

Complete Genome Sequence of *Nitrosospira multiformis*, an Ammonia-Oxidizing Bacterium from the Soil Environment^{∇†}

Jeanette M. Norton,^{1*} Martin G. Klotz,² Lisa Y. Stein,³ Daniel J. Arp,⁴ Peter J. Bottomley,⁴ Patrick S. G. Chain,^{5,6,7} Loren J. Hauser,⁸ Miriam L. Land,⁸ Frank W. Larimer,⁸ Maria W. Shin,^{5,6} and Shawn R. Starkenburg⁴

*Department of Plants, Soils, and Climate, Utah State University, Logan, Utah 84322*¹; *Departments of Biology and Microbiology and Immunology, University of Louisville, Louisville, Kentucky 40292*²; *Department of Environmental Sciences, University of California, Riverside, California 92521*³; *Department of Botany and Plant Pathology, Oregon State University, Corvallis, Oregon 97331*⁴; *Chemistry, Materials, and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, California 94550*⁵; *Microbial Program, Joint Genome Institute, Walnut Creek, California 94598*⁶; *Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, Michigan 48824*⁷; and *Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831*⁸

Received 3 December 2007/Accepted 7 March 2008

The complete genome of the ammonia-oxidizing bacterium *Nitrosospira multiformis* (ATCC 25196^T) consists of a circular chromosome and three small plasmids totaling 3,234,309 bp and encoding 2,827 putative proteins. Of the 2,827 putative proteins, 2,026 proteins have predicted functions and 801 are without conserved functional domains, yet 747 of these have similarity to other predicted proteins in databases. Gene homologs from *Nitrosomonas europaea* and *Nitrosomonas eutropha* were the best match for 42% of the predicted genes in *N. multiformis*. The *N. multiformis* genome contains three nearly identical copies of *amo* and *hao* gene clusters as large repeats. The features of *N. multiformis* that distinguish it from *N. europaea* include the presence of gene clusters encoding urease and hydrogenase, a ribulose-bisphosphate carboxylase/oxygenase-encoding operon of distinctive structure and phylogeny, and a relatively small complement of genes related to Fe acquisition. Systems for synthesis of a pyoverdine-like siderophore and for acyl-homoserine lactone were unique to *N. multiformis* among the sequenced genomes of ammonia-oxidizing bacteria. Gene clusters encoding proteins associated with outer membrane and cell envelope functions, including transporters, porins, exopolysaccharide synthesis, capsule formation, and protein sorting/export, were abundant. Numerous sensory transduction and response regulator gene systems directed toward sensing of the extracellular environment are described. Gene clusters for glycogen, polyphosphate, and cyanophycin storage and utilization were identified, providing mechanisms for meeting energy requirements under substrate-limited conditions. The genome of *N. multiformis* encodes the core pathways for chemolithoautotrophy along with adaptations for surface growth and survival in soil environments.

Nitrification is a key process in the nitrogen cycle of terrestrial, wastewater, and marine systems. The first step in the aerobic process is the oxidation of ammonia, mediated by ammonia-oxidizing bacteria (AOB) or ammonia-oxidizing archaea. Because we are particularly interested in the genetic complement adaptive for ammonia-based chemolithotrophy in the soil environment, we completed the genome sequence of the soil AOB *Nitrosospira multiformis* (ATCC 25196^T). Obtaining the *N. multiformis* genome sequence offers a unique opportunity for comparison to the available genomes of other betaproteobacterial AOB (beta-AOB), *Nitrosomonas europaea*

(16), and *Nitrosomonas eutropha* (63). The AOB isolated or detected by noncultural methods in aerobic surface soils all have been members of the *Betaproteobacteria* (order *Nitrosomonadales*, family *Nitrosomonadaceae*). Recent evidence suggests that *Crenarchaeota* may also contribute to ammonia oxidation in soils (43).

The sequenced AOB, *Nitrosospira multiformis* ATCC 25196^T, was isolated from soil near Paramaribo, Surinam, by enrichment culturing, followed by serial dilution to extinction (71). Originally, this isolate was the type strain for *Nitrosolobus multiformis*, with a genus name indicative of its lobular morphology. *N. multiformis* and closely related cluster 3 *Nitrosospira* (41, 64) are commonly identified as important members of the AOB community in agricultural soils from a range of geographical locations (3, 4, 12, 45, 52).

Soil AOB survive in a discontinuous environment subject to rapid changes in water potential, diffusional limitation of substrate supply (61), and competition from a range of heterotrophic bacteria and plant roots using ammonium as a nitrogen

* Corresponding author. Mailing address: Department of Plants, Soils, and Climate, Utah State University, Logan, UT 84322-4820. Phone: (435) 797-2166. Fax: (435) 797-3376. E-mail: jennyn@cc.usu.edu.

† Supplemental material for this article may be found at <http://aem.asm.org/>.

∇ Published ahead of print on 4 April 2008.

TABLE 1. General characteristics of the *Nitrosospora multiformis* genome (IMG version 2.4, 2007)

General characteristic	Value
No. of chromosomes	1
No. of plasmids	3
No. of base pairs	3,234,309
G+C content (%)	53.88
Coding density (%)	85.6
No. of predicted protein-coding genes	2,827
No. of predicted proteins with putative function (%)	2,026 (70.5)
No. of predicted proteins of unknown function	801
No. of predicted proteins with similarity to other proteins of unknown function	747
No. of predicted RNA coding genes	49
tRNAs	43
16S-Ala tRNA _{TGC} -Ile tRNA _{GAT} -23S-5S	1 (16S), 1 (23S), 1 (5S)
Miscellaneous RNAs	3

source. The soil pH (21) and oxygen availability (10) are known selective factors for AOB in the environment. The AOB grow strongly attached to soil particle surfaces embedded in an exopolysaccharide (EPS) matrix (1) often in association with nitrite-oxidizing bacteria, such as those in the *Nitrobacter* and *Nitrospira* genera (27). Adaptation to and evolution in this environment are reflected in the gene complement found in *N. multiformis* described herein.

MATERIALS AND METHODS

Strains and culture conditions. *Nitrosospora multiformis* ATCC 25196^T was obtained from the American Type Culture Collection and maintained on ATCC medium 929 at 28°C in the dark as described previously (50).

Construction, isolation, and sequencing of small-insert and large-insert libraries. Genomic DNA isolated from *N. multiformis* was sequenced using a conventional whole-genome shotgun strategy. Briefly, three libraries were constructed using randomly sheared genomic DNA fragments of approximately 3 kb, 8 kb, and 40 kb in size that were ligated into pUC18, pMCL200, and pCC1Fos cloning vectors, respectively. Double-ended plasmid sequencing reactions were performed at the Department of Energy Joint Genome Institute (JGI) using ABI 3730xl DNA analyzers and MegaBACE 4500 genetic analyzers as described on the JGI website <http://www.jgi.doe.gov/>. After quality control of the 44,363 total initial reads of draft sequence, 37,538 sequences were assembled, producing an average of 11.1-fold coverage across the entire genome. The reads were assembled into 39 high-quality draft sequence contigs, which were linked into 23 larger scaffolds by paired-end sequence information. Gaps in the sequence were closed by primer walking on gap-spanning library clones or with PCR products from genomic DNA. Remaining physical (uncaptured) gaps, some of which are regions suspected of being lethal in *Escherichia coli*, were closed by combinatorial PCR and sequencing. Sequence finishing and polishing added 868 reads, and final assessment of the completed genome (one chromosome and three plasmids) was performed as previously described (16).

Sequence analysis and annotation. Automated gene modeling and functional assignments were completed using multiple databases and modeling packages as described elsewhere (17, 37). Manual curation of automated assignment was completed on an individual gene-by-gene basis as needed. Comparative analyses of bacterial genomes and gene neighborhoods were completed using the JGI Integrated Microbial Genomes web-based interface (<http://img.jgi.doe.gov/>). The sequence and results of automatic annotations are available at http://genome.ornl.gov/microbial/nmul/29apr05/kegg_summary.html. Putative Rho-independent transcriptional terminators were predicted using TransTermHP (36).

Nucleotide sequence accession numbers. The sequence of the complete *N. multiformis* genome is available under NCBI accession numbers NC_007614 (chromosome 1), NC_007615 (plasmid 1 [18.8 kb]), NC_007616 (plasmid 2 [17.0 kb]), and NC_007617 (plasmid 3 [14.2 kb]).

TABLE 2. COG classifications of the genes in the *Nitrosospora multiformis* genome (IMG version 2.4, 2007)

Protein category based on COGs	No. of genes with COG (% in COG category) ^a
Any COG category	2,102 (73.1)
Cell wall, membrane, and envelope	194 (6.8)
DNA replication and repair	182 (6.3)
Energy production and conversion	160 (5.6)
Amino acid transport and metabolism	154 (5.4)
Translation, ribosome structure, and biogenesis	151 (5.3)
Posttranslational modification and chaperones	129 (4.5)
Signal transduction mechanisms	124 (4.3)
Coenzyme transport and metabolism	108 (3.8)
Inorganic ion transport and metabolism	105 (3.7)
Transcription	99 (3.4)
Intracellular traffic and secretion	96 (3.3)
Carbohydrate transport and metabolism	94 (3.3)
Lipid metabolism	77 (2.7)
Cell motility	64 (2.2)
Secondary metabolites	59 (2.0)
Nucleotide transport and metabolism	54 (1.9)
Defense	31 (1.1)
General function prediction only	227 (7.9)
Unknown function	197 (6.8)

^a Classifications of genes are based on clusters of orthologous genes (COGs). The percentage given is based on a total gene count of 2,876. Some genes may be counted in more than one COG.

RESULTS AND DISCUSSION

Genome organization, general features, and analysis for lateral gene transfers. The genome of *N. multiformis* ATCC 25196 consists of a single circular chromosome of 3,184,243 bp (G+C content of 53.9%) along with three previously unknown plasmids of 18,871, 17,036, and 14,159 bp in size and G+C contents of 49.5%, 50.0%, and 49.6%, respectively. General features of the genome are listed in Table 1, and detailed circular maps are shown in Fig. S1 in the supplemental material. The genes are distributed evenly around the chromosome, with 1,337 transcribed from the forward strand and 1,420 transcribed from the complementary strand. Plasmids 1, 2, and 3 have 17, 16, and 15 genes identified, respectively. There were 49 RNA genes with a single copy of the rRNA operon of the 16S-Ala tRNA_{TGC}-Ile tRNA_{GAT}-23S-5S type and 43 tRNAs (representing all 20 amino acids). A total of 2,827 coding genes averaging 980 bp in length emerged from the modeling effort, 22 coding sequences (CDS) are fragmentary, frameshifted, or interrupted by insertion sequence elements (ISE); these have been designated pseudogenes. Of the 2,827 putative proteins, 2,026 have similarity to a protein with a functional assignment. Other protein searches give similar results: 2,119 proteins match InterPro profiles; 2,014 match a Pfam hmm profile; 2,102 can be assigned to a cluster of orthologous gene (COG) group (Table 2).

