.
- 4 89. **Ward N., O. Larsen, J. Sakwa, L. Bruseth, H. Khouri, A. S. Durkin, G. Dimitrov, L.**  
5 **X. Jiang, D. Scanlan, K. H. Kang, M. Lewis, K. E. Nelson, B. Methé, M. Wu, J. F.**  
6 **Heidelberg, I. T. Paulsen, D. Fouts, J. Ravel, H. Tettelin, Q. Ren, T. Read, R. T.**  
7 **DeBoy, R. Seshadri, S. L. Salzberg, H. B. Jensen, N. K. Birkeland, W. C. Nelson, R.**  
8 **J. Dodson, S. H. Grindhaug, I. Holt, I. Eidhammer, I. Jonasen, S. Vanaken, T.**  
9 **Utterback, T. V. Feldblyum, C. M. Fraser, J. R. Lillehaug, and J. A. Eisen.** 2004.  
10 Genomic insights into methanotrophy: the complete genome sequence of *Methylococcus*  
11 *capsulatus* (Bath). *PLoS Biology* **2**: e303 (<http://www.plosbiology.com>).
- 12 90. **Whittenbury R., and D. P. Kelly.** 1977. Autotrophy: a conceptual phoenix. *Symp. Soc.*  
13 *Gen. Microbiol.* **27**: 121-149.
- 14 91. **Wodara, C., S. Kostka, M. Egert, D. P. Kelly, and C. G. Friedrich.** 1994.  
15 Identification and sequence analysis of the *soxB* gene essential for sulfur oxidation of  
16 *Paracoccus denitrificans* GB17. *J. Bacteriol.* **176**:6188-6191.
- 17 92. **Wood, A. P., J. P. Aurikko, and D. P. Kelly.** 2004. A challenge for 21<sup>st</sup> century  
18 molecular biology and biochemistry: what are the causes of obligate autotrophy and  
19 methanotrophy? *FEMS Microbiol. Rev.* **28**: 335-352.
- 20 93. **Zaman, Z., S. B. Bowman, G. D. Kornfeld, A. J. Brown, and I. W. Dawes.** 1999.  
21 Transcription factor GCN4 for control of amino acid biosynthesis also regulates the  
22 expression of the gene for lipoamide dehydrogenase. *Biochem. J.* **340**: 855-862.

- 1 94. **Zumft, W. G.** 1997. Cell biology and molecular basis of denitrification. *Microbiol. Mol.*
- 2 *Biol. Rev.* **61**:533-616.

TABLE 1. General features of the *T. denitrificans* strain ATCC 25259 genome.

Chromosome size, base pairs	2,909,809
G+C ratio	66.07 %
Coding density	92.5 %
Number of predicted protein coding genes	2827
Average CDS length, bp	952
Number of predicted proteins unique to <i>T. denitrificans</i> (%)	89 (3.1%)
rRNA operons	2
tRNA genes	43
Other small RNAs	3
Number of predicted proteins with putative function (%)	2183 (77.2%)
Number of predicted proteins with unknown function (%)	644 (22.8%)
Protein categories (%)	
Energy production	6.6 %
Inorganic ion transport	5.7 %
Cell envelope biogenesis	6.0 %
BLASTP comparison against the KEGG completed microbial genomes database (number of top KEGG hits)	
β-proteobacteria	1624
γ-proteobacteria	455
α-proteobacteria	119
δ-proteobacteria	92
Archaea	22

TABLE 2. Regions with uncharacteristic G+C content and Karlin signatures<sup>a</sup>.

Location in genome	% G+C	CDS	BLAST hits
394058..402433	58.87	Tbd0363-0373	Phage integrase, regulatory protein, hypothetical proteins
979890..1004773	57.39	Tbd0925-0943	Methylase, transposase, regulatory protein, helicase (Snf2/Rad54 family), type III restriction/modification system (methylase, restriction enzyme), phage integrases and several hypothetical proteins
1264595..1271704	56.99	Tbd1207-1215	Phage integrase, conserved imported protein, and hypothetical proteins
1395068..1400129	52.43	Tbd1314-1318	Phage integrase, and hypothetical and conserved hypothetical proteins
1424272..1431083	60.88	Tbd1342-1354	Mostly hypothetical proteins, and a conserved imported protein
1575347..1583813	57.71	1487-1492	Type II/III restriction/modification system (helicase, methylase), conserved hypothetical and hypothetical proteins



