Proteomics and comparative genomics of *Nitrososphaera viennensis* reveal the core genome and adaptations of archaeal ammonia oxidizers

Melina Kerou[a], Pierre Offre[a,1], Luis Valledor[a,2], Sophie S. Abby[a], Michael Melcher[a], Matthias Nagler[a], Wolfram Weckwerth[b,3], and Christa Schleper[a,3]

[a]Department of Ecogenomics and Systems Biology, University of Vienna, A-1090 Vienna, Austria; and [b]Vienna Metabolomics Center, University of Vienna, A-1090 Vienna, Austria

Edited by Edward F. DeLong, University of Hawaii at Manoa, Honolulu, HI, and approved October 11, 2016 (received for review January 22, 2016)

Ammonia-oxidizing archaea (AOA) are among the most abundant microorganisms and key players in the global nitrogen and carbon cycles. They share a common energy metabolism but represent a heterogeneous group with respect to their environmental distribution and adaptions, growth requirements, and genome contents. We report here the genome and proteome of *Nitrososphaera viennensis* EN76, the type species of the archaeal class Nitrososphaeria of the phylum Thaumarchaeota encompassing all known AOA. *N. viennensis* is a soil organism with a 2.52-Mb genome and 3,123 predicted protein-coding genes. Proteomic analysis revealed that nearly 50% of the predicted genes were translated under standard laboratory growth conditions. Comparison with genomes of closely related species of the predominantly terrestrial Nitrososphaerales as well as the more streamlined marine Nitrosopumilales [*Candidatus* (*Ca.*) order] and the acidophile "*Ca.* Nitrosotalea devanaterra" revealed a core genome of AOA comprising 860 genes, which allowed for the reconstruction of central metabolic pathways common to all known AOA and expressed in the *N. viennensis* and "*Ca.* Nitrosopelagicus brevis" proteomes. Concomitantly, we were able to identify candidate proteins for as yet unidentified crucial steps in central metabolisms. In addition to unraveling aspects of core AOA metabolism, we identified specific metabolic innovations associated with the Nitrososphaerales mediating growth and survival in the soil milieu, including the capacity for biofilm formation, cell surface modifications and cell adhesion, and carbohydrate conversions as well as detoxification of aromatic compounds and drugs.

ammