Nitrosopumilus maritimus genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea

C. B. Walker[a,b], J. R. de la Torre[a], M. G. Klotz[c], H. Urakawa[a], N. Pinel[a], D. J. Arp[d], C. Brochier-Armanet[e], P. S. G. Chain[f,g,h], P. P. Chan[i], A. Gollabgir[j], J. Hemp[k], M. Hügler[l,m], E. A. Karr[n], M. Könneke[o], M. Shin[f,g], T. J. Lawton[p], T. Lowe[i], W. Martens-Habbena[a], L. A. Sayavedra-Soto[d], D. Lang[f,g], S. M. Sievert[q], A. C. Rosenzweig[p], G. Manning[j], and D. A. Stahl[a,1]

[a]Department of Civil and Environmental Engineering, University of Washington, Seattle, WA 98195; [b]Geosyntec Consultants, Seattle, WA 98101; [c]Department of Biology, University of Louisville, Louisville, KY 40292; [d]Department of Botany and Plant Pathology, Oregon State University, Corvalis, OR 97331; [e]Université de Provence Aix-Marseille I, Laboratoire de Chimie Bactérienne, Centre National de la Recherche Scientifique Unité Propre de Recherche, Marseille, 13402 France; [f]Biosciences Division, Lawrence Livermore National Laboratory, Livermore, CA 94550; [g]Microbial Program, Joint Genome Institute, Walnut Creek, CA 94598; [h]Center for Microbial Ecology, Michigan State University, East Lansing, MI 48824; [i]Department of Biomolecular Engineering, University of California, Santa Cruz, CA 95064; [j]Razavi Newman Center for Bioinformatics, Salk Institute for Biological Studies, La Jolla, CA 92037; [k]School of Chemical Sciences, University of Illinois, Urbana, IL 61801; [l]Leibniz-Institut für Meereswissenschaften, Kiel, 24105 Germany; [m]Water Technology Center, Karlsruhe, 76139 Germany; [n]Department of Botany and Microbiology, University of Oklahoma, Norman, OK 73019; [o]Institut für Chemie und Biologie des Meeres, Universität Oldenburg, Oldenburg, 26129 Germany; [p]Departments of Biochemistry, Molecular Biology and Cell Biology, and Chemistry, Northwestern University, Evanston, IL 60208; and [q]Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA 02543

Edited by David Karl, University of Hawaii, Honolulu, HI, and approved April 2, 2010 (received for review December 6, 2009)

Ammonia-oxidizing archaea are ubiquitous in marine and terrestrial environments and now thought to be significant contributors to carbon and nitrogen cycling. The isolation of *Candidatus* "*Nitrosopumilus maritimus*" strain SCM1 provided the opportunity for linking its chemolithotrophic physiology with a genomic inventory of the globally distributed archaea. Here we report the 1,645,259-bp closed genome of strain SCM1, revealing highly copper-dependent systems for ammonia oxidation and electron transport that are distinctly different from known ammonia-oxidizing bacteria. Consistent with in situ isotopic studies of marine archaea, the genome sequence indicates *N. maritimus* grows autotrophically using a variant of the 3-hydroxypropionate/4-hydroxybutyrate pathway for carbon assimilation, while maintaining limited capacity for assimilation of organic carbon. This unique instance of archaeal biosynthesis of the osmoprotectant ectoine and an unprecedented enrichment of multicopper oxidases, thioredoxin-like proteins, and transcriptional regulators points to an organism responsive to environmental cues and adapted to handling reactive copper and nitrogen species that likely derive from its distinctive biochemistry. The conservation of *N. maritimus* gene content and organization within marine metagenomes indicates that the unique physiology of these specialized oligophiles may play a significant role in the biogeochemical cycles of carbon and nitrogen.

ammonia oxidation | marine microbiology | archaea | nitroxyl

Marine Group I archaea are among the most abundant microorganisms in the global oceans (1–3). Originally discovered through ribosomal RNA gene sequencing (3, 4), recent metagenomic, biogeochemical, and microbiological studies established the capacity of these organisms to oxidize ammonia, thus linking this abundant microbial clade to one of the key steps of the global nitrogen cycle (5–9). For a century following the discovery of autotrophic ammonia oxidizers, only *Bacteria* were thought to catalyze this generally rate-limiting transformation in the two-step process of nitrification (10). Despite recent enrichment of mesophilic as well as thermophilic ammonia-oxidizing archaea (AOA) (6, 11, 12), only a single Group I-related strain, isolated from a gravel inoculum from a tropical marine aquarium, has thus far been successfully obtained in pure culture (7).

The isolation of *Nitrosopumilus maritimus* strain SCM1 ultimately confirmed an archaeal capacity for chemoautotrophic growth on ammonia. More detailed characterization of this strain revealed cytological and physiological adaptations critical for life in an oligotrophic open ocean environment, most notably one of the highest substrate affinities yet observed (13). Among characterized ammonia oxidizers, only *N. maritimus* is capable growing at the extremely low concentrations of ammonia generally found in the open ocean (7, 13). This strain therefore provided an excellent opportunity to investigate the core genetic inventory for ammonia-based chemoautotrophy by Group I crenarchaea.

The gene content and gene order of *N. maritimus* is highly similar to environmental populations represented in marine bacterioplankton metagenomes, confirming on a genomic level its close relationship to many oceanic crenarchaea. Thus, an evaluation of the genomic inventory of *N. maritimus* should offer a framework to identify features shared among ammonia-oxidizing Group I crenarchaea, resolve physiological diversity among AOA, and refine understanding of their ecology in relationship to the larger assemblage of marine archaea—not all of which are ammonia oxidizers. In support of this expectation, the physiological and genomic profiles together show that many of the "non-extreme" archaea identified in metagenomic studies, and currently assigned to the *Crenarchaeota* kingdom, are AOA that contribute to global carbon and nitrogen cycling, possibly determining rates of nitrification in a variety of environments (6, 8, 9, 13).

Results and Discussion

Primary Sequence Characteristics. *N. maritimus* strain SCM1 contains a single chromosome of 1,645,259 bp encoding 1,997 predicted genes and no extrachromosomal elements or complete prophage sequences (Table 1). No unambiguous origin of replication could be determined on the basis of local gene content or GC skew, as commonly observed for other archaeal genomes (14). Approximately 61% of the *N. maritimus* open-reading

Author contributions: C.B.W., J.R.d.l.T., P.S.G.C., and D.A.S. designed research; C.B.W., J.R.d.l.T., M.G.K., H.U., N.P., C.B-A., P.P.C., A.G., M.H., E.A.K., M.K., M.S., T.L., W.M-H., M.S., D.L., S.M.S., A.C.R., G.M., and D.A.S. performed research; C.B.W. and J.R.d.l.T. contributed new reagents/analytic tools; C.B.W., J.R.d.l.T., M.G.K., H.U., N.P., D.J.A., C.B.-A., P.P.C., A.G., J.H., M.H., E.A.K., M.K., T.J.L., T.L., W.M.-H., L.A.S.-S., S.M.S., A.C.R., G.M., and D.A.S. analyzed data; and C.B.W., J.R.d.l.T., M.G.K., H.U., N.P., C.B.-A., J.H., M.H., E.A.K., M.K., T.L., S.M.S., A.C.R., G.M., and D.A.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The sequence reported in this paper has been deposited in the NCBI database (accession no. NC_010085).

[1]To whom correspondence should be addressed. E-mail: dastahl@u.washington.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.0913533107/-/DCSupplemental.

Table 1. Genome features of *N. maritimus* SCM1, *C. symbiosum*, sequenced AOB, and crenarchaeal genome fragments

	Nitrosopumilus maritimus SCM1	*Cenarchaeum symbiosum*	*Nitrosococcus oceani* ATCC 19707	*Nitrosomonas europaea* ATCC 19718	*Nitrosomonas eutropha* C91	*Nitrosospira multiformis* ATCC 25196	Fosmid 4B7	Cosmid DeepAnt-EC39	Fosmid 74A4
Size (bp)	1,645,259	2,045,086	3,481,691	2,812,094	2,661,057	3,184,243	39,297	33,347	43,902
Percent coding	91.90%	91.20%	86.80%	88.40%	85.60%	85.60%	89.10%	86.10%	84.00%
GC content	34.20%	57.70%	50.30%	50.70%	48.50%	53.90%	34.40%	34.10%	32.60%
ORFs	1,997	2,066	3,186	2,628	2,578	2,827	41	41	51
ORF density (ORF/kb)	1.19	0.986	0.889	0.876	0.952	0.86	0.992	1.17	1.12
Avg. ORF length (bp)	757	924	964	1009	890	980	898	737	753
Standard tRNAs	44	45	45	41	41	43	0	0	0
rRNAs	1	1	2	1	1	1	1	1	1
Plasmids	0	0	1	0	2	3	NA	NA	NA

NA, not analyzed.

frames (ORFs) could be assigned to clusters of orthologous groups of proteins (COGs), a lower percentage than for genomes of ammonia-oxidizing bacteria (AOB) (Table S1) but similar to *Cenarchaeum symbiosum* (15). The genome possesses a relatively high coding density (91.9%), with a larger fraction dedicated to energy production/conservation, coenzyme transport/metabolism, and translation genes than other characterized *Crenarchaeota*, but similar to two common species of photoautotrophic marine *Bacteria*, *Prochlorococcus*, and *Synechococcus*.

Energy Metabolism. The stoichiometry of ammonia oxidation to nitrite is similar to that of characterized aerobic, obligate chemolithoautotrophic AOB (13), yet the contributing biochemistry is distinctly unique. All AOB share a common pathway where hydroxylamine, produced by an ammonia monooxygenase (AMO), is oxidized to nitrite by a heme-rich hydroxylamine oxidoreductase (HAO) complex; the oxidation of hydroxylamine supplies electrons to both the AMO and a typical electron transport chain composed of cytochrome *c* proteins. *N. maritimus* lacks genes encoding a recognizable AOB-like HAO complex and pertinent cytochrome *c* proteins, indicating an alternative archaeal pathway. The numerous copper-containing proteins, including multicopper oxidases and small blue copper-containing proteins (similar to plasto-, halo-, and sulfocyanins), suggest an alternative electron transfer mechanism (Table S2). These predicted periplasmic proteins likely serve functionally similar roles to soluble cytochrome *c* proteins in other *Archaea*, including *Natronomonas pharaonis* and thermoacidophiles such as *Sulfolobus* (16). This apparent reliance on copper for redox reactions is a major divergence from the iron-based electron transfer system of AOB.

The *N. maritimus* genome contains genes coding for six soluble periplasmic multicopper oxidase (MCO) proteins: two nearly identical NO-forming nitrite reductase proteins (NirK; Nmar_1259 and -1667), each with three cupredoxin domains; two NcgA-like (*nirK* cluster gene A) MCOs (Nmar_1131 and -1663) with two cupredoxin domains; one MCO (Nmar_1136) with three cupredoxin domains; and one MCO (Nmar_1354) with two domains fused to a blue copper-containing protein (Table S2). Two (Nmar_1131 and Nmar_1663) of the three genes that are classified as belonging to the emerging class of two-domain MCOs (2dMCOs) resemble the general architecture of the 2dMCO NcgA present in *Nitrosomonas europaea*. Although the overall sequence identity is low, clustering of NcgA with a nitrite reductase suggests it may play a supporting role in nitrite reduction (17). Genes Nmar_1131 and Nmar_1663 are also colocated with a member of the DtxR family of metal regulators and a member of the ZIP metal transport family, suggesting a role in metal homeostasis (see below). The third 2dMCO (Nmar_1354), possessing a fused blue copper-containing domain, has not been found in AOB and appears to be unique to AOA. Redox interactions with the MCOs (and other predicted redox proteins) are likely mediated by eight soluble and nine membrane-anchored copper-binding proteins containing plastocyanin-like domains (Table S2). The corresponding genes appear to be the result of a series of duplications within the *N. maritimus* lineage (Fig. S1).