1767440..1787406 <sup>b</sup>	60.46	Tbd1679- 1695	Phage integrase, conjugal transfer region (TraWBAY), a number of hypothetical proteins, and imported AAA superfamily ATPase and conserved hypothetical protein
1955614..1961331	61.96	Tbd1861- 1865	Type IV pilus proteins (PileWXV and FimT)
2052544..2061073	59.64	Tbd1958- 1969	Transposases, conserved imported hypothetical proteins and hypothetical proteins
2089032..2093604 <sup>b</sup>	57.82	Tbd2000- 2005	Phage integrase, plasmid recombination protein, and several hypothetical proteins
2160137..2166983 <sup>b</sup>	62.36	Tbd2066- 2073	Phage integrase, phage replication protein, DNA helicase and several hypothetical proteins
2750091..2764591	57.46	Tbd2675- 2685	Phage integrase, prophage regulatory protein, Type I restriction/modification system (HsdMS, R)
2900270..2905890	61.02	Tbd2817- 2823	Phage integrase, phage primase and phage regulatory protein

<sup>a</sup> All regions are also supported by Karlin signature difference.

<sup>b</sup> Flanked by a tRNA.

TABLE 3. Differential transcription of selected genes in *T. denitrificans* that occur in multiple copies

<b>Genes compared</b>	<b>Thiosulfate-induced<sup>a</sup> (fold difference)</b>	<b>FeCO<sub>3</sub>-induced<sup>b</sup> (fold difference)</b>
<i>dsrC (putative)</i>		
Tbd2480/Tbd2480 <sup>c</sup>	1 <sup>d</sup>	0.38
Tbd2488/Tbd2480 <sup>c</sup>	1.4	0.025
Tbd2658/Tbd2480 <sup>c</sup>	0.069	0.006
Tbd2326/Tbd2480 <sup>c</sup>	0.015	0.002
Tbd2327/Tbd2480 <sup>c</sup>	0.009	0.002
Tbd1365/Tbd2480 <sup>c</sup>	0.039	0.009
Tbd1408/Tbd2480 <sup>c</sup>	0.82	0.031
Tbd1926/Tbd2480 <sup>c</sup>	0.001	0.002
<i>Adenylylsulfate reductase</i>		
$\alpha$ subunit		
Tbd0872/Tbd2282	80	NA <sup>e</sup>
$\beta$ subunit		
Tbd0873/Tbd2283	740	NA
<i>ATP sulfurylase</i>		
Tbd0874/Tbd0210	5	NA
<i>Nitric oxide reductase</i>		
<i>norC</i>		
Tbd0562/Tbd0822	63	NA
<i>norB</i>		

<sup>a</sup> Cells harvested for RNA while carrying out thiosulfate oxidation and denitrification (see Materials and Methods).

<sup>b</sup> Cells harvested for RNA while carrying out oxidation of Fe(II) (in FeCO<sub>3</sub>) and denitrification (see Materials and Methods).

<sup>c</sup> Comparison made to the number of transcripts for Tbd2480 under thiosulfate-induced conditions. The transcript copy number for Tbd2480 under these conditions was relatively high (in the copy number range of Tbd0562 and 0561, which encode subunits of nitric oxide reductase, a key enzyme involved in denitrification).

<sup>d</sup> By definition.

<sup>e</sup> Not analyzed.

TABLE 4. CDS potentially encoding *c*-type cytochromes in the *T. denitrificans* genome<sup>a</sup>