The three plasmids have lower G+C content than the chromosome and carry few genes of known function besides those encoding plasmid replication initiation, partitioning, and mobilization functions. Functional plasmid genes did not have significant similarity to each other or to genes on the chromosome. However, all three plasmids do encode putative peptidases that may confer some advantage to the organism in the

TABLE 3. Complex repetitive sequences including insertion sequence elements in *N. multiformis*

Gene cluster or region or ISE and transposase	Size (bp) of repeat	No. of copies	Gene locus no. ^a	Designation
Gene cluster or region				
<i>amoCAB-orf4-orf5 copC copD</i>	8,114	2	A2326—A2320, A2766—A2760	
<i>amoCAB</i> adjacent to <i>hao</i> (supercluster)	3,729	1	A0798—A0800	
<i>amoC</i> single	844	1	A0177	
<i>amoC</i> divergent at 75%	565	1	A2467	
<i>hao-orf2-C₅₅₄-Cyt_{m552}</i>	5,090	3	A0805—A0802, A1082—A1085, A2662—A2659	
Translation elongation factor Tu	1,180	2	A0752, A0765	
Membrane protein, Kef-type K ⁺ transport system NAD-binding component; ion transport protein	1,354	2	A0075 and A0076, A1690 and A1691	
Diguanylate cyclase/phosphodiesterase [GGDEF and EAL domains with PAS/PAC sensor(s)]	1,131	2 ^b	A0630, A1605	
Amino acid adenylation condensation domains	1,727	2 ^c	A1829, A1830, A1832	
ISE and transposases				
Transposase, IS3/IS911; putative transposase; putative transposase	1,003	5		ISNmu4
Transposase, IS3/IS911; integrase, catalytic region	1,274	5		ISNmu5
Transposase, IS3/IS911; integrase, catalytic region	1,214	3		ISNmu6
Transposase, IS4; transposase	848	13 ^d		ISNmu1
IS298 transposase OrfA; transposase, IS4	815	5		ISNmu3
Integrase, catalytic region	1,228	6		ISNmu2
Transposase, IS3/IS911; integrase, catalytic region, one copy on plasmid 2	1,438	3		ISNmu8
Transposase, IS3/IS911; integrase, catalytic region, one copy on plasmid 3	1,244	4		ISNmu7

^a Nmul gene locus numbers are shown abbreviated: A2326—A2320, Nmul_A2326 to Nmul_A2320.

^b Three other copies with domains of 62 to 80% identity and one partial copy with only the GGDEF domain.

^c Ninety-three percent partial hit within the same protein, 70% identical portion in another protein.

^d Two degenerate copies of 75 to 80% identity.

soil environment. Plasmid 1 also carries a postsegregational killing system (Nmul_B2801 and Nmul_B2802) and a phage integrase (Nmul_B2807). Plasmid 3 carries a resolvase/recombinase (Nmul_D2813), as well as an ISE that is repeated three times within the chromosome. Plasmid 2 encodes a restriction modification system (Nmul_C2785 and Nmul_C2786), a site-specific recombinase (Nmul_C2795), and an ISE (ISNmu8) that is also found twice in the chromosome. It is unclear if this ISE was introduced into the genome via the plasmid, as both copies in the chromosome are associated with regions of unknown origin as well: one lies within a 14.4-kb region (2309003 to 2323384) along with four different repeated ISE (ISNmu1, -3, -4, and -5) and a phage integrase, while the other is adjacent to a site-specific recombinase and a phage-related integrase gene and lies within a 16.6-kb aberrantly low (43.2%) G+C region (113521 to 130139). It is possible that both of these entire regions may have been recently acquired via lateral gene transfer. This scenario is substantiated for the 16.6-kb region in that it encompasses a number of hypothetical proteins, a cytidine deaminase, and an endonuclease, and it interrupts a Mg²⁺ chelatase (Nmul_A0124, N-terminal fragment; Nmul_A0106, C-terminal fragment), conserved in many *Betaproteobacteria*.

Additional regions with nucleotide composition anomalies (G+C content and dinucleotide and trinucleotide frequencies) were analyzed for evidence of recent lateral acquisition by *N. multiformis*. One such region (Nmul_A0922 to Nmul_A0934) encodes the largest gene product (Nmul_A0927, 15,651 bp) in *N. multiformis*. This very large hypothetical protein (5,216 amino acids [aa]) shows little similarity to entries in the

GenBank NR database. This region also encodes putative phage products, i.e., an integrase (Nmul_A0922), a reverse transcriptase (Nmul_A0934), and a transcriptional regulator (Nmul_A0932), and lies directly downstream of a tRNA. We interpret the presence and arrangement of this inventory as evidence of a recent lateral gene acquisition(s).

Complex repetitive sequences. The chromosome has eight families of transposase-encoding ISE, repeated from 2 to 13 times spread randomly throughout the genome; two of these ISE are also found on plasmids as mentioned above (Table 3). All 13 copies of the most numerous ISE are 100% identical to one another and were found to occasionally interrupt CDS. Similar to other beta-AOB, *N. multiformis* harbors repeated copies of *amo* and *hao* gene clusters (Fig. 1 and see below). In addition, several other repeated genes or gene fragments were identified (Table 3) including two nearly identical K⁺ transport systems. One copy (Nmul_A1690 to Nmul_A1691) lies between an ISNmu1 and ISNmu4, with a putative cyanophycinase unique to this AOB and a nonduplicated ISE. In accordance with its large genomic inventory of signaling proteins, multiple genes encoding highly similar proteins with distinctive signaling domains were found.

Taxonomic distribution of gene homologs. The majority of the identified protein-encoding genes have best matches (top KEGG hits) to those from *Betaproteobacteria*, with 753 matches to genes in *N. europaea* and 415 matches to genes in *N. eutropha* (together comprising 42% of all matches with cultivated organisms), 139 to *Thiobacillus denitrificans*, 93 to *Methylobacillus flagellatus*, 65 to *Dechloromonas aromatica*, and

CbbS sequences are 92% and 82% identical to those in *N. multiformis*. In contrast, *N. europaea* and *N. eutropha* have form I A (green-like) *cbb* operons (16, 63) as reported for *Acidithiobacillus ferrooxidans* (59). It is presently not clear whether the *cbb* operons in *Nitrosomonas europaea* as proposed by Wei et al. (73) or in *Nitrospira* are the result of lateral gene transfer. Multiple events of lateral gene transfer, gene duplication, and loss of paralogs are believed to have occurred during the evolution of the RuBisCO-encoding genes in *Proteobacteria* (65).

Unequivocal candidates for genes encoding a Calvin cycle-specific regeneration of pentose phosphates by conversion of triosephosphates via the F1P6Pase/sedoheptulose-1,7-bisphosphatase reactions were not identified in the genome. *N. multiformis* instead may use one of the enzymes described above in combination with a fructose-1,6-biphosphate aldolase (encoded in Nmul_A384) for these functions. A similar class II aldolase was found to be induced during autotrophic growth of *Xanthobacter flavus* (59) and is implicated in *M. capsulatus* (MCA3045) (35). The further pathway to ribulose-5-phosphate mediated by transketolase and pentose-5-phosphate-3-epimerase is likely encoded in Nmul_A0388 and Nmul_A2371, respectively. Phosphoribulokinase (encoded in Nmul_A0562) generates the substrate for RuBisCO irreversibly. Overall, it appears that the central pathway inventories that often contain paralogs, functional analogs, and irreversible enzymes in chemorganoheterotrophs and facultative chemolithoautotrophs have been streamlined in AOB by reductive genome evolution. Future experiments may characterize individual enzymes for their reversibility, substrate specificity, and turnover capacity to more fully resolve the carbon fixation pathway in *N. multiformis*.

(iii) Chemolithotrophy. The *N. multiformis* genome encodes a complete TCA cycle, including alpha-ketoglutarate dehydrogenase (alpha-KGDH; Nmul_A855 to Nmul_A857) and succinate dehydrogenase (Nmul_A0862 to Nmul_A0860). Two divergent copies of the genes encoding succinyl-coenzyme A synthetase (Nmul_A1079 to Nmul_A1080 and Nmul_A1995 to Nmul_A1996) were identified. Genes encoding malate synthase and isocitrate lyase, enzymes unique to the glyoxylate shunt, were not found. In the pregenomic era, failure to detect significant activity of alpha-KGDH and succinate dehydrogenase had been associated with obligate lithoautotrophy (71, 74); however, the presence of genes encoding both of these enzymes in all sequenced AOB genomes (16, 37, 63) and recent experiments with *N. europaea* (31) have dispelled this theory with regard to autotrophy. Even though the TCA cycle is complete, energy generation by oxidation of ammonia was still strictly required for the growth of *N. europaea* on organic carbon sources (32), thereby indicating obligate lithotrophy. Perhaps the role of alpha-KGDH in *N. multiformis* is associated with fitness during the stationary phase as found for *N. europaea* (31). We suggest that in *N. multiformis*, the TCA cycle operates to secure stable pools of precursor metabolites and to utilize stored glycogen.

The genes and genomic context for the central enzymes in energy production, ammonia monooxygenase and hydroxylamine oxidoreductase, are shown in Fig. 1. The *N. multiformis* genome contains three copies of the *amoCAB* operon and two singleton copies of *amoC* (Nmul_A0177 and Nmul_A2467) on

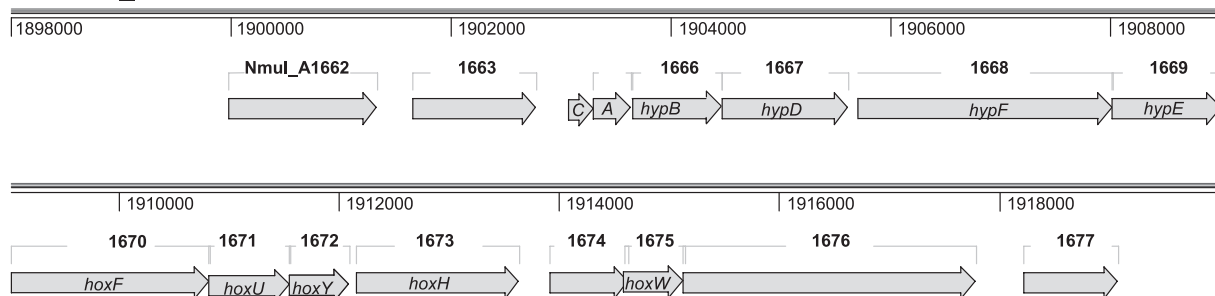
the chromosome as previously noted (50). Two of the *amoCAB* operons reside on 5,440-bp regions that are nearly identical with only 5 nucleotide mismatches in both coding and intergenic regions. Similar to the beta-AOB of the *Nitrosomonas* lineage, these *amoCAB* operons are succeeded by two conserved genes, *orf4* and *orf5*, whose expression products are likely associated with the plasma membrane and involved in ammonia catabolism (2). Furthermore, the *orf4* and *orf5* genes are followed by two genes, *copCD*, that encode copper tolerance or copper sequestration proteins. The *copCD* genes also follow *amoCAB-orf4-orf5* operons of *N. europaea* and *N. eutropha* (16, 63). This arrangement of the *amo* gene cluster seems typical for beta-AOB and was not found in the genome of the gamma-AOB *N. oceani* (2, 37). The third *amoCAB* copy, not found in the nitrosomonads, is nearly 100% identical with the other copies for 3,224 bp but lacks the *orf4-orf5* and *copCD* downstream genes.