A second family of predicted redox active periplasmic proteins, composed of 11 thiol-disulfide oxidoreductases from the thioredoxin family (Nmar_0639, _0655, _0829, _0881, _1140, _1143, _1148, _1150, _1181, _1658, and Nmar_1670), show low (but recognizable) identity with the better characterized disulfide bond oxidases/isomerases found in *Bacillus subtilis* (BdbD) and *Escherichia coli* (DsbA, DsbC, and DsbG). The mean percentage of sequence identity between the *N. maritimus* proteins and BdbD is 21 ± 3%. The significantly lower mean percentage of sequence identity to DsbA, DsbC, and DsbG (9.2, 10.7, and 10.4%, respectively) is comparable to that shared between the *E. coli* proteins (10–11%). Although functional equivalency cannot be established, all but Nmar_0881 preserve the conserved thioredoxin-like active site FX_4CXXC sequence (18–20). In *E. coli*, both DsbA and DsbC rectify nonspecific disulfide bonds catalyzed by copper (21), whereas up-regulation of *dsbA* by the Cpx regulon occurs during copper stress (22, 23). Eukaryotic protein disulfide isomerase (PDI) homologs sequester and/or reduce oxidized Cu(II), perhaps serving as copper acceptors/donors for copper-containing proteins (24). Another described function of PDIs is the capture and transport of nitric oxide (25, 26), a possible intermediate or by-product of ammonia oxidation. The related protein family in *N. maritimus* may function in part to alleviate copper and nitric oxide toxicity.

Pathways for Ammonia Oxidation and Electron Transfer. The three genes (Nmar_1500, _1503, and _1502) annotated as *amoA*, *amoB*, and *amoC* and coding for a putative ammonia monooxygenase complex are the only recognizable genetic hallmarks of ammonia oxidation in the genome sequence. However, the *N. maritimus* sequences are no more similar (in either content or organizational structure) to bacterial *amo* genes than they are to the genes encoding bacterial particulate methane monooxygenases (pMMO), suggestive of functional differences between the archaeal and bacterial versions of AMO (7, 27). Notably, mapping the sequence encoded by *amoB* onto the pmoB crystal structure of *Methylococcus capsulatus* (Bath) (28) reveals the conservation of the ligands to the pMMO metal centers and the complete absence of both a transmembrane helix and a C-terminal cupredoxin domain predicted to be present in bacterial AMO (Fig. S2).

The structural differences in the archaeal AMO, the lack of genes encoding the hydroxylamine–ubiquinone redox module (29), and a periplasm enriched in redox active proteins together suggest significant divergence from the bacterial pathway of

ammonia oxidation. There are two hypothetical mechanistic alternatives (Fig. 1, Table S2): either a unique biochemistry exists for the oxidation of hydroxylamine or the divergent AMO does not actually produce hydroxylamine. If the former is true, hydroxylamine oxidation may occur via one of the periplasmic MCOs (CuHAO). Given the lack of cytochrome *c* proteins, the four electrons would then be transferred to a quinone reductase (QRED) via small blue copper-containing plastocyanin-like electron carriers. The protein encoded by Nmar_1226, which contains four transmembrane-spanning regions and two plastocyanin-like domains, may serve as an analog of the membrane-bound cytochrome c_M552 quinone reductase present in AOB (29) and is a good candidate for the QRED.

In an alternative scenario, the archaeal AMO produces not hydroxylamine, but the reactive intermediate nitroxyl (nitroxyl hydride, HNO). Nitroxyl is a highly toxic and reactive compound recently recognized as having biological significance in a number of systems (30, 31). During archaeal ammonia oxidation, nitroxyl might be formed by a unique monooxygenase function of archaeal AMO. Alternatively, the archaeal AMO may act as a dioxygenase and insert two oxygen atoms into ammonia, producing nitroxyl from the spontaneous decay of HNOHOH. Both reaction sequences eliminate the requirement for reductant recycling during the initial oxygenase reaction, a simplification offering significant ecological advantage (when compared with AOB) in nutrient poor environments. In this pathway, one of the MCO-like proteins may act as a nitroxyl oxidoreductase (NXOR) and facilitate the oxidation of nitroxyl to nitrite with the extraction of two protons and two electrons in the presence of water. The proposed NXOR would relay the two extracted electrons into the quinone pool via the QRED pathway described above.

In this proposed model, the electrons extracted by either a CuHAO or a NXOR (and transferred into the quinone pool) would generate a proton motive force (PMF) through complexes III (plastocyanin-like subunit, Nmar_1542; Rieske-type subunit, Nmar_1544; transmembrane subunit, Nmar_1543) and IV (Nmar_0182-5), driving the generation of ATP by an F_0F_1-type ATP synthase (Nmar_1688–1693). The production of reductant (i.e., NADH) would require the reverse operation of complex I (NuoABCDHIJKMLN, Nmar_0276–286) as a quinol oxidase driven by a PMF. The proposed biochemistry involving nitroxyl produces the same net gain as bacterial ammonia oxidation, providing two electrons for reduction of the quinone pool and subsequent linear electron flow and the generation of a PMF. The presence of a copper-containing (versus heme) complex III and the unique evolutionary placement of terminal oxidase (complex IV) between two of the heme–copper oxygen reductase families further distinguish this proposed ammonia oxidation pathway from that in AOB.