<b>CDS</b>	<b>Annotation<sup>b</sup></b>	<b>Molecular mass (Da)<sup>c</sup></b>	<b>No. of hemes</b>
Tbd0055	Cytochrome <i>c</i> family protein	20116	1
Tbd0064	Cytochrome <i>c</i> -553	23727	2
Tbd0070	Probable <i>nirN</i>	63589	1
Tbd0076	Probable <i>nirC</i>	10547	1
Tbd0077	<i>nirS</i> (cytochrome <i>cd</i> <sub>1</sub> )	62992	1
Tbd0094	Hypothetical protein	17109	1
Tbd0128	Cytochrome <i>c</i>	38000	2
Tbd0129	Cytochrome <i>c</i>	21784	2
Tbd0137	Diheme cytochrome <i>c</i>	19427	2
Tbd0138	Cytochrome <i>c</i> -type protein	14395	1
Tbd0146	Probable cytochrome <i>c</i> 5	26734	2
Tbd0187	Cytochrome <i>c</i>	21387	2
Tbd0219	FAD/FMN-containing dehydrogenase	138251	1
Tbd0325	<i>aa</i> <sub>3</sub> -type cytochrome <i>c</i> oxidase, subunit II	41279	1
Tbd0339	<i>cb</i> <sub>3</sub> -type cytochrome <i>c</i> oxidase, subunit II	28037	1
Tbd0341	<i>cb</i> <sub>3</sub> -type cytochrome <i>c</i> oxidase, subunit III	33124	2
Tbd0436	Excinuclease ATPase subunit	103145	1
Tbd0562	<i>norC</i>	15845	1
Tbd0564	<i>soxA</i>	30960	1
Tbd0567	<i>soxX</i>	12712	1
Tbd0571	DMSO reductase chain B	25909	1

Tbd0640	<i>cbb</i> <sub>3</sub> -type cytochrome <i>c</i> oxidase, subunit III	33873	2
Tbd0642	<i>cbb</i> <sub>3</sub> -type cytochrome <i>c</i> oxidase, subunit II	22314	1
Tbd0723	Possible high-affinity Fe <sup>2+</sup> /Pb <sup>2+</sup> permease	69511	1
Tbd0752	MSHA pilin biogenesis ATPase protein MshE	62481	1
Tbd0820	Cytochrome <i>c</i> (in/near nonfunctional <i>nor</i> cluster)	57261	2
Tbd0822	<i>norC</i> -related (potentially not functional)	30200	2
Tbd0840	Probable cytochrome <i>c</i> <sub>5</sub>	16718	1
Tbd0917	<i>soxX</i>	13246	1
Tbd0918	<i>soxA</i>	30875	1
Tbd1169	Ferredoxin, 2Fe-2S	12259	1
Tbd1357	Unknown	16635	1
Tbd1398	Putative cytochrome <i>c</i> -type protein	15842	1
Tbd1404	<i>narH</i>	59191	1
Tbd1484	Cytochrome <i>c</i>	9544	1
Tbd1520	Putative Fe-S protein	48430	1
Tbd1542	ATPase involved in DNA replication	61297	1
Tbd1564	Probable ribonuclease E	95354	1
Tbd1585	Putative pyruvate formate-lyase-activating enzyme	40800	1
Tbd1831	Putative cytochrome <i>c</i> <sub>1</sub>	27211	1
Tbd1840	Unknown	11061	1
Tbd2026	Possible cytochrome <i>c</i> <sub>4</sub> or <i>c</i> -553	11253	1
Tbd2027	Cytochrome <i>c</i> , class IC	11521	1
Tbd2034	Possible cytochrome subunit of sulfide dehydrogenase	10395	1

Tbd2060	Possible alpha-mannosidase	64835	1
Tbd2157	Cytochrome <i>c</i>	18742	1
Tbd2170	Activase of anaerobic class III ribonucleotide reductase	24585	1
Tbd2181	Unknown	20904	1
Tbd2476	<i>dsrJ</i>	17948	3
Tbd2477	<i>dsrL</i>	71328	1
Tbd2545	Diheme cytochrome <i>c</i>	37973	2
Tbd2726	Cytochrome <i>c</i>	11096	1
Tbd2727	Conserved protein of unknown function	74616	1
Tbd2738	Zinc-dependent hydrolase	26287	1

<sup>a</sup> As defined by the presence of at least one CXXCH heme-binding motif. Tbd0039 (which encodes ribosomal protein L31) and Tbd1539 (which encodes DnaJ) include the CXXCH motif but were excluded from this table.

<sup>b</sup> Best attempt at annotation based on examination of best BLASTP hits, top ten PSI-BLAST hits, and genomic context.

<sup>c</sup> Molecular mass predicted for the unprocessed gene product without cofactors.

TABLE 5. Comparison of encoded efflux-mediated heavy metal resistance systems among the genomes of *T. denitrificans* and other selected bacteria.