In addition to the clustered *amo* and accessory genes, *N. multiformis* encodes singleton copies of *amoC* (Nmul_A0177 and Nmul_A2467), *orf4* (Nmul_A1398), and *orf5* also known as *amoD* (24) (Nmul_A0227, Nmul_A0945, and Nmul_A1933). Whereas Nmul_A0177 is nearly identical to the clustered *amoC* genes, Nmul_A2467 is 25% divergent from the other four *amoC* gene copies. Multiple copies of *amoCAB-orf4-orf5* clusters and singleton *amoC* genes may extend flexibility for expression of ammonia catabolic inventory under fluctuating ammonia concentrations frequently encountered in the soil environment. In *N. europaea*, for example, the three *amoC* gene copies were differentially expressed during recovery from starvation (9), and mutants with inactivated individual *amo* copies exhibited different growth phenotypes (33).

The *hao* gene cluster (*hao-orf2-cycAB*) also exists in three copies (Fig. 1), one of which is contiguous with the third *amoCAB* operon in a single ammonia catabolic supercluster (2). All three copies of the *hao* gene cluster in *N. multiformis* include the *cycB* gene, whereas one of three copies in the *N. europaea* and *N. eutropha* genomes lack this gene (16, 63). Of further interest is that all *hao* gene clusters whose products constitute the hydroxylamine-ubiquinone-reductase module (38), are adjacent to groups of genes that are involved in energy transformation (Fig. 1). This association may have functional implications, as the products of clustered genes conserved in sequence and synteny are very likely to be functionally related (60).

(iv) Electron transport. The gene profile is consistent with electron flow between NADH and the ubiquinone pool via NADH-ubiquinone oxidoreductase (NUO) (complex I), between the ubiquinone pool and cytochrome *c* via the cytochrome *bc*₁ complex III, and from cytochrome *c* to oxygen via cytochrome *c* oxidase (complex IV). The genome of *N. multiformis* contains two complete but distinct sets of genes encoding NUO. Genes Nmul_A1091 to Nmul_A1104 are most similar to the corresponding single-copy genes in *N. europaea* (16), which encodes a NUO that is expected to function in reverse electron flow. The second complex I, encoded in genes Nmul_A1013 to Nmul_A1025, is most similar to NADH dehydrogenase 1 (NDH-1) in *Methylococcus capsulatus* Bath. The role for this second NDH-1 complex is not known, but we speculate that it might be involved with electron flow in the "forward" direction (toward ubiquinone) as it was proposed

A. Nmul_A1662-1677



B. Nmul_A1233-1246

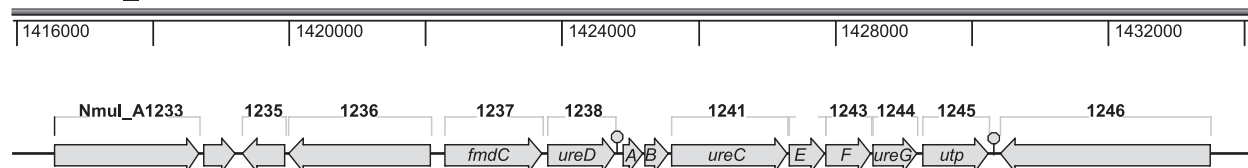


FIG. 2. Unique hydrogenase and urease gene clusters of *N. multiformis*. (A) The hydrogenase-encoding gene cluster from Nmul_A1662 to Nmul_A1677 (Nmul_A1662-1677; genes are abbreviated by their last four numbers) and its putative products and functions are as follows: Nmul_A1662, exonuclease metallo-beta-lactamase RNA processing enzyme; Nmul_A1663, response regulator GAF:metal-dependent phosphohydrolase HD domain (sigma-54 related); Nmul_A1664 (*hypC*), chaperone; Nmul_A1665 (*hypA*), hydrogenase expression and synthesis; Nmul_A1666 (*hypB*), hydrogenase accessory protein; Nmul_A1667 (*hypD*), hydrogenase formation; Nmul_A1668 (*hypF*), [NiFe] hydrogenase maturation protein; Nmul_A1669 (*hypE*), hydrogenase expression/formation protein; Nmul_A1670 to Nmul_A1673, four subunits of hydrogen dehydrogenase (EC 1.1.2.1) (Nmul_A1670 [*hoxF*], NAD-reducing hydrogenase [diaphorase] and 51-kDa alpha subunit; Nmul_A1671 [*hoxU*] ferredoxin gamma subunit Nmul_A1672 [*hoxY*], coenzyme F420-reducing delta 20-kDa small-subunit; Nmul_A1673 [*hoxH*], large-subunit); Nmul_A1674, unknown with N-terminal signal sequence and C-terminal PEP motif; Nmul_A1675 (*hoxW*), hydrogenase maturation peptidase; Nmul_A1676, GCN5-related acetyltransferase coenzyme A-binding transcriptional activation; Nmul_A1677, unknown function with N-terminal signal sequence and C-terminal PEP motif. (B) Urease-encoding cluster and regulatory regions, including the genes from Nmul_A1233 to Nmul_A1246 (Nmul_A1233-1246; genes are abbreviated by their last four numbers), and their products or functions are as follows: Nmul_A1233, TonB-dependent receptor; Nmul_A1234, putative transcriptional regulator and CopG family of nickel-responsive regulators; Nmul_A1235, two-component transcriptional regulator LuxR (regulatory protein, response regulator receiver containing a CheY-like receiver domain and a helix-turn-helix DNA-binding domain); Nmul_A1236, periplasmic sensor signal transduction histidine kinase; Nmul_A1237, FmdC precursor and urea-responsive OM porin; Nmul_A1238, accessory protein UreD; Nmul_A1239 to Nmul_A1241, structural subunits UreA, UreB, and UreC, respectively; Nmul_A1242 to Nmul_A1244, accessory proteins UreE, UreF, and UreG, respectively; Nmul_A1245, urea transporter; Nmul_A1246, transcriptional regulator, putative ATPase, winged helix family.

for the second NDH-1 complex in *N. oceanii* (37). In contrast to the scenario in *N. oceanii*, for which Klotz et al. (37) proposed that the forward NDH-1 function was coupled to NADH generated via a sodium circuit, the forward extension of the electron transport chain in *N. multiformis* might be coupled to H₂ as the electron donor because *N. multiformis* uniquely harbors genes encoding a hydrogenase (see below). The *N. multiformis* genome encodes several heme-copper oxidases (HCOs) that likely function as the terminal oxidase for respiratory electron disposal (complex IV). The genome contains three gene clusters that encode *caa*₃-type HCO: one cluster of genes encoding four subunits (Nmul_A0183 to Nmul_A0185 and Nmul_A0187) and two clusters that lack subunit IV (Nmul_A0458 to Nmul_A0460 and Nmul_A1775 to Nmul_A1777). In addition to gene clusters encoding *caa*₃-type HCOs, the *N. multiformis* genome also contains a gene cluster (Nmul_A2666 to Nmul_A2668) that encodes a novel HCO, termed sNOR, since it is likely reducing NO rather than oxygen (63). This gene cluster *norSY-senC* has been found in all of the AOB genomes and in the genomes of a few other chemolithotrophs including the nitrite oxidizer *Nitrobacter hamburgensis*. The SenC pro-

tein has been implicated in the function of HCO in other organisms (63). Interestingly, one of the gene clusters encoding a *caa*₃-type HCO (Nmul_A0458 to Nmul_A0460) is preceded by genes encoding an additional SenC (Nmul_A0454) and a class I cytochrome *c* (Nmul_A0456), which may constitute yet another functional module. Overall, the genome contains 34 CDS that contain CxxCH heme-coordination motifs indicative of cytochrome *c* (see Fig. S1 in the supplemental material).

(v) **Hydrogenase.** *N. multiformis* is the first AOB genome found to contain a gene cluster encoding a putative [NiFe]-hydrogenase (EC 1.1.2.1). The putative regulon includes 15 genes (Nmul_A1662 to Nmul_A1676), encoding structural and accessory proteins as well as transcriptional regulators (Fig. 2A). The deduced sequence of the large subunit (Nmul_A1673) is most similar to proteins deduced from the genome sequences of *Magnetospirillum magnetotacticum* strains Ms-1 and AMB-1 (Magn03008384 and Amb3396), *Dechloromonas aromatica* RCB (Daro_0982), and *Methylococcus capsulatus* Bath (MCA0114); however, all of these organisms contain the inventory for more than one hydrogenase. In *M. capsulatus* Bath, hydrogenase was involved with membrane-associated H₂ uptake (20, 29). Signature

sequences (L1 and L2) in the large-subunit protein place the *N. multiformis* hydrogenase in class 3d of bidirectional NAD-linked hydrogenases (69). These enzymes are heterotetrameric and are soluble or loosely attached to membranes. In the facultative lithoautotroph *Ralstonia eutrophus* H16, this type of hydrogenase couples hydrogen oxidation to NAD reduction, supplying reducing power (56). As mentioned above, the hydrogenase in *N. multiformis* may function as the donor of reducing power that drives a complete electron transport chain including a forward-operating complex I, thereby decreasing the need for reverse electron flow and increasing the overall energy yield from ammonia oxidation. This may be considered a niche-specific adaptation to the aerobic/anaerobic interface common to soil and sediment environments.

(vi) Nitrogen oxide metabolism. *N. multiformis*, like *N. europaea* and *N. eutropha*, is capable of reducing nitrite to NO to N₂O under oxic to microoxic conditions in a process known as nitrifier denitrification (23, 58). The observed levels of N₂O production from nitrite by *N. multiformis* were generally less than those by *N. europaea* but were still significant (58). Orthologs to copper-containing nitrite reductase, *nirK* (Nmul_A1998), and nitric oxide reductase, *norCBQD* (Nmul_A1256 to Nmul_A1253) were identified in the *N. multiformis* genome sequence. As with *N. europaea*, no CDS were identified with strong similarity to known dissimilatory nitrate (EC 1.7.99.4) or nitrous oxide reductases (EC 1.7.99.6), consistent with the physiological evidence that *N. multiformis* does not reduce nitrate or N₂O.