Carbon Fixation and Mixotrophy. *N. maritimus*, like all known AOB, grows chemolithoautotrophically by using inorganic carbon as the sole carbon source (7, 32). However, whereas AOB use the Calvin–Bassham–Benson cycle with the CO_2-fixing enzyme ribulose bisphosphate carboxylase/oxygenase (RubisCO) as the key enzyme, the absence of genes in *N. maritimus* coding for RubisCO and other enzymes of this cycle points to an alternative pathway for carbon fixation. The most likely mechanism supported by the genome sequence is the 3-hydroxypropionate/4-hydroxybutyrate pathway elucidated in the thermophilic crenarchaote *Metallosphaera sedula* and suggested as a potential pathway of carbon fixation in *C. symbiosum* (33). The pathway has two parts: a sequence including two carboxylation reactions transforming acetyl-CoA to succinyl-CoA and a multistep sequence converting succinyl-CoA into two molecules of acetyl-CoA. Genes identified in the *N. maritimus* genome coding for key enzymes of the pathway (Fig. S3) include a biotin-dependent acetyl-CoA/propionyl-CoA carboxylase (Nmar_0272–0274), methylmalonyl-CoA epimerase and mutase (Nmar_0953, _0954, and _0958), and 4-hydroxybutyrate dehydratase (Nmar_0207). With the exception of one gene (Nmar_1608), all of the genes implicated in the 3-hydroxyproprionate/4-hydrobutyrate pathway for carbon assimilation in *N. maritimus* are present and show highest similarity to the genes of *C. symbiosum* (34, 35). Although *N. maritimus* and *M. sedula* most likely use the same CO_2-fixation reaction sequences, not all individual reactions appear to be catalyzed by identical enzymes. In one instance, the stepwise reductive transformation of malonyl-CoA to propionyl-CoA involves five enzymes in *M. sedula* (33, 36, 37). Although the *N. maritimus* genome lacks any close homologs of the *M. sedula* genes, it contains alternative alcohol dehydrogenases, aldehyde dehydrogenases, acyl-CoA synthetases, and enoyl-CoA hydratases possibly fulfilling the same functions. Similarly, *M. sedula* catalyzes the activation of 3-hydroxypropionate to 3-hydroxypropionyl-CoA, using an AMP-forming 3-hydroxypropionyl-CoA synthetase (37). The *N. maritimus* genome lacks an obvious homolog, although does code for an ADP-forming acyl-CoA synthetase (Nmar_1309) that suggests a more energy efficient alternative.

In addition to the genes coding for the 3-hydroxypropionate/4-hydroxybutyrate pathway, the genome of *N. maritimus* contains a number of genes encoding enzymes of the tricarboxylic acid (TCA) cycle. No homologs for genes coding for a citrate-cleaving enzyme (ATP citrate lyase or citryl-CoA lyase) (38) were identified, permitting exclusion of the reductive TCA cycle as a pathway for carbon fixation. The lack of these genes suggests that *N. maritimus* utilizes either an incomplete (or horseshoe-type) TCA cycle for strictly biosynthetic purposes or possibly a complete oxidative TCA cycle.

N. maritimus grows on a completely inorganic medium, indicating the genes coding for essential biosynthetic capacity (*SI Materials and Methods*), yet its genomic inventory also suggests some flexibility in the utilization of organic sources of phosphorus and carbon. Two systems for phosphorous acquisition are suggested: the high-affinity, high-activity phosphate transport system (*pstSCAB*, Nmar_0479, Nmar_0481–0483) and a phosphonate transporter (Nmar_0873–0875). However, because the genome lacks genes encoding known C-P lyases and hydrolases (39), and phosphate limitation is not alleviated by supplementation with phosphonates common in the marine environment

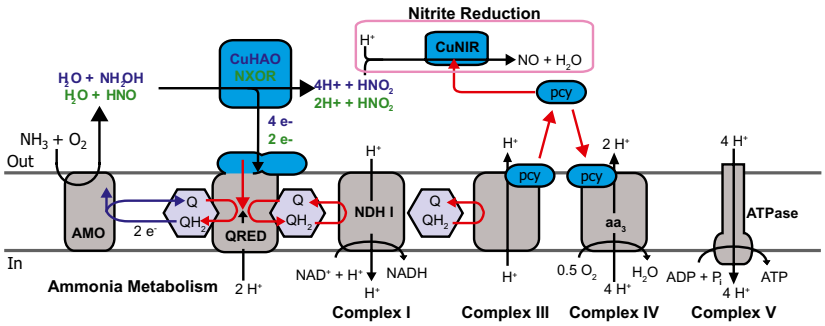

Fig. 1. Proposed AOA respiratory pathway. Text indicates the described possible hydroxylamine (blue text and arrows) and nitroxyl (green) pathways. Red arrows indicate electron flow not involved in ammonia oxidation. Blue shading denotes blue copper-containing proteins. Pink box indicates possible alternative respiratory electron sink. Hexagons containing Q and QH2 represent the oxidized and reduced quinone pools, respectively.

(e.g., aminoethylphosphonate), there is as yet no support for a functional phosphonate utilization pathway. Numerous organic transport functions are also evident, broadly encompassing transporters for different amino acids, dipeptides/oligopeptides, sulfonates/taurine, and glycerol. Additional physiological characterization will likely demonstrate some capacity for mixotrophic growth, as suggested by isotopic studies of natural populations (40–42). No genomic evidence exists for growth on urea, as *N. maritimus* lacks the homologs of the putative urease and urea transporter genes identified in *C. symbiosum* (35).

Noncoding RNA Genes. The *N. maritimus* genome contains a full complement of essential noncoding RNA (ncRNA) genes, including one copy each of 5S/16S/23S ribosomal RNAs, RNase P, SRP RNA, and 44 transfer RNAs (Table S3). In addition to six normally placed canonical tRNA introns, noncanonical introns were found at different positions in six of the tRNAs (Val_{CAC}, Met, Trp, Arg_{CCT}, Leu_{TAA}, and Glu_{TTC}), a phenomenon previously observed only in thermophiles and *C. symbiosum* (34, 43). All other sequenced crenarchaea (including *C. symbiosum*) contain at least 46 tRNAs. *N. maritimus* lacks tRNA sequences coding for Pro_{CGG} or Arg_{CCG}, perhaps resulting from (or related to) the low G + C content of the genome and preference for protein codons ending in A/T. Other archaeal genomes with relatively low G + C content, such as the euryarchaeon *Haloquadratum walsbyi*, also lack these tRNAs, while possessing the exact complement of 44 tRNAs found in *N. maritimus* (43). This occurrence likely reflects a difference in posttranscriptional modification of the wobble base of tRNAs Pro_{TGG} and Arg_{TCG}, allowing more efficient decoding of the rare codons CCG and CGG, respectively (44).