Bacterial species	Genome size (Mbp)	HME RND	CDF	CPx-type ATPases
<i>Thiobacillus denitrificans</i> <sup>a</sup>	2.9	5	3	3
<i>Geobacter sulfurreducens</i> PCA <sup>a</sup>	3.8	3	2	2
<i>Ralstonia metallidurans</i> <sup>b</sup>	6.9	12	3	5
<i>Pseudomonas aeruginosa</i> <sup>b</sup>	6.3	1	3	4
<i>Escherichia coli</i> <sup>b</sup>	4.6	1	2	2

<sup>a</sup>Putative heavy metal-exporting protein families for *T. denitrificans* and *G. sulfurreducens* PCA (GenBank AE01780) were identified as described by Nies (55). Confirmatory data on metal transport in *G. sulfurreducens* was obtained from an unpublished source (H. A. Vrionis and D. R. Lovley, 2005, Poster I-054 at the ASM 105<sup>th</sup> General Meeting).

<sup>b</sup>From Nies (55).



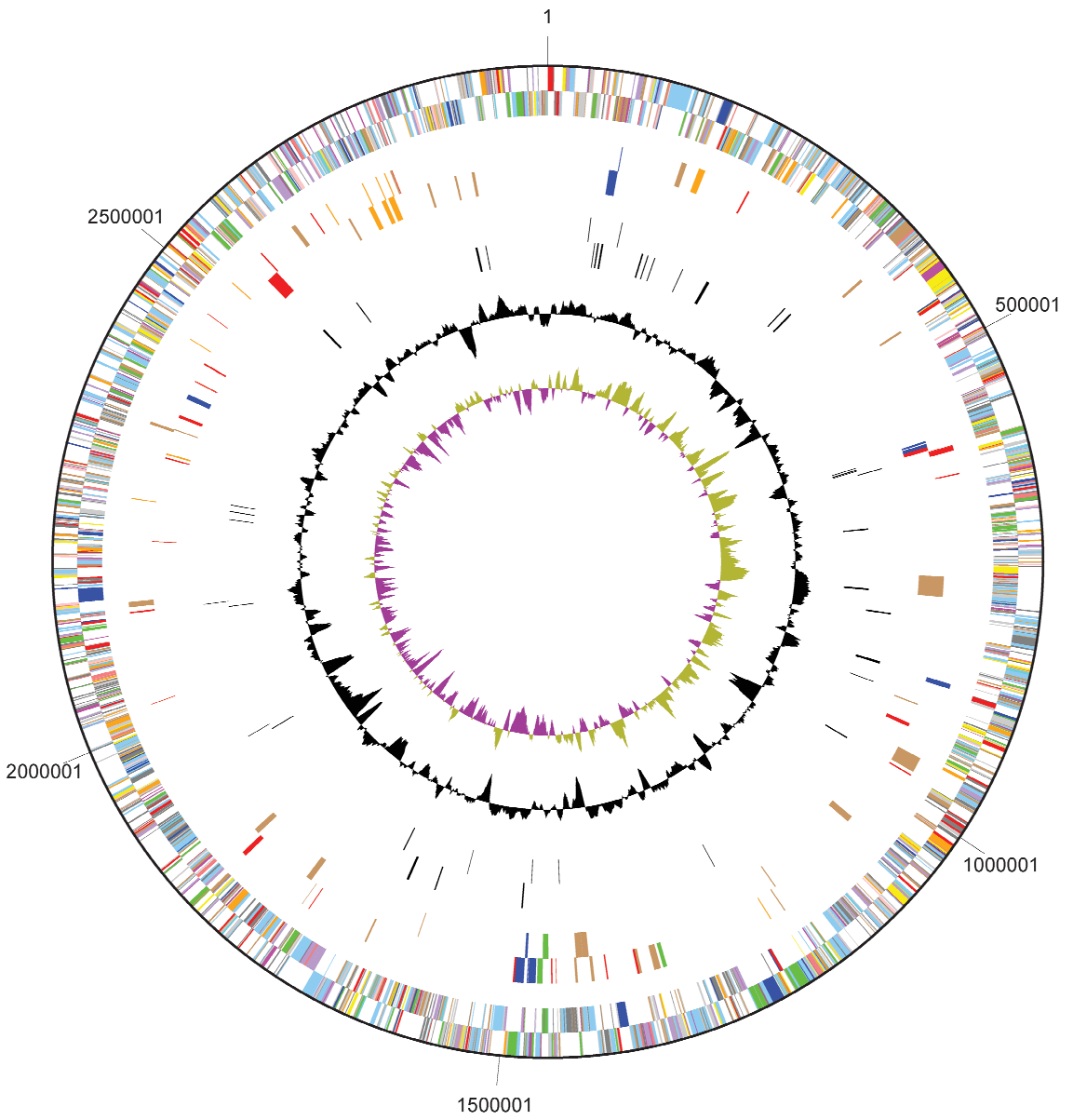
## FIGURE LEGENDS

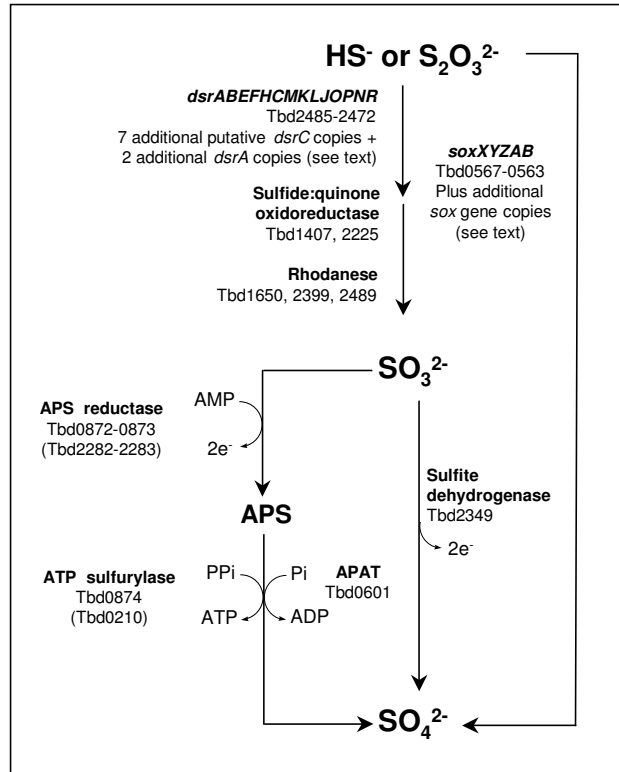
FIG. 1. Schematic circular diagram of the *T. denitrificans* ATCC 25259 genome. Outer circle, predicted coding regions on the forward strand, color-coded by role categories (dark grey - hypothetical proteins, light grey- conserved hypothetical and unknown function, brown - general function, red – DNA replication and repair, green - energy metabolism, blue - carbon and carbohydrate metabolism, cyan - lipid metabolism, magenta - transcription, yellow - translation, orange - amino acid metabolism, pink - metabolism of cofactors and vitamins, light red - purine and pyrimidine metabolism, lavender - signal transduction, sky blue - cellular processes, pale green - structural RNAs). Second circle, predicted coding regions on the reverse strand, color-coded as for outer circle. Third and fourth circles, coding regions (on forward and reverse strands) predicted to be involved in denitrification (blue), sulfur-compound oxidation (red), hydrogen oxidation (green), autotrophic carbon assimilation (orange), and metal ion transport/resistance (brown). Fifth and sixth circles, coding regions found to have a CXXCH heme-binding motif and therefore potentially encoding *c*-type cytochromes. Seventh circle, deviation from the average G+C. Eighth circle, GC skew (positive – olive; negative – purple).