The *nirK* gene of *N. multiformis* is phylogenetically distinct from *nirK* of *N. europaea* and *N. oceani*, indicating that this gene originated from a different evolutionary or gene transfer event (15). The *N. multiformis nirK* gene exists as a singleton, rather than as a member of a multigene cluster as found in *N. europaea* and *N. eutropha*, and its promoter region lacks a regulatory binding motif for NO₂⁻ or NO-responsive transcription factors, such as NsrR, DNR (dissimilatory nitrate respiration regulator), or NnrR (8, 15, 55). NirK of *N. multiformis* is predicted to differ structurally from NirK in *N. europaea* and *N. oceani* in that it contains two (rather than one) type 1 Cu ligands, one related to plastocyanin at the N terminus and one related to nitrite reductase at the C terminus (15). These distinctive features indicate that NirK of *N. multiformis*, along with related NirK of other *Nitrosospira* spp. (15), may function differently and/or be expressed in response to different environmental signals than NirK of *N. europaea* or *N. oceani*.

The nitric oxide reductase *norCBQD* genes of *N. multiformis* are highly divergent from those of other AOB. For example, the catalytic NorB subunit is only 41 and 42% similar to the NorB of *N. europaea* and *N. oceani*, respectively, and 71% similar to NorB of the chemolithotrophic sulfur oxidizer *Thiobacillus denitrificans*. No studies have yet been conducted on NO reduction in *N. multiformis* or any other *Nitrosospira* spp. to assign a specific role to this enzyme in aerobic or anaerobic metabolism of NO. *N. multiformis* lacks an ortholog to the pan1-type multicopper oxidase that is associated with *nirK* in a multigene cluster in *N. europaea* and *N. eutropha* and resides close to a gene encoding cytochrome P460 in *N. oceani*. Unlike the other sequenced AOB, *N. multiformis* lacks an ortholog encoding cytochrome P460. However, an ortholog encoding the periplasmic nitrosocyanin protein (Nmul_A1601) was found; this ortholog is exclusive to the

AOB and has structural features implicating involvement in NO or N₂O metabolism (6).

(vii) Energy and carbon storage. *N. multiformis* has been shown to deposit glycogen primarily in the peripheral compartments of the cell (71). Genes encoding the functions for glycogen formation and utilization, including ADP-glucose pyrophosphorylase, glycogen synthase, 1,4 alpha-glucan branching enzyme, alpha-glucan phosphorylase, and glycoside hydrolase, are clustered (Nmul_A0715 to Nmul_A0719). Gene Nmul_A0382 encodes a polyphosphate kinase, indicating that *N. multiformis* can use polyphosphate as a source of energy. Additionally, an ATP-NAD kinase may catalyze the phosphorylation of NAD to NADP utilizing inorganic polyphosphate as a source of phosphorous (Nmul_A2422). The potential integration of polyphosphate and pyrophosphate into energy metabolism has also been reported recently for *N. oceani* (37). No genes for the synthesis of polyhydroxyalkonates (i.e., polyhydroxybutyrate) were found. As in the other AOB (2, 44), there are genes in *N. multiformis* that encode sucrose synthase (Nmul_A2266) and sucrose phosphate synthase/phosphatase (Nmul_A2267). Sucrose is known to be a compatible solute for resistance to moderate osmotic stress in bacteria (42) and also functions as a storage reserve in cyanobacteria (44).

Nitrogen metabolism. (i) Nitrogen assimilation. The presence of a gene coding for an NADP⁺-specific glutamate dehydrogenase (Nmul_A2447) suggests that when the external ammonium concentration is high (i.e., >1 mM), assimilation would be primarily via this low-affinity system. Since these levels are unlikely in soil solution except after recent fertilizer additions, it is not surprising that several options exist for high-affinity ammonia assimilation. In *N. multiformis*, these include a GS-GltS system (glutamine synthetase-glutamate synthase, which is also known as GS-GOGAT), with a ferredoxin-dependent glutamate synthase (68) and several asparagine synthases. The GS-GltS system genes include those encoding glutamine synthetase *glnA* (Nmul_A2288) and two possible genes for glutamate synthase (GltS) of the ferredoxin type more typical of cyanobacteria (Nmul_A1804 and Nmul_A1542). Nmul_A1542 is orthologous to genes found in *N. europaea* and *N. eutropha*. Members in the complex regulatory pathway include *glnB* (Nmul_A2536) encoding the regulatory protein PII, a *glnD* gene (Nmul_A2633) encoding the PII uridylyl transferase, and a putative *glnE* (Nmul_A1058) encoding the glutamine synthetase adenylyating enzyme. The PII-encoding gene (*glnB*) occurs in a cluster with other genes potentially associated with N metabolism (peptidase and transaminase) although not with *glnA* as in many proteobacteria. No PII-encoding gene was identified in *N. europaea* (16). Multiple pathways and complex regulatory schemes for N assimilation are likely crucial for responsiveness to variable and often limiting ammonia/ammonium supplies in soil environments.

Two nonclustered genes with similarity to assimilatory NAD(P)H-nitrite reductase, *nirB*, were identified (Nmul_A0171 and Nmul_A0355). Assimilatory nitrite reductases are generally encoded by *nirBD*; however, the *N. multiformis* genome lacks a *nirD* homolog. *nirB* genes are present in the *N. europaea* and *N. eutropha* genomes but absent from the genome of *N. oceani*. In a dissimilatory nitrite reductase (*nirK*)-deficient mutant of *N. europaea*, expression of the *nirB* gene was significantly down-regulated relative to its expression in wild-type *N.*

europaea (18), suggesting a functional, albeit uncharacterized, role.

(ii) Amino acid metabolism. Amino acid metabolism pathways are quite similar to those identified in *N. europaea* (16). For example, a superoperon for aromatic amino acid synthesis has gene synteny with those found in *N. europaea* and *Pseudomonas aeruginosa* (16, 75), except that *N. multiformis* also has a gene (Nmul_A2194) encoding a 3-deoxy-D-arabinoheptulosonate-7-phosphate synthase (EC 2.5.1.54) that is missing from the cluster in *N. europaea*.

Interestingly, two pathways for Asn synthesis are indicated; Nmul_A2519 and Nmul_A2668 encode two distinct Asn synthases (EC 6.3.5.4); both are amidotransferases using glutamine to convert L-Asp to Asn. These two *asnB* genes encode proteins that are only 29% identical to each other, similar to the two *asnB* genes found in *N. europaea*. Nmul_A2166 encodes an asparaginase (EC 3.5.1.1); this is one of the few degradative amidohydrolases identified in the sequenced AOB genomes. Together these observations suggest an important role for Asn in N metabolism and possibly N storage in *N. multiformis*. *N. multiformis* has both genes encoding the enzyme cyanophycin synthetase (Nmul_A2250 to Nmul_A2251) and two paralogs for the exopeptidase cyanophycinase (Nmul_A1689 and Nmul_A1900); the peptidase function is not encoded in the other sequenced AOB genomes. In cyanobacteria, cyanophycin acts as a storage compound for N because its Arg-poly(Asp) structure stores five N for every Arg-Asp (5).

In *N. multiformis*, 20 aminoacyl-tRNA synthetases have been identified (11). A specific GlnRS type (EC 6.1.1.18) (Nmul_A2083) and two nondiscriminatory type GlxRS (Nmul_A0797 and Nmul_A1604) were identified. No candidate for AsnRS (EC 6.1.1.22) was found in the genome. The AspRS encoded by gene Nmul_A0603 does contain the GAD domain typically found in AspRS involved in the indirect transamination route to Asn-tRNA-^{ASN} synthesis. Genes encoding a possible aspartyl/glutamyl-tRNA-Asn/Gln amidotransferase (*gatCAB* [Nmul_A0321 to Nmul_A0323]) are found, so indirect routes for charging both Asn and Gln are possible. It is somewhat surprising but not without precedent (46) that multiple *asnB* genes are present in an organism without an identifiable AsnRS. This suggests that Asn synthesized by the various AsnB proteins are incorporated by direct enzymatic, nonribosomal pathways for specialized functions.

(iii) Urea and polyamine cycling. Urease- and urea carboxylase-encoding genes are present in the *N. multiformis* genome. The *ure* operon (*ureABC*) and accessory genes (*ureD* and *ureEFG*) (Fig. 2B) are very similar in sequence and arrangement to those described for *Nitrosospora* sp. strain NpAV (40). Genes encoding an outer membrane (OM) porin responsive to urea (*fmdC*) and urea transporter membrane protein (Utp) were identified. The predicted Utp (321 aa) has 10 transmembrane α -helix segments with conserved loops that likely play a functional role in pore formation (47) and is 71% similar to the Utp from *N. europaea*. The vicinity of Utp in *N. europaea* is characterized by evidence of genomic rearrangement, including an integrase, frameshifts, and truncated remnant genes indicative of the loss of the urease-encoding genes through reductive genome evolution (2). In *N. multiformis*, a complex regulatory control of urea transport and *ure* gene transcription

by a two-component regulatory system and sensor proteins is indicated by the neighboring genes (Fig. 2B).

The biotin-containing urea carboxylase (Nmul_A0943) is followed by a putative allophanate (carboxyurea) hydrolase (Nmul_A0944). These two genes are similar to those found in *Oleomonas sagaranensi* (34) and *Pseudomonas* sp. strain ADP (57) that have verified urea carboxylase and allophanate hydrolase functions. Together, these may function as an ATP-dependent pathway for urea conversion to ammonium and bicarbonate such as mediated by urea amidolyase. *N. europaea*, *N. europaea*, and *N. oceani* lack the ortholog for allophanate hydrolase, although their putative urea (biotin) carboxylase-encoding genes are 69% identical to their orthologs in *N. multiformis*. The presence of ureolytic capacity in *N. multiformis* in contrast to the nitrosomonads (*N. europaea* and *N. europaea*) (40, 63) is adaptive for soil environments experiencing fluctuating urea concentrations and/or acidic pH (14, 51).

While *N. oceani* possesses the inventory for a complete urea cycle (37), in *N. multiformis* the gene encoding arginase (EC 3.5.3.1) was not identified, suggesting that the urea cycle is incomplete as was the case in *N. europaea* (16) and *N. europaea* (63). The *N. multiformis* genome, however, encodes other enzymes involved in Arg metabolism, including two putative Arg decarboxylases (EC 4.1.1.19), acetylornithine transferase (EC 2.6.1.11), ornithine carbonyltransferase (EC 2.1.3.3), and arginosuccinate synthase (EC 6.3.4.5) (Nmul_A1040 to Nmul_A1043). The presence of a second arginine decarboxylase-encoding gene (Nmul_A2669 [see Fig. S2 in the supplemental material]) and this putative enzyme complement are distinct from those in the other AOB and are likely related to differences in urea, polyamine (48), and cyanophycin metabolism in *N. multiformis*. Arg decarboxylases are also known to function in acid tolerance which may be an important ecological trait in soil AOB. The polyamine synthesis and transport functions that are encoded in the genome may have additional roles in the metabolism of cytoplasmic protectants under water potential stress conditions experienced in soil environments.