Six candidates for C/D box small RNAs (sRNAs) were identified (Nmar_sR1–sR6). Most C/D box sRNAs guide the precise positioning of posttranscriptional 2′-O-methyl group addition to rRNAs or tRNAs, a process also occurring in eukaryotes, but not *Bacteria*. In *N. maritimus*, predicted targets of 2′-O-methylation may include the wobble base of the Leu_{CAA} tRNA and two different positions separated by 26 nucleotides in the 23S rRNA. Before the *N. maritimus* and *C. symbiosum* genomes, multiple C/D box sRNAs were found almost exclusively in hyperthermophilic archaea (45). Detection of these conserved, syntenic guide sRNAs in the two mesophilic crenarchaeal genomes supports an RNA stabilization/chaperone function not seen in other archaeal mesophiles and possibly more similar to their predicted function(s) in eukaryotes (46).

Regulation of Transcription. The genome contains at least eight transcription factor B (TFB) and two TATA-box binding protein (TBP) (Table S3) genes required for starting site-specific transcription initiation, making *N. maritimus* among the densest and richest archaeal genomes for these transcription factors. TFB and TBP are thought to serve functions similar to the bacterial sigma factors (e.g., modulating cellular function in response to fluctuating environmental conditions) in *Archaea* with genomes coding for multiple copies, with optimal TFB/TPB partners (47). Although many other archaeal genomes contain multiple copies of these transcription factors, only the haloarchaea have more than five TFB genes (47). The functional significance of this exceptionally high density of regulatory factors in an apparently metabolically specialized organism will likely be informed by future transcriptional analyses of different growth states. Genes for two widely distributed types of archaeal chromatin proteins are present, an archaeal histone (Nmar_0683) and two Alba genes (Nmar_0255 and Nmar_0933). These are thought to maintain chromosomal material in a state permitting polymerase accessibility, with differential expression possibly providing for altered global transcription (48).

Unique Cell Division Machinery and Previously Uncharacterized Instance of Archaeal Biosynthesis of Hydroxyectoine. The *N. maritimus* genome contains elements homologous with two systems of cell division: *ftsZ* (Nmar_1262) and *cdvABC* (Nmar_0700, _0816, and _1088). The *cdvABC* operon, induced at the onset of genome segregation and cell division, codes for machinery related to the eukaryotic endosomal sorting complex (49, 50). With the exception of *N. maritimus*, *C. symbiosum*, and the Thermoproteales (where the cell division machinery remains uncharacterized), all available archaeal genomes have either the FtsZ or the Cdv cell division machinery, but not both. The two cell division systems may comprise a hybrid mechanism or serve two distinct processes in marine *Crenarchaeota*.

The genome of *N. maritimus* also encodes for synthesis of the compatible solute hydroxyectoine: ectoine synthase (*ectC*, Nmar_1344), a L-2,4-diaminobutyrate transaminase (*ectB*, Nmar_1345), a L-2,4-diaminobutyrate acetyltransferase (*ectA*, Nmar_1346), and an ectoine hydroxylase (*ectD*, Nmar_1343). Although widely distributed among *Bacteria* (in particular the genome sequences of marine organisms), the presence of these genes in *N. maritimus* represents a unique indication of archaeal biosynthesis.

Relationship to *C. symbiosum*. The genome of *N. maritimus* differs significantly in G + C content and size from that of the closely related sponge symbiont (~97% 16S rRNA gene sequence identity). Despite the differences in overall genomic G + C content (34.2% for *N. maritimus* versus 57.7% for *C. symbiosum*), the G + C content

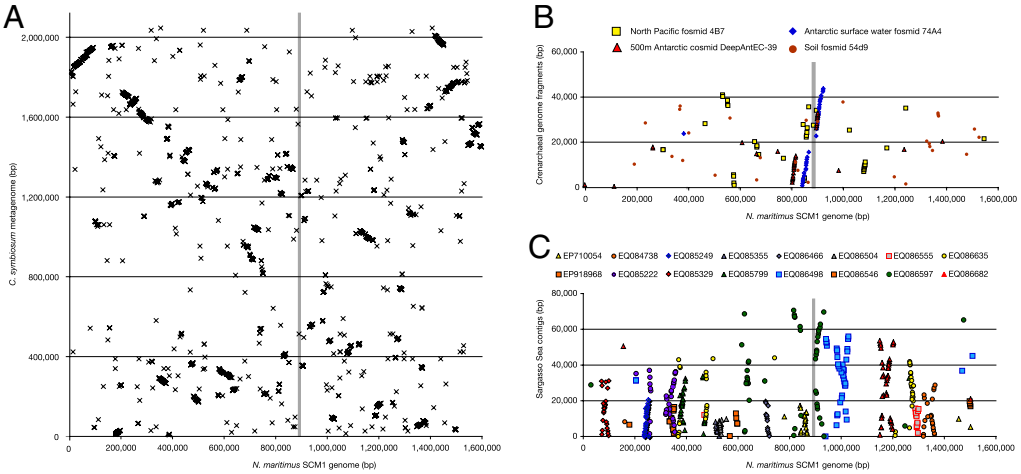

Fig. 2. Synteny plots comparing the *N. maritimus* genome with (*A*) the *Cenarchaeum symbiosum* A type genome, (*B*) crenarchaeal genome fragments, and (*C*) Sargasso Sea fosmids. Vertical gray bar indicates location of ribosomal RNA operon.