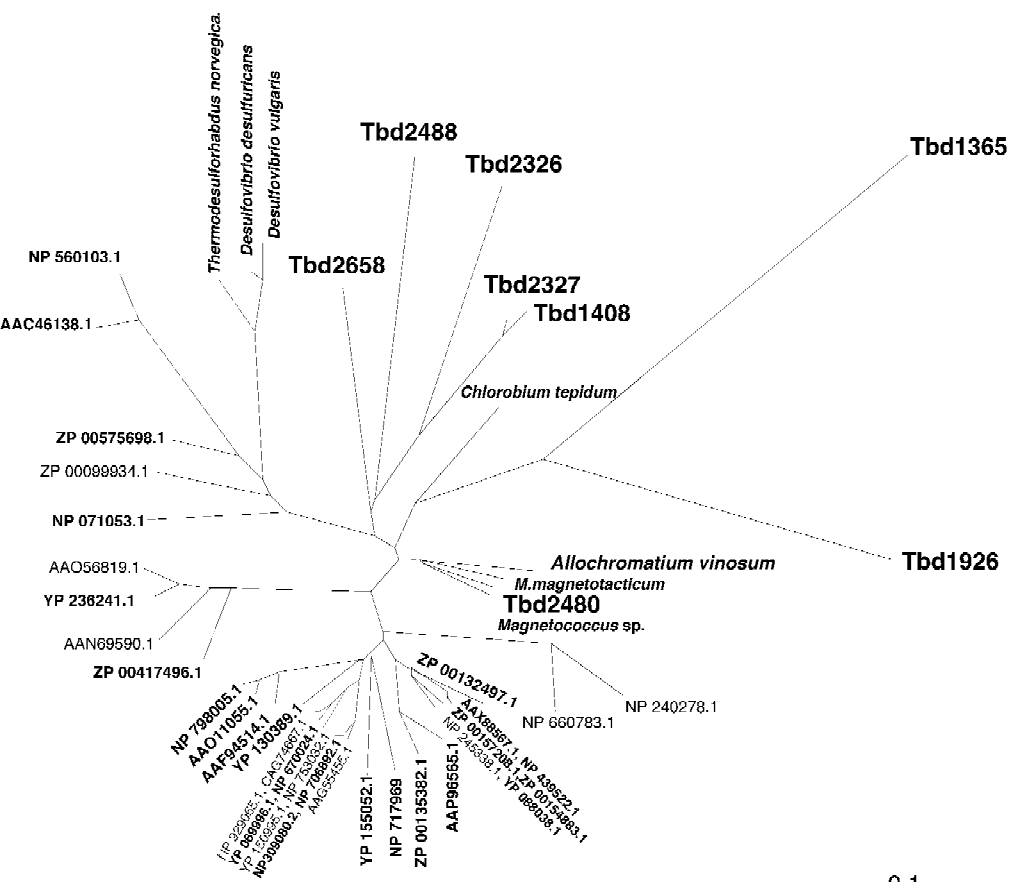
FIG. 2. Schematic overview of key genes/enzymes putatively associated with sulfur-compound oxidation in *T. denitrificans*. Genes in parenthesis have been shown to be lesser expressed paralogs (this study). The biochemical roles of a number of gene products represented in this figure have not been experimentally demonstrated in *T. denitrificans* and are uncertain. Sulfide:quinone oxidoreductase is not proposed to catalyze the direct oxidation of sulfide to sulfite, but rather may participate in an indirect pathway (20). The arrow between thiosulfate and sulfate (right side) represents the possibility that SoxB catalyzes a sulfate thiohydrolase reaction (28) in *T. denitrificans*.

FIG. 3. Phylogenetic relationships among the eight putative DsrC proteins encoded in *T. denitrificans* ATCC 25259 and the top BLASTP matches from the GenBank nr database for Tbd2480. Of the proteins represented in this figure that are not from *T. denitrificans*, more than 70 percent are known or predicted to be DsrC or more broadly related to sulfite reductases (indicated in bold faced type). For limbs that show species names rather than GenBank accession numbers, the corresponding accession numbers are as follows: *A. vinosum* (AAC35399.1), *M. magnetotacticum* (ZP 00052645.1), *Magnetococcus* sp. (ZP 00287929.1), *C. tepidum* (NP 663123.1), *T. norvegica* (CAC36215.1), *D. desulfuricans* (ZP 00130056.2), and *D. vulgaris* (YP 011988).