Cell structure and motility. (i) OM, capsule, and exopolysaccharides. Lipopolysaccharide (LPS) is a complex molecule found in the OMs of gram-negative bacteria generally consisting of the O side chain, core oligosaccharide, and lipid A. The lipid A and 2-keto-3-deoxyoctonic acid components are thought to be required for viability (5, 22). Pathway- and enzyme-encoding genes involved in the synthesis of the O side chain-specific and core oligosaccharides were identified. Although a *lpxA* gene (Nmul_A2199) encoding the essential enzyme for the first step of biosynthesis of lipid A was found, genes encoding a disaccharide synthetase (LpxB), deacetylase (LpxC), or acylase (LpxD) for biosynthesis of lipid X were not clearly identified, suggesting that the LPS structure may be different or synthesized differently in this soil organism. The gene complement for fatty acid biosynthesis is similar to that in *N. europaea* (16).

Bacteria growing in microcolonies and in biofilms on soil surfaces often have capsules and are embedded in large amounts of EPS. Capsular polysaccharides are typically glycolipids with phospholipid membrane anchors in the OM. Three clusters of genes encoding putative proteins with roles in production of EPS, LPS, and capsule synthesis were identified (Fig. 3). One cluster spanning from Nmul_A0238 to

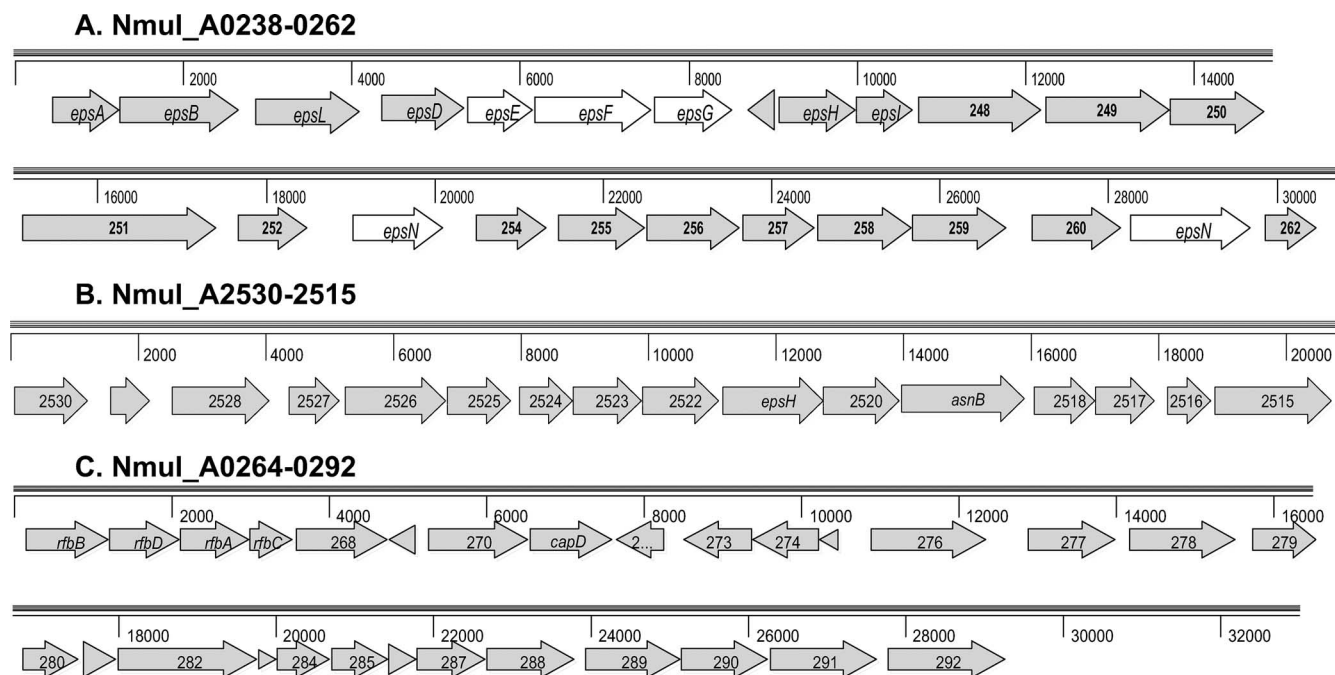


FIG. 3. EPS, LPS, and capsule clusters in *N. multiformis*. (A) Nmul_A0238 to Nmul_A0262 (Nmul_A0238-0262). EPS cluster with similarity to cluster in *Methylobacillus* sp. strain 12S (AB062506). Gene designations for orthologs in *Methylobacillus* sp. strain 12S are given, and genes essential for the production of the EPS methanolan in *Methylobacillus* sp. strain 12S are highlighted. (B) Putative EPS and LPS cluster including Nmul_A2530 to Nmul_A2515 (Nmul_A2530-2515) and their putative products and functions are as follows: Nmul_A2529, polysaccharide export to OM protein; Nmul_A2528, EPS biosynthesis chain length determinant similar to Wzz; Nmul_A2527, capsular polysaccharide biosynthesis protein; Nmul_A2526, signal peptide motif for secreted protein of unknown function; Nmul_A2525, type II transport ATPase; Nmul_A2524, polysaccharide deacetylase; Nmul_A2523, conserved hypothetical protein; Nmul_A2522, glycosyltransferase; Nmul_A2521 (*epsH*), EpsH (EPS locus protein H) involved in processing proteins to OM locations; Nmul_A2520, glycosyltransferase group 1, similar to *wapH* in *P. putida* (PP_4943); Nmul_A2519, asparagine synthase; Nmul_A2516, methyltransferase; Nmul_A2515, FkbH domain membrane protein involved in bacterial cell division. (C) Nmul_A0264-0292 putative EPS and LPS gene clusters. Nmul_A0264 to Nmul_A0292 (Nmul_A0264-0292) putative EPS and LPS gene clusters encoding proteins are shown as follows: Nmul_A0264 to Nmul_A0267, synthesis of dTDP-L-rhamnose (*rfbBDAC*); Nmul_A0268, fatty acid desaturase; Nmul_A0270, UDP-glucose/GDP-mannose dehydrogenase; Nmul_A0271, polysaccharide biosynthesis CapD; Nmul_A0277, dolichyl-P-beta-D-mannosyl transferase; Nmul_A0278 to Nmul_A0292, several glycosyltransferases, *capK* gene, and a polysaccharide biosynthesis protein often implicated in polysaccharide capsule biosynthesis.

Nmul_A0262 is unique among the sequenced AOB but has a similar gene arrangement and highest encoded product similarity to genes identified as essential to the synthesis of the EPS methanolan (a polymer of glucose, mannose, and galactose) in *Methylobacillus* sp. strain 12S (76) (Fig. 3A). The two clusters shown in Fig. 3B and C are at least partially represented in the other sequenced AOB. The production of N-linked glycoproteins may be facilitated by the presence of Asn synthase in the cluster of EPS-related and capsular genes (Fig. 3B). Also in this cluster is Nmul_A2520 that has similarity to the *Pseudomonas putida* (PP4943) *wapH* gene that has been implicated in LPS core synthesis and competitiveness in the biofilm niche (30). The third cluster (Fig. 3C) contains genes encoding the pathway for the synthesis of dTDP-L-rhamnose and several *cap* genes often implicated in polysaccharide capsule biosynthesis. In addition to these three clusters, two genes (Nmul_A0410 and Nmul_A0408) that encode synthesis of N-acetyl neuraminic acid and the capsular homopolysaccharide sialic acid were found. Overall, *N. multiformis* has the highest number (70) of CDS associated with carbohydrate active enzymes or binding modules among the sequenced AOB (<http://www.cazy.org/> [19]).

N. multiformis has 39 proteins that belong to The Institute

for Genomic Research (TIGR) PEP-CTERM family (TIGR 02595), a family identified by the novel conserved C-terminal domain with characteristics of protein sorting signals (28). Proteins with this designation are associated with gram-negative bacteria from soil and sediment environments capable of EPS synthesis, and all have orthologs encoding EPS locus protein H (EpsH) (Fig. 3). The putative function for the PEP-CTERM proteins is to aid in the targeting and transport of molecules through the inner membrane functioning similarly to the LPXTG/sortase systems in gram-positive bacteria (49). Many of the transported molecules may travel further to the OM and the cell exterior. Periplasmic glucans are adaptive for bacteria that undergo hypo-osmotic stress (upshock after rainfall) but may have additional roles in virulence and motility (54, 66). A gene encoding a putative cyclic beta-1,2-glucan synthetase was found (Nmul_A1183) containing five transmembrane helices and a glycosyltransferase motif.

(ii) **Motility, chemotaxis, and attachment.** Planktonically grown *N. multiformis* have 1 to 20 peritrichously located flagella (71), but these may not be evident in more typical biofilm growth. The motility and chemotaxis genes are contained in numerous operons within a three-tiered regulatory structure similar to that in other proteobacteria. A large region

TABLE 4. Summary of the 187 regulatory proteins in *N. multiformis*^a

Category	No. of genes
Transcription/elongation/termination factors.....	98
Sigma factors.....	8
Anti/anti-anti-sigma factors.....	6
Termination/antitermination factor.....	4
Elongation factors.....	3
Transcription factors.....	77
Signal transduction proteins.....	89
Chemotaxis signal transduction proteins.....	6
Chemotaxis sensory transducers (receptors).....	1
Other chemotaxis signal transduction proteins.....	5
Nonchemotaxis signal transduction.....	83
Signal transduction histidine kinases.....	23
Cyclic nucleotide signal transduction.....	21
PTS NTR regulator ^b	2
Miscellaneous.....	37

^a A detailed listing is available in Table S1 in the supplemental material. Associations and details of individual sensory transduction and regulatory response systems are available in Table S2 in the supplemental material.

^b PTS NTR regulator, phosphotransferase system nitrogen regulator.

(Nmul_A1305 to Nmul_A1357) starting with global regulatory transcriptional activators (FlhD and FlhC) is associated with flagellar synthesis, assembly, and regulatory control. Genes encoding a putative major pilus assembly protein PilA (Nmul_A2475) and several for PilF were found, although these may be involved in as yet uncharacterized membrane structures and filamentous secretory proteins. Nmul_A2575 encodes a large exoprotein (3,409 aa) similar to *M. capsulatus* MCA2227 with a hemagglutination activity domain followed by hemagglutinin repeat. Similar proteins have been implicated in adhesion, cell aggregation, and heme utilization. Adjacent genes (Nmul_A2569 to Nmul_A2591) encode secretion system proteins, surface antigens, pseudopilins, and other functions associated with the OM.