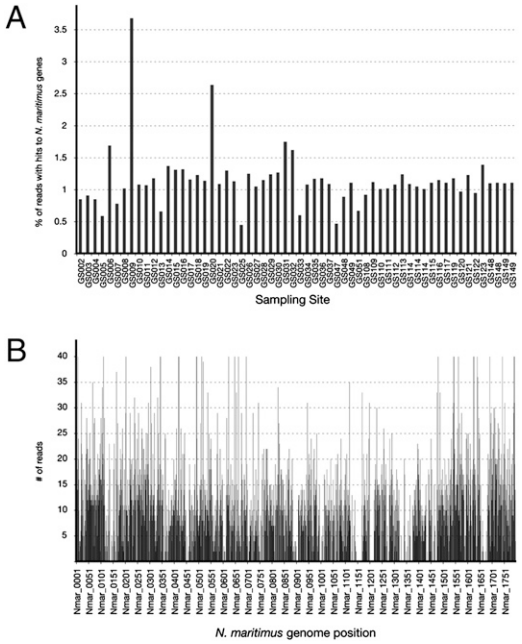

Fig. 3. Comparison of *N. maritimus* genome to metagenomic sequence reads from GOS. (*A*) Percentage of reads from each GOS site that align to the *N. maritimus* genome by protein sequence similarity. (*B*) Number of GOS reads homologous to each *N. maritimus* protein-coding gene. Counts of 40 are mostly due to highly conserved proteins that include contaminants from other clades.

for the rRNAs is largely similar between both organisms (50–52%). The higher ORF density (1.19 ORF/kb) relative to *C. symbiosum* (0.986 ORF/kb) results principally from the 0.4 Mbp smaller genome of *N. maritimus*, not from a large disparity in the number of predicted ORFs. The two genomes share 1,267 genes in common (when compared via reciprocal BLAST with expectation cutoff values of 10^{-4}), yet there is little conservation of synteny (Fig. 2*C*, Table S4). Most of the increased size of the *C. symbiosum* genome and much of the divergence in gene content are associated with discrete regions ("islands"), although no obvious functionality could be assigned to individual islands. Homologs for a majority of genes implicated in the archaeal ammonia oxidation pathway (51 of 69 listed in Table S2) appear present in *C. symbiosum*.

Phylogeny and Evolution. The widely distributed Group I archaeal lineage with which *N. maritimus* is affiliated was earlier assigned to the hyperthermophilic *Crenarchaeota* (3). Questions regarding this association arose through phylogenetic analysis of *C. symbiosum* ribosomal proteins, indicating possible divergence before the *Crenarchaeota–Euryarchaeota* split and therefore deserving provisional assignment to a new archaeal kingdom, the Thaumarchaeota (51). The basal position of the Group I archaea previously inferred from protein sequences encoded by the *C. symbiosum* genome was reexamined by reanalysis of the combined dataset, using patterns of gene distribution (Table S5) and phylogeny inference. Maximum-likelihood analyses confirmed the basal branching with significant statistical support (bootstrap value = 90%, Fig. S4). Bayesian analysis of a selection of species from the same dataset produced results linking *C. symbiosum* and *N. maritimus* as sister taxa of *Crenarchaeota*, albeit with nonsignificant support (posterior probability = 0.88). Although a definitive placement within the Archaea still must be confirmed by inclusion of genomes from more distantly related lineages, both analyses strongly support a lineage distinct from all other cultivated *Crenarchaeota*.

High Similarity to the Metagenome of the Globally Distributed Marine Group I Archaea. The genome of *N. maritimus* shares remarkable conservation of gene content and gene order with numerous archaeal sequences previously recovered in fosmid libraries and recent oceanic surveys. The Antarctic genome fragments DeepAnt-EC39 (taken from 500 m depth) (52) and cosmid 74A4 (from surface waters) (53) both share very high synteny with portions of the *N. maritimus* genome (Fig. 2*B* and Table S6) despite significant differences in rRNA sequence identity (93 and 98%, respectively). Sixteen Sargasso Sea contigs (93% 16S rRNA gene sequence identity with *N. maritimus*) also have high synteny (Fig. 2*C* and Table S6). Retrieval of DNA fragments from the Global Ocean Sampling (GOS) database using differential protein sequence similarity showed that *N. maritimus*-like sequences constituted an average of 1.15% of all sequences across a wide range of temperatures (9–29 °C), salinities (freshwater to seawater, 0.1–63 practical salinity units), two open oceans, and several coastal environments (Fig. 3*A*). The Block Island, NY, coastal site and the Lake Gatun, Panama Canal, site (neither of which share any notable physical/chemical characteristics other than sample depth) both exhibited notably large increases in density (>2.5%). The available GOS sequences provided almost complete and uniform coverage of the *N. maritimus* genome (Fig. 3*B*), although at least three significant gaps in coverage exist (possibly corresponding with unique *N. maritimus* genomic islands). Whereas some of the coverage may result in matches to bacterial sequences, particularly for very highly conserved proteins, the majority of recruited proteins had >50% sequence identity to *N. maritimus* proteins. Together, these shared genomic features suggest *N. maritimus* is representative of many of the globally abundant marine Group I Crenarchaeaota and that it should provide a useful model for developing an understanding of the basic physiology of these abundant and cosmopolitan organisms.

A comparison of the sequence content and genome organization hints at functionally more divergent marine population types. *N. maritimus* has limited syntenic similarity to a deep-water population represented by the North Pacific fosmid 4B7 (93% 16S rRNA sequence identity), but shares proteins highly similar to most of those coded on this fosmid. Previous comparison of marine crenarchaeal genomic fragments reported changes in genomic organization with sampling depths, suggestive of depth-related habitat types (54). Coupled with recent evidence indicating varied physiological lifestyles along depth and latitudinal gradients, distinct crenarchaeal ecotypes likely exist, analogous to that observed in marine cyanobacteria (2, 55). However, no clear correlations currently exist between environmental parameters and the similarity of recruited fragments.

The genome sequence presented here also offers further insight into the ecological success of AOA. For example, using the likely more energy-efficient 3-hydroxypropionate/4-hydroxybutyrate pathway for CO_2 fixation rather than the Calvin–Bassham–Benson cycle used by AOB could provide a growth advantage. Further ecological advantage may be conferred by their potential capacity for mixotrophic growth or the use of copper as a major redox-active metal for respiration in generally iron-limited oceans. However, a deeper understanding of the remarkable success of this archaeal lineage will come only from more detailed physiological, biochemical, and genetic characterization of *N. maritimus* and additional environmental isolates.