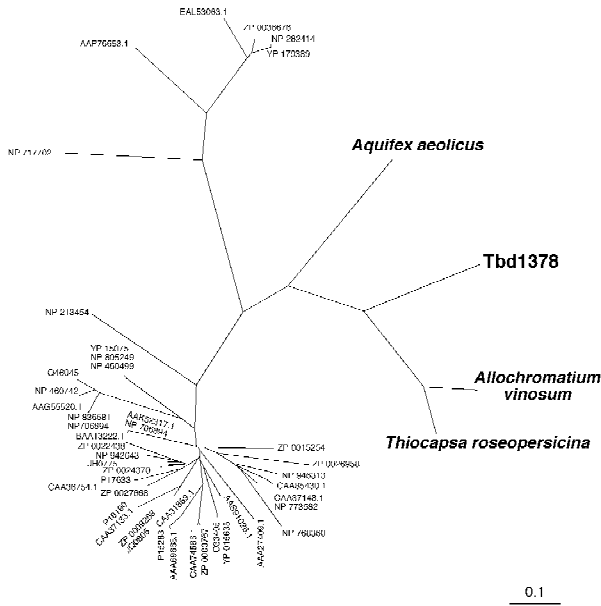
FIG. 4. Phylogenetic relationships among predicted amino acid sequences for HynS (A), Isp1 (B), Isp2 (C), and HynL (D) in *T. denitrificans* and the best BLASTP matches from the GenBank nr database. For limbs that show species names rather than GenBank accession numbers, the corresponding accession numbers are as follows: (HynS) *A. vinosum* (AAU93828.1), *T. roseopersicina* (AAC38281.1), *A. aeolicus* (NP 213658.1); (Isp1) *A. vinosum* (AAU93829.2), *T. roseopersicina* (AAC38283.1), *A. aeolicus* (NP 213657.1), *D. vulgaris* Hmc5 (YP 009755), *A. ambivalens* (CAC86885.1), *D. desulfuricans* NarI (ZP 00128546.1); (Isp2) *A. vinosum* (AAY89333.1), *T. roseopersicina* (AAC38284.1), *A. aeolicus* (NP 213656.1), *A. ambivalens* (CAC86886.1), *Polaromonas* sp. (ZP 00503323.1), *C. aurantiacus* (ZP 00356812), *A. vinosum* DsrK (AAC35401.2), *C. tepidum* DsrK (NP 663117.1), *D. vulgaris* Hmc6 (YP 009754.1); (HynL) *A. vinosum* (AAY89334.1), *T. roseopersicina* (AAC38282.1), *A. aeolicus* (NP 213655.1).



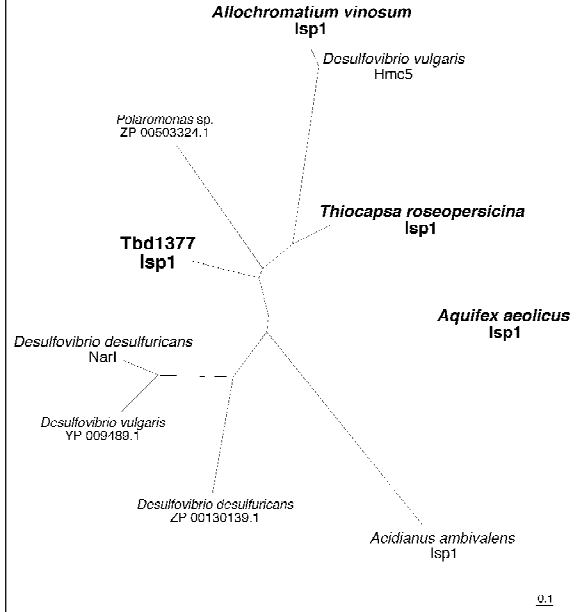




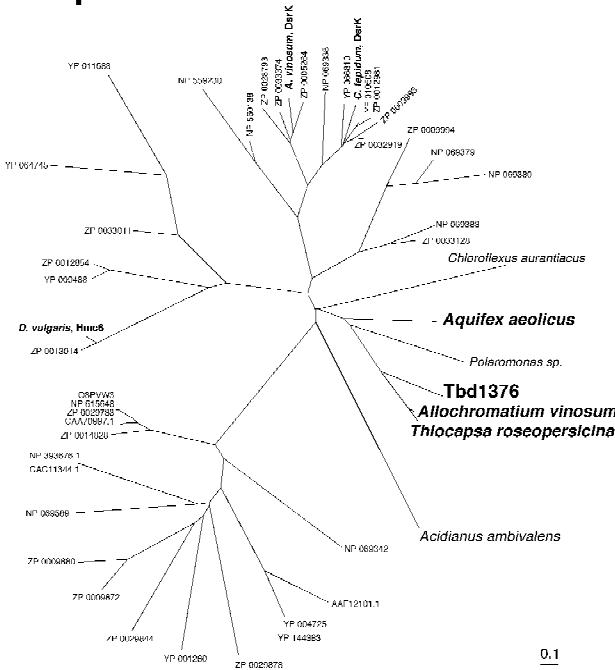
# A. HynS



# B. Isp1



# C. Isp2



# D. HynL