Communication and interaction with the environment. (i) Sensory and response regulator systems. The genome of *N. multiformis* contains a large number of signal transduction and sensory response systems of several types (39) (Tables 1 and 4). There are multiple examples of two-component signal transduction pathways that are used by bacteria to relay environmental signals and regulate cellular functions, often by functioning as transcription factors. The number and relative abundance of signal transducers may indicate the ability of an organism to adapt to diverse environmental conditions; soil bacteria typically have complex systems (25). Of the AOB examined to date, *N. multiformis* has both the highest absolute number and relative abundance of these systems (Table 4). In *N. multiformis*, signal transduction proteins include histidine kinases, methyl-accepting chemotaxis proteins, Ser/Thr protein kinases, adenylate and diguanylate cyclases, di-GMP phosphodiesterases, and adenylate cyclases. Some processes known to be regulated by these systems include EPS synthesis, biofilm formation, motility, and cell differentiation.

Acyl-homoserine lactones (AHL) are signal compounds that have been previously identified as active in the AOB (7). In contrast to the lack of a clear autoinducer synthase protein in *N. europaea*, Nmul_A2390 encodes an autoinducer synthesis

protein with more than 60% similarity to LasI found in *Burkholderia* and *Pseudomonas*. Sequence motifs suggest that this synthetase would produce a 3-oxo-homoserine lactone of indeterminate chain length (26, 72). Interestingly, *N. europaea* was responsive to 3-oxo-C6-homoserine lactone (7), although 3-oxo-homoserine lactone was not produced by *N. europaea* under the assay conditions examined (13). AHL signaling is known to be subject to cross talk and signal degradation in soil environments (70), so investigations of AHL identity and function in the environment are warranted for AOB.

(ii) Transport and protein secretion. Similar to *N. oceanii*, approximately 10.5% (298 CDS) of the *N. multiformis* genome is comprised of genes involved in active transport and protein secretion (Table 5). Thirty percent (87) of these CDS are subunits of type I ATP-binding cassette (ABC) transporters. Nine complete uptake ABC transporters were identified for sulfate/molybdate, phosphate, organic solvent, di- and oligopeptides, polyamine, ferric iron, and divalent cation transport. Nineteen ABC efflux systems were identified; these systems included numerous multidrug efflux pumps and multiple systems for export of the OM and cell wall components.

Mechanisms for iron transport are essential for maintaining the many cytochromes and putative heme-binding proteins involved in ammonia-oxidizing metabolism. Approximately 29 genes were identified for active transport of iron in *N. multiformis*, compared to ca. 90 in *N. europaea*, 28 in *N. eutropha*, and 22 in *N. oceanii* (16, 37, 63) (Table 5). Of the iron transporters, nine TonB-dependent iron siderophore receptors and 12 genes for the TonB/ExbB/ExbD-type energy transducers were identified, most presumably functioning in uptake of ferric iron. The other eight iron uptake genes were related to components of ABC transporters for FeS cluster biosynthesis, ferric iron uptake, siderophores, and ferric hydroxymate uptake (see Table S3 in the supplemental material). Like *N. oceanii*, no orthologs to *fecI-fecR-fecA* gene clusters were identified in the genome of *N. multiformis*, compared to 22 such iron transporter gene clusters found in *N. europaea* and a single gene cluster found in *N. eutropha*. However, three unclustered FecI-like sigma-24 (Nmul_A1839, Nmul_A1051, and Nmul_A1746) and one FecR-like response regulator genes (Nmul_A1046) were identified. A gene cluster for the transport and synthesis of a complex siderophore similar to the system producing pyoverdines in fluorescent pseudomonads (53) was identified (see Fig. S3 in the supplemental material); this was not found in any other sequenced AOB. This region is preceded by a gene encoding FecI (Nmul_A1839), further supporting a link to Fe-regulated siderophore production.

Transporters for inorganic nutrients, ions, and metals besides iron were also identified (Table 5) (see Table S3 in the supplemental material). For uptake of inorganic N, an ammonia permease, a NarK nitrate/nitrite transporter of the major facilitator superfamily (MFS), and a FNT-type nitrate/nitrite transporter were identified. Sulfate transporters include an ABC transporter, three periplasmic binding proteins, and a MFS transporter. Phosphate transporters include a complete ABC transporter, unlinked components of ABC transporters, and three phosphate-selective porins. For small-ion transport, an ABC transporter for Mn²⁺/Zn²⁺ was identified, as were ion channels for Mg²⁺/Co²⁺ and heavy metals and five P-type

TABLE 5. Transport-, efflux-, and secretion-related genes (265 total genes)^a

Class ^b	No. of genes	General function ^c
ABC transporter	87	Uptake and efflux substrates/toxins
Complete uptake systems (9) with ATPase, inner membrane, and periplasmic binding subunits	30	Sulfate/molybdate, spermidine/putrescine, ferric iron, phosphate, Mn ²⁺ /Zn ²⁺ , resistance to organic solvents (3), dipeptide/oligopeptide/nickel, cation/multidrug
Efflux pumps	35	Iron (cobalamin, Fe-S assembly, heme, siderophore), antimicrobial peptide, multidrug, lysophospholipase L1 biosynthesis, polysaccharide/polyol phosphate, protease/lipase, bacteriocin/antibiotic export
Periplasmic binding protein	7	Iron (hydroxymate), sulfate, branched-chain amino acid, organic solvents
Auxiliary proteins	4	
Unknown or orphans	10	
Pseudogene	1	
P-ATPase	5	Cation uptake/efflux ATPase
Iron uptake	21	
TonB/ExbB/ExbD	12	
TonB-dependent receptors	9	Siderophore/cobalamalamin
Transport of simple organic N compounds	14	
Urea	1	
FNT	1	Formate/nitrite
Na ⁺ /metabolite symporter	3	
Na ⁺ /H ⁺ antiporter	2	NhaA, NhaD
K ⁺ antiporter	4	TrkA
Antiporter flippases	3	
Efflux	22	
CDF	1	Co/Zn/Cd
CopD	2	Copper
TerC	2	Tellurium resistance
DMT superfamily	4	Cationic drug resistance
RND	13	Multidrug/cation/metals
Permease	24	
Ammonia	1	
MFS	10	Nitrate/nitrite, sulfate, and six uncharacterized efflux proteins
Tellurite resistance	1	
Amino acid	2	
YjgP/YjgQ	2	
Uncharacterized	8	
Potassium uptake	1	KUP
Transport of ions	13	
Potassium	2	
MgtE	6	Magnesium and cobalt transport
ZIP	1	Heavy metal cation transport
Mechanosensitive channel	2	
Unknown	2	
Porins	14	
Phosphate	3	OprP
Carbohydrate	1	OprB
Lipoprotein	3	
Polysaccharide	2	
Other	5	
Phosphotransferase	3	Fructose IIa, HPr, phosphoenolpyruvate phosphotransferase
Protein secretion	17	Sec pathway and TAT translocases
Type II secretion	28	General protein secretion/pilus formation
Type III secretion	11	Flagellum biosynthesis
Type IV secretion	5	TraG; DNA transfer during conjugation

^a Details of genes and COG designations are available in Table S3 in the supplemental material.

^b TNT, formate nitrate transporter; CDF, cation diffusion facilitator; DMT, drug metabolite transporter; RND, resistance nodulation division; MFS, major facilitator superfamily; ZIP, zinc transport (Zrt/Irt-like proteins).

^c Nha, Na⁺/H⁺ antiporter; Trk, transmembrane potassium transporter; KUP, potassium uptake protein; TAT, twin-arginine translocation.

ATPases for cation uptake/efflux. A single NhaA Na⁺/H⁺ antiporter was identified, along with a Na⁺ symporter.

K⁺ transport is thought to be an important property of soil bacteria that must acclimate to rapid changes in osmotic stress during wetting and drying of soil. A TrkA (low-affinity) K⁺ transport system was identified in *N. multiformis* along with a KUP-type K⁺ uptake system. An ortholog to this K⁺ transporter was found in the genome of *N. eutropha*, but not in the genomes of *N. europaea* and *N. oceani*. Two genes encoding mechanosensitive ion channels were found, which are thought to be involved in protecting bacterial cells from hypo-osmotic shock.

Systems for uptake of simple organic N compounds were identified including the following: a urea transporter (described above), an ABC transporter for uptake of dipeptides/oligopeptides, a periplasmic binding protein for branched-chain amino acids, and two amino acid transporters. A carbohydrate-selective barrel porin most similar to those from *Ralstonia* (OprB [Nmul_A2120]) was found in the genome of *N. multiformis*, but not in the genomes of other sequenced AOB. OprB porins are generally used for sugar transport across bacterial membranes. Like *N. europaea* and *N. oceani*, *N. multiformis* also has orthologs to the three components of the phosphotransferase-type phosphotransferase sugar transport system: a mannose/fructose-type IIa specific component, a HPr kinase/phosphorylase, and a phosphoenolpyruvate protein kinase (encoded by Nmul_A0218 to Nmul_A0220). This relatively small complement of genes involved in organic nutrient uptake verifies the specialization of *N. multiformis* for a chemolithotrophic lifestyle with some flexibility for uptake of simple sugars possibly used directly in EPS.

Mechanisms for efflux of organic compounds and other toxins are numerous in the *N. multiformis* genome (see Table S3 in the supplemental material). Among the ABC transporters, gene clusters containing all three components were identified for organic solvents and multidrug export. ABC efflux pumps were identified for multidrug export, heme, antimicrobial peptides, lipoproteins, bacteriocin, polysaccharide/polyol phosphate, organic solvents, protease/lipase, and lysophospholipase L1 biosynthesis. Efflux pumps were also found for heavy metals, copper, and tellurium.

The *N. multiformis* genome encodes protein secretion components of the Sec-dependent pathway, as well as components of type II, III, and IV protein secretion systems (see Table S3 in the supplemental material). Transport of OM proteins across the periplasm and targeting to the OM may be mediated by the Skp-like chaperone (Nmul_A0666) and the associated OM protein (Nmul_A0665) similar to YaeT in *E. coli* (22). Two *tolC* genes that potentially code for OM secretion proteins are present.

Conclusions: *N. multiformis* gene complement for the soil niche. The genome of *N. multiformis* exhibits several features that may be interpreted as adaptations for the soil niche. The high number and complexity of response regulatory networks and transporters indicate an extensive capacity for extracellular sensing and response and the acquisition of metals and inorganic nutrients. Numerous systems for efflux of metals, antimicrobial peptides, toxins, multidrug, and organic solvents may represent an adaptation for responding to stresses found in soils. The presence of both urease and hydrogenase gene

clusters indicate a flexibility in electron donors for possible energy production. The multiple copies of both *amo* and *hao* gene clusters and their regulatory elements indicate responsiveness to fluctuating ammonium availability. Glycogen, polyphosphate, sucrose, and cyanophycin may act in storage of carbon, energy, and nitrogen useful during substrate limitations for cell maintenance functions. The biofilm mode of existence depends on EPS production, and significant gene clusters are dedicated to this function. While the overall gene complement in *N. multiformis* reflects its close phylogenetic and functional relationship to *N. europaea*, its unique aspects reflect its adaptation and evolution in the soil habitat.