Materials and Methods

Genome sequencing was performed on high-molecular-weight DNA extracted from two cultures of *N. maritimus*. Whole-genome shotgun sequencing of 3-, 8-, and 40-kb DNA libraries by the Joint Genome Institute produced at least 8× coverage. Annotation of the closed genome was performed using Department of Energy (DOE) computational support at Oak Ridge National Laboratory and The Institute for Genomic Research (TIGR) Autoannotation Service in conjunction with Manatee visualization software. Complete details describing high-molecular-weight DNA purification and sequence analysis are found in *SI Materials and Methods*.

ACKNOWLEDGMENTS. The authors thank David Bruce and Paul Richardson from the Joint Genome Institute for facilitating genome sequencing. This

work was supported by the Department of Energy Microbial Genome Program, by National Science Foundation Microbial Interactions and Processes Grant MCB-0604448 (to D.A.S. and J.R.d.l.T.), by National Science Foundation Molecular and Cellular Biosciences Grant MCB-0920741 (to D.A.S.), by National Science Foundation Biological Oceanography Grants OCE-0623174 (to D.A.S.) and OCE-0623908 (to S.M.S.), by National Science Foundation Grant EF-0412129 (to M.G.K.), by incentive funds from the University of Louisville VP Research office (to M.G.K.), by the Deutsche Forschungsgemeinschaft (M.K.), by US Department of Agriculture Grant 2010-65115-20380 (to A.C.R.), and by a Salk Institute Innovation Grant (to G.M.).