ACKNOWLEDGMENTS

Funding for this project was provided by the Department of Energy through the Microbial Genomes Program of the Office of Science. The Joint Genome Institute managed the overall sequencing effort, which was carried out by Lawrence Livermore National Laboratory under contract W-7405-Eng-48 under the auspices of the U.S. Department of Energy. Computational annotation was carried out at Oak Ridge National Laboratory (ORNL) and the Production Genomics Facility. Our thanks to Janet Chang (ORNL) for the use of her transporter identification tool prior to publication. A consortium of investigators from four universities and ORNL assisted in the analysis of the information made available from the sequencing effort. J.M.N. was supported by the Utah Agricultural Experiment Station and by grant US-3377-02 from BARD (U.S.-Israel Binational Agricultural Research and Development Fund). L.Y.S. was supported by the University of California at Riverside Agricultural Experiment Station. M.G.K. was supported in part, by incentive funds provided by the University of Louisville EVPR office, the Kentucky Science and Engineering Foundation (KSEF-787-RDE-007), and the National Science Foundation (EF-0412129). Publication was supported by NSF grant 0541797.

REFERENCES

- Aakra, A., M. Hesselsoe, and L. R. Bakken. 2000. Surface attachment of ammonia-oxidizing bacteria in soil. *Microb. Ecol.* **39**:222–235.
- Arp, D. J., P. S. G. Chain, and M. G. Klotz. 2007. The impact of genome analyses on our understanding of ammonia-oxidizing bacteria. *Annu. Rev. Microbiol.* **61**:503–528.
- Avrahami, S., and B. J. A. Bohannon. 2007. Response of *Nitrosospira* sp. strain AF-like ammonia oxidizers to changes in temperature, soil moisture content, and fertilizer concentration. *Appl. Environ. Microbiol.* **73**:1166–1173.
- Avrahami, S., W. Liesack, and R. Conrad. 2003. Effects of temperature and fertilizer on activity and community structure of soil ammonia oxidizers. *Environ. Microbiol.* **5**:691–705.
- Barton, L. L. 2005. Structural and functional relationships in prokaryotes. Springer Science+Business Media, Inc., New York, NY.
- Basumallick, L., R. Sarangi, S. D. George, B. Elmore, A. B. Hooper, B. Hedman, K. O. Hodgson, and E. I. Solomon. 2005. Spectroscopic and density functional studies of the red copper site in nitrosocyanin: role of the protein in determining active site geometric and electronic structure. *J. Am. Chem. Soc.* **127**:3531–3544.
- Batchelor, S. E., M. Cooper, S. R. Chhabra, L. A. Glover, G. Stewart, P. Williams, and J. I. Prosser. 1997. Cell density-regulated recovery of starved biofilm populations of ammonia-oxidizing bacteria. *Appl. Environ. Microbiol.* **63**:2281–2286.
- Beaumont, H. J. E., S. I. Lens, W. N. M. Reijnders, H. V. Westerhoff, and R. J. M. van Spanning. 2004. Expression of nitrite reductase in *Nitrosomonas europaea* involves NsrR, a novel nitrite-sensitive transcription repressor. *Mol. Microbiol.* **54**:148–158.
- Berube, P. M., R. Samudrala, and D. A. Stahl. 2007. Transcription of all *amoC* copies is associated with recovery of *Nitrosomonas europaea* from ammonia starvation. *J. Bacteriol.* **189**:3935–3944.
- Bodelier, P. L. E., J. A. Libochant, C. Blom, and H. J. Laanbroek. 1996. Dynamics of nitrification and denitrification in root-oxygenated sediments and adaptation of ammonia-oxidizing bacteria to low-oxygen or anoxic habitats. *Appl. Environ. Microbiol.* **62**:4100–4107.
- Brown, J. R., and W. F. Doolittle. 1999. Gene descent, duplication, and horizontal transfer in the evolution of glutamyl- and glutaminyl-tRNA synthetases. *J. Mol. Evol.* **49**:485–495.
- Bruns, M. A., J. R. Stephen, G. A. Kowalchuk, J. I. Prosser, and E. A. Paul.