1. Karner MB, DeLong EF, Karl DM (2001) Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* 409:507–510.
2. Agogué H, Brink M, Dinasquet J, Herndl GJ (2008) Major gradients in putatively nitrifying and non-nitrifying *Archaea* in the deep North Atlantic. *Nature* 456:788–791.
3. DeLong EF (1992) *Archaea* in coastal marine environments. *Proc Natl Acad Sci USA* 89: 5685–5689.
4. Fuhrman JA, McCallum K, Davis AA (1992) Novel major archaebacterial group from marine plankton. *Nature* 356:148–149.
5. Venter JC, et al. (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 304:66–74.
6. Wuchter C, et al. (2006) Archaeal nitrification in the ocean. *Proc Natl Acad Sci USA* 103:12317–12322.
7. Könneke M, et al. (2005) Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* 437:543–546.
8. Leininger S, et al. (2006) *Archaea* predominate among ammonia-oxidizing prokaryotes in soils. *Nature* 442:806–809.
9. Prosser JI, Nicol GW (2008) Relative contributions of archaea and bacteria to aerobic ammonia oxidation in the environment. *Environ Microbiol* 10:2931–2941.
10. Prosser JI (1989) Autotrophic nitrification in bacteria. *Adv Microb Physiol* 30:125–181.
11. de la Torre JR, Walker CB, Ingalls AE, Könneke M, Stahl DA (2008) Cultivation of a thermophilic ammonia oxidizing archaeon synthesizing crenarchaeol. *Environ Microbiol* 10:810–818.
12. Hatzenpichler R, et al. (2008) A moderately thermophilic ammonia-oxidizing crenarchaeote from a hot spring. *Proc Natl Acad Sci USA* 105:2134–2139.
13. Martens-Habbena W, Berube PM, Urakawa H, de la Torre JR, Stahl DA (2009) Ammonia oxidation kinetics determine niche separation of nitrifying Archaea and Bacteria. *Nature* 461:976–979.
14. Sernova NV, Gelfand MS (2008) Identification of replication origins in prokaryotic genomes. *Brief Bioinform* 9:376–391.
15. Makarova KS, Sorokin AV, Novichkov PS, Wolf YI, Koonin EV (2007) Clusters of orthologous genes for 41 archaeal genomes and implications for evolutionary genomics of archaea. *Biol Direct* 2:33.
16. Schäfer G (2004) *Respiration in Archaea and Bacteria*, ed Zannoni D (Springer, Dordrecht, The Netherlands), pp 1–33.
17. Beaumont HJ, Lens SI, Westerhoff HV, van Spanning RJ (2005) Novel *nirK* cluster genes in *Nitrosomonas europaea* are required for NirK-dependent tolerance to nitrite. *J Bacteriol* 187:6849–6851.
18. Andersen CL, Matthey-Dupraz A, Missiakas D, Raina S (1997) A new *Escherichia coli* gene, *dsbG*, encodes a periplasmic protein involved in disulphide bond formation, required for recycling DsbA/DsbB and DsbC redox proteins. *Mol Microbiol* 26:121–132.
19. Meima R, et al. (2002) The *bdbDC* operon of Bacillus subtilis encodes thiol-disulfide oxidoreductases required for competence development. *J Biol Chem* 277:6994–7001.
20. Raina S, Missiakas D (1997) Making and breaking disulfide bonds. *Annu Rev Microbiol* 51:179–202.
21. Hiniker A, Collet JF, Bardwell JC (2005) Copper stress causes an in vivo requirement for the *Escherichia coli* disulfide isomerase DsbC. *J Biol Chem* 280:33785–33791.
22. Pogliano J, Lynch AS, Belin D, Lin EC, Beckwith J (1997) Regulation of *Escherichia coli* cell envelope proteins involved in protein folding and degradation by the Cpx two-component system. *Genes Dev* 11:1169–1182.
23. Kershaw CJ, Brown NL, Constantinidou C, Patel MD, Hobman JL (2005) The expression profile of *Escherichia coli* K-12 in response to minimal, optimal and excess copper concentrations. *Microbiology* 151:1187–1198.
24. Narindrasorasak S, Yao P, Sarkar B (2003) Protein disulfide isomerase, a multifunctional protein chaperone, shows copper-binding activity. *Biochem Biophys Res Commun* 311: 405–414.
25. Sliskovic I, Raturi A, Mutus B (2005) Characterization of the S-denitrosation activity of protein disulfide isomerase. *J Biol Chem* 280:8733–8741.
26. Ramachandran N, Root P, Jiang XM, Hogg PJ, Mutus B (2001) Mechanism of transfer of NO from extracellular S-nitrosothiols into the cytosol by cell-surface protein disulfide isomerase. *Proc Natl Acad Sci USA* 98:9539–9544.
27. Nicol GW, Tscherko D, Chang L, Hammesfahr U, Prosser JI (2006) Crenarchaeal community assembly and microdiversity in developing soils at two sites associated with deglaciation. *Environ Microbiol* 8:1382–1393.
28. Lieberman RL, Rosenzweig AC (2005) Crystal structure of a membrane-bound metalloenzyme that catalyses the biological oxidation of methane. *Nature* 434: 177–182.
29. Klotz MG, Stein LY (2008) Nitrifier genomics and evolution of the nitrogen cycle. *FEMS Microbiol Lett* 278:146–156.
30. Fukuto JM, Switzer CH, Miranda KM, Wink DA (2005) Nitroxyl (HNO): Chemistry, biochemistry, and pharmacology. *Annu Rev Pharmacol Toxicol* 45:335–355.
31. Miranda KM, et al. (2003) A biochemical rationale for the discrete behavior of nitroxyl and nitric oxide in the cardiovascular system. *Proc Natl Acad Sci USA* 100:9196–9201.
32. Arp DJ, Chain PS, Klotz MG (2007) The impact of genome analyses on our understanding of ammonia-oxidizing bacteria. *Annu Rev Microbiol* 61:503–528.
33. Berg IA, Kockelkorn D, Buckel W, Fuchs G (2007) A 3-hydroxypropionate/4-hydroxybutyrate autotrophic carbon dioxide assimilation pathway in Archaea. *Science* 318:1782–1786.
34. Hallam SJ, et al. (2006) Genomic analysis of the uncultivated marine crenarchaeote *Cenarchaeum symbiosum*. *Proc Natl Acad Sci USA* 103:18296–18301.
35. Hallam SJ, et al. (2006) Pathways of carbon assimilation and ammonia oxidation suggested by environmental genomic analyses of marine *Crenarchaeota*. *PLoS Biol* 4: e95.
36. Alber B, et al. (2006) Malonyl-coenzyme A reductase in the modified 3-hydroxypropionate cycle for autotrophic carbon fixation in archaeal *Metallosphaera* and *Sulfolobus* spp. *J Bacteriol* 188:8551–8559.
37. Alber BE, Kung JW, Fuchs G (2008) 3-Hydroxypropionyl-coenzyme A synthetase from *Metallosphaera sedula*, an enzyme involved in autotrophic CO2 fixation. *J Bacteriol* 190:1383–1389.
38. Hügler M, Huber H, Molyneaux SJ, Vetriani C, Sievert SM (2007) Autotrophic CO2 fixation via the reductive tricarboxylic acid cycle in different lineages within the phylum *Aquificae*: Evidence for two ways of citrate cleavage. *Environ Microbiol* 9: 81–92.
39. Quinn JP, Kulakova AN, Cooley NA, McGrath JW (2007) New ways to break an old bond: The bacterial carbon-phosphorus hydrolases and their role in biogeochemical phosphorus cycling. *Environ Microbiol* 9:2392–2400.
40. Herndl GJ, et al. (2005) Contribution of *Archaea* to total prokaryotic production in the deep Atlantic Ocean. *Appl Environ Microbiol* 71:2303–2309.
41. Ingalls AE, et al. (2006) Quantifying archaeal community autotrophy in the mesopelagic ocean using natural radiocarbon. *Proc Natl Acad Sci USA* 103:6442–6447.
42. Ouverney CC, Fuhrman JA (2000) Marine planktonic archaea take up amino acids. *Appl Environ Microbiol* 66:4829–4833.
43. Chan PP, Lowe TM (2009) GtRNAdb: A database of transfer RNA genes detected in genomic sequence. *Nucleic Acids Res* 37 (Database issue):D93–D97.
44. Grosjean H, Marck C, de Crécy-Lagard V (2007) The various strategies of codon decoding in organisms of the three domains of life: Evolutionary implications. *Nucleic Acids Symp Ser (Oxf)* 51:15–16.
45. Dennis PP, Omer A, Lowe T (2001) A guided tour: Small RNA function in *Archaea*. *Mol Microbiol* 40:509–519.
46. Maxwell ES, Fournier MJ (1995) The small nucleolar RNAs. *Annu Rev Biochem* 64: 897–934.
47. Facciotti MT, et al. (2007) General transcription factor specified global gene regulation in *archaea*. *Proc Natl Acad Sci USA* 104:4630–4635.
48. Sandman K, Reeve JN (2005) Archaeal chromatin proteins: Different structures but common function? *Curr Opin Microbiol* 8:656–661.
49. Lindås AC, Karlsson EA, Lindgren MT, Ettema TJ, Bernander R (2008) A unique cell division machinery in the *Archaea*. *Proc Natl Acad Sci USA* 105:18942–18946.
50. Samson RY, Obita T, Freund SM, Williams RL, Bell SD (2008) A role for the ESCRT system in cell division in *archaea*. *Science* 322:1710–1713.
51. Brochier-Armanet C, Boussau B, Gribaldo S, Forterre P (2008) Mesophilic *Crenarchaeota*: Proposal for a third archaeal phylum, the *Thaumarchaeota*. *Nat Rev Microbiol* 6:245–252.
52. Stein JL, Marsh TL, Wu KY, Shizuya H, DeLong EF (1996) Characterization of uncultivated prokaryotes: Isolation and analysis of a 40-kilobase-pair genome fragment from a planktonic marine archaeon. *J Bacteriol* 178:591–599.
53. Béjà O, et al. (2002) Comparative genomic analysis of archaeal genotypic variants in a single population and in two different oceanic provinces. *Appl Environ Microbiol* 68:335–345.
54. López-García P, Brochier C, Moreira D, Rodríguez-Valera F (2004) Comparative analysis of a genome fragment of an uncultivated mesopelagic crenarchaeote reveals multiple horizontal gene transfers. *Environ Microbiol* 6:19–34.
55. Johnson ZI, et al. (2006) Niche partitioning among Prochlorococcus ecotypes along ocean-scale environmental gradients. *Science* 311:1737–1740.