1999. Comparative diversity of ammonia oxidizer 16S rRNA gene sequences in native, tilled, and successional soils. *Appl. Environ. Microbiol.* **65**:2994–3000.
13. **Burton, E. O., H. W. Read, M. C. Pelletteri, and W. J. Hickey.** 2005. Identification of acyl-homoserine lactone signal molecules produced by *Nitrosomonas europaea* strain Schmidt. *Appl. Environ. Microbiol.* **71**:4906–4909.
14. **Burton, S. A. Q., and J. I. Prosser.** 2001. Autotrophic ammonia oxidation at low pH through urea hydrolysis. *Appl. Environ. Microbiol.* **67**:2952–2957.
15. **Cantera, J. J. L., and L. Y. Stein.** 2007. Molecular diversity of nitrite reductase genes (*nirK*) in nitrifying bacteria. *Environ. Microbiol.* **9**:765–776.
16. **Chain, P., J. Lamerdin, F. Larimer, W. Regala, V. Lao, M. Land, L. Hauser, A. Hooper, M. Klotz, J. Norton, L. Sayavedra-Soto, D. Arciero, N. Hommes, M. Whittaker, and D. Arp.** 2003. Complete genome sequence of the ammonia-oxidizing bacterium and obligate chemolithoautotroph *Nitrosomonas europaea*. *J. Bacteriol.* **185**:2759–2773.
17. **Chain, P. S. G., V. J. Denef, K. T. Konstantinidis, L. M. Vergez, L. Agullo, V. L. Reyes, L. Hauser, M. Cordova, L. Gomez, M. Gonzalez, M. Land, V. Lao, F. Larimer, J. J. Lipuma, E. Mahenthiralingam, S. A. Malfatti, C. J. Marx, J. J. Parnell, A. Ramette, P. Richardson, M. Seeger, D. Smith, T. Spilker, W. J. Sul, T. V. Tsoi, L. E. Ulrich, I. B. Zhulin, and J. M. Tiedje.** 2006. *Burkholderia xenovorans* LB400 harbors a multi-replicon, 9.73-Mbp genome shaped for versatility. *Proc. Natl. Acad. Sci. USA* **103**:15280–15287.
18. **Cho, C. M. H., T. F. Yan, X. D. Liu, L. Y. Wu, J. Z. Zhou, and L. Y. Stein.** 2006. Transcriptome of a *Nitrosomonas europaea* mutant with a disrupted nitrite reductase gene (*nirK*). *Appl. Environ. Microbiol.* **72**:4450–4454.
19. **Coutinho, P. M., and B. Henrissat.** 1999. Carbohydrate-active enzymes: an integrated database approach, p. 3–12. *In* G. D. H. J. Gilbert, B. Henrissat, and B. Svensson (ed.), Recent advances in carbohydrate bioengineering. The Royal Society of Chemistry, Cambridge, United Kingdom.
20. **Csaki, R., T. Hanczar, L. Bodrossy, J. C. Murrell, and K. L. Kovacs.** 2001. Molecular characterization of structural genes coding for a membrane bound hydrogenase in *Methylococcus capsulatus* (Bath). *FEMS Microbiol. Lett.* **205**:203–207.
21. **De Boer, W., and G. A. Kowalchuk.** 2001. Nitrification in acid soils: microorganisms and mechanisms. *Soil Biol. Biochem.* **33**:853–866.
22. **Doerrler, W. T.** 2006. Lipid trafficking to the outer membrane of Gram-negative bacteria. *Mol. Microbiol.* **60**:542–552.
23. **Dundee, L., and D. W. Hopkins.** 2001. Different sensitivities to oxygen of nitrous oxide production by *Nitrosomonas europaea* and *Nitrosolobus multiformis*. *Soil Biol. Biochem.* **33**:1563–1565.
24. **El Sheikh, A. F., A. T. Poret-Peterson, and M. G. Klotz.** 2008. Characterization of two new genes, *amoR* and *amoD*, in the *amo* operon of the marine ammonia oxidizer *Nitrosococcus oceanii* ATCC 19707. *Appl. Environ. Microbiol.* **74**:312–318.
25. **Galperin, M. Y.** 2005. A census of membrane-bound and intracellular signal transduction proteins in bacteria: bacterial IQ, extroverts and introverts. *BMC Microbiol.* **5**:35.
26. **Gould, T. A., J. Herman, J. Krank, R. C. Murphy, and M. E. A. Churchill.** 2006. Specificity of acyl-homoserine lactone synthases examined by mass spectrometry. *J. Bacteriol.* **188**:773–783.
27. **Grundmann, G. L., A. Dechesne, F. Bartoli, J. P. Flandrois, J. L. Chasse, and R. Kizungu.** 2001. Spatial modeling of nitrifier microhabitats in soil. *Soil Sci. Soc. Am. J.* **65**:1709–1716.
28. **Haft, D. H., I. T. Paulsen, N. Ward, and J. D. Selengut.** 2006. Exopolysaccharide-associated protein sorting in environmental organisms: the PEP-CTERM/EpsH system. Application of a novel phylogenetic profiling heuristic. *BMC Biol.* **4**:16.
29. **Hanczar, T., R. Csaki, L. Bodrossy, J. C. Murrell, and K. L. Kovacs.** 2002. Detection and localization of two hydrogenases in *Methylococcus capsulatus* (Bath) and their potential role in methane metabolism. *Arch. Microbiol.* **177**:167–172.
30. **Hansen, S. K., J. A. J. Haagensen, M. Gjermansen, T. M. Jorgensen, T. Tolker-Nielsen, and S. Molin.** 2007. Characterization of a *Pseudomonas putida* rough variant evolved in a mixed-species biofilm with *Acinetobacter* sp. strain C6. *J. Bacteriol.* **189**:4932–4943.
31. **Hommes, N. G., E. G. Kurth, L. A. Sayavedra-Soto, and D. J. Arp.** 2006. Disruption of *sucA*, which encodes the E1 subunit of α -ketoglutarate dehydrogenase, affects the survival of *Nitrosomonas europaea* in stationary phase. *J. Bacteriol.* **188**:343–347.
32. **Hommes, N. G., L. A. Sayavedra-Soto, and D. J. Arp.** 2003. Chemolithoautotrophic growth of *Nitrosomonas europaea* on fructose. *J. Bacteriol.* **185**:6809–6814.
33. **Hommes, N. G., L. A. Sayavedra-Soto, and D. J. Arp.** 1998. Mutagenesis and expression of *amo*, which codes for ammonia monooxygenase in *Nitrosomonas europaea*. *J. Bacteriol.* **180**:3353–3359.
34. **Kanamori, T., N. Kanou, H. Atomi, and T. Imanaka.** 2004. Enzymatic characterization of a prokaryotic urea carboxylase. *J. Bacteriol.* **186**:2532–2539.
35. **Kelly, D. P., C. Anthony, and J. C. Murrell.** 2005. Insights into the obligate methanotroph *Methylococcus capsulatus*. *Trends Microbiol.* **13**:195–198.
36. **Kingsford, C. L., K. Ayanbule, and S. L. Salzberg.** 2007. Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. *Genome Biol.* **8**:R22.
37. **Klotz, M. G., D. J. Arp, P. S. G. Chain, A. F. El-Sheikh, L. J. Hauser, N. G. Hommes, F. W. Larimer, S. A. Malfatti, J. M. Norton, A. T. Poret-Peterson, L. M. Vergez, and B. B. Ward.** 2006. Complete genome sequence of the marine, chemolithoautotrophic, ammonia-oxidizing bacterium *Nitrosococcus oceanii* ATCC 19707. *Appl. Environ. Microbiol.* **72**:6299–6315.
38. **Klotz, M. G., and L. Y. Stein.** 2008. Nitrifier genomics and evolution of the nitrogen cycle. *FEMS Microbiol. Lett.* **278**:146–156.
39. **Konstantinidis, K. T., and J. M. Tiedje.** 2004. Trends between gene content and genome size in prokaryotic species with larger genomes. *Proc. Natl. Acad. Sci. USA* **101**:3160–3165.
40. **Koper, T. E., A. F. El-Sheikh, J. M. Norton, and M. G. Klotz.** 2004. Urease-encoding genes in ammonia-oxidizing bacteria. *Appl. Environ. Microbiol.* **70**:2342–2348.
41. **Kowalchuk, G. A., and J. R. Stephen.** 2001. Ammonia-oxidizing bacteria: a model for molecular microbial ecology. *Annu. Rev. Microbiol.* **55**:485–529.
42. **Kunte, H. J.** 2006. Osmoregulation in bacteria: compatible solute accumulation and osmosensing. *Environ. Chem.* **3**:94–99.
43. **Leininger, S., T. Urich, M. Schloter, L. Schwark, J. Qi, G. W. Nicol, J. I. Prosser, S. C. Schuster, and C. Schleper.** 2006. Archaea predominate among ammonia-oxidizing prokaryotes in soils. *Nature* **442**:806–809.
44. **Lunn, J. E.** 2002. Evolution of sucrose synthesis. *Plant Physiol.* **128**:1490–1500.
45. **Mendum, T. A., and P. R. Hirsch.** 2002. Changes in the population structure of beta-group autotrophic ammonia oxidizing bacteria in arable soils in response to agricultural practice. *Soil Biol. Biochem.* **34**:1479–1485.
46. **Min, B., J. T. Pelaschier, D. E. Graham, D. Tumbula-Hansen, and D. Soll.** 2002. Transfer RNA-dependent amino acid biosynthesis: an essential route to asparagine formation. *Proc. Natl. Acad. Sci. USA* **99**:2678–2683.
47. **Minocha, R., K. Studley, and M. H. Saier.** 2003. The urea transporter (UT) family: bioinformatic analyses leading to structural, functional, and evolutionary predictions. *Receptors Channels* **9**:345–352.
48. **Nakada, Y., and Y. Itoh.** 2003. Identification of the putrescine biosynthetic genes in *Pseudomonas aeruginosa* and characterization of arginine deiminase and *N*-carbamoylputrescine amidohydrolase of the arginine decarboxylase pathway. *Microbiology* **149**:707–714.
49. **Navarre, W. W., and O. Schneewind.** 1999. Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. *Microbiol. Mol. Biol. Rev.* **63**:174–229.
50. **Norton, J. M., J. J. Alzerreca, Y. Suwa, and M. G. Klotz.** 2002. Diversity of ammonia monooxygenase operon in autotrophic ammonia-oxidizing bacteria. *Arch. Microbiol.* **177**:139–149.
51. **Pommerening-Roser, A., and H. P. Koops.** 2005. Environmental pH as an important factor for the distribution of urease positive ammonia-oxidizing bacteria. *Microbiol. Res.* **160**:27–35.
52. **Prosser, J. I., and T. M. Embley.** 2002. Cultivation-based and molecular approaches to characterisation of terrestrial and aquatic nitrifiers. *Antonie van Leeuwenhoek* **81**:165–179.
53. **Ravel, J., and P. Cornelis.** 2003. Genomics of pyoverdine-mediated iron uptake in pseudomonads. *Trends Microbiol.* **11**:195–200.
54. **Rediers, H., P. B. Rainey, J. Vanderleyden, and R. De Mot.** 2005. Unraveling the secret lives of bacteria: use of in vivo expression technology and differential fluorescence induction promoter traps as tools for exploring niche-specific gene expression. *Microbiol. Mol. Biol. Rev.* **69**:217–261.
55. **Rodionov, D. A., I. L. Dubchak, A. P. Arkin, E. J. Alm, and M. S. Gelfand.** 2005. Dissimilatory metabolism of nitrogen oxides in bacteria: comparative reconstruction of transcriptional networks. *PLoS Comp. Biol.* **1**:415–431.
56. **Schwartz, E., U. Gerischer, and B. Friedrich.** 1998. Transcriptional regulation of *Alcaligenes eutrophus* hydrogenase genes. *J. Bacteriol.* **180**:3197–3204.
57. **Shapir, N., M. J. Sadowsky, and L. P. Wackett.** 2005. Purification and characterization of allophanate hydrolase (AtzF) from *Pseudomonas* sp. strain ADP. *J. Bacteriol.* **187**:3731–3738.
58. **Shaw, L. J., G. W. Nicol, Z. Smith, J. Fear, J. I. Prosser, and E. M. Baggs.** 2006. *Nitrosospira* spp. can produce nitrous oxide via a nitrifier denitrification pathway. *Environ. Microbiol.* **8**:214–222.
59. **Shively, J. M., G. van Keulen, and W. G. Meijer.** 1998. Something from almost nothing: carbon dioxide fixation in chemoautotrophs. *Annu. Rev. Microbiol.* **52**:191–230.
60. **Snel, B., P. Bork, and M. A. Huynen.** 2002. The identification of functional modules from the genomic association of genes. *Proc. Natl. Acad. Sci. USA* **99**:5890–5895.
61. **Stark, J. M., and M. K. Firestone.** 1995. Mechanisms for soil moisture effects on activity of nitrifying bacteria. *Appl. Environ. Microbiol.* **61**:218–221.
62. **Stec, B., H. Y. Yang, K. A. Johnson, L. J. Chen, and M. F. Roberts.** 2000. MJ0109 is an enzyme that is both an inositol monophosphatase and the 'missing' archaeal fructose 1,6-bisphosphatase. *Nat. Struct. Biol.* **7**:1046–1050.
63. **Stein, L. Y., D. J. Arp, P. M. Berube, P. S. G. Chain, L. Hauser, M. S. M. Jetten, M. G. Klotz, F. W. Larimer, J. M. Norton, H. J. M. Op den Camp, M. Shin, and X. Wei.** 2007. Whole-genome analysis of the ammonia-oxidizing bacterium, *Nitrosomonas europaea* C91: implications for niche adaptation. *Environ. Microbiol.* **9**:2993–3007.
64. **Stephen, J. R., A. E. McCaig, Z. Smith, J. I. Prosser, and T. M. Embley.** 1996. Molecular diversity of soil and marine 16S rRNA gene sequences related to

- β -subgroup ammonia-oxidizing bacteria. *Appl. Environ. Microbiol.* **62**:4147–4154.
65. **Tabita, F. R., T. E. Hanson, H. Li, S. Satagopan, J. Singh, and S. Chan.** 2007. Function, structure, and evolution of the RubisCO-like proteins and their RubisCO homologs. *Microbiol. Mol. Biol. Rev.* **71**:576–599.
66. **Talaga, P., V. Coge, J. M. Wieruszkeski, B. Stahl, J. Lemoine, G. Lippens, and J. P. Bohin.** 2002. Osmoregulated periplasmic glucans of the free-living photosynthetic bacterium *Rhodobacter sphaeroides*. *Eur. J. Biochem.* **269**:2464–2472.
67. **Utåker, J. B., K. Andersen, A. Aakra, B. Moen, and I. F. Nes.** 2002. Phylogeny and functional expression of ribulose 1,5-bisphosphate carboxylase/oxygenase from the autotrophic ammonia-oxidizing bacterium *Nitrosospira* sp. isolate 40KI. *J. Bacteriol.* **184**:468–478.
68. **Vanoni, M. A., L. Dossena, R. H. H. van den Heuvel, and B. Curti.** 2005. Structure-function studies on the complex iron-sulfur flavoprotein glutamate synthase: the key enzyme of ammonia assimilation. *Photosynthesis Res.* **83**:219–238.
69. **Vignais, P. M., B. Billoud, and J. Meyer.** 2001. Classification and phylogeny of hydrogenases. *FEMS Microbiol. Rev.* **25**:455–501.
70. **Wang, Y. J., and J. R. Leadbetter.** 2005. Rapid acyl-homoserine lactone quorum signal biodegradation in diverse soils. *Appl. Environ. Microbiol.* **71**:1291–1299.
71. **Watson, S. W., L. B. Graham, C. C. Remsen, and F. W. Valois.** 1971. A lobular, ammonia-oxidizing bacterium, *Nitrosolobus multiiformis* nov. gen. nov. sp. *Arch. Microbiol.* **76**:183–203.
72. **Watson, W. T., T. D. Minogue, D. L. Val, S. B. von Bodman, and M. E. A. Churchill.** 2002. Structural basis and specificity of acyl-homoserine lactone signal production in bacterial quorum sensing. *Mol. Cell* **9**:685–694.
73. **Wei, X. M., L. A. Sayavedra-Soto, and D. J. Arp.** 2004. The transcription of the *ccb* operon in *Nitrosomonas europaea*. *Microbiology* **150**:1869–1879.
74. **Wood, A., J. Aurikko, and D. Kelly.** 2004. A challenge for 21st century molecular biology and biochemistry: what are the causes of obligate autotrophy and methanotrophy? *FEMS Microbiol. Rev.* **28**:335–352.
75. **Xie, G., C. A. Bonner, and R. A. Jensen.** 1999. A probable mixed-function supraoperon in *Pseudomonas* exhibits gene organization features of both intergenomic conservation and gene shuffling. *J. Mol. Evol.* **49**:108–121.
76. **Yoshida, T., Y. Ayabe, M. Yasunaga, Y. Usami, H. Habe, H. Nojiri, and T. Omori.** 2003. Genes involved in the synthesis of the exopolysaccharide methanolan by the obligate methylotroph *Methylobacillus* sp. strain 12S. *Microbiology* **149**:431–444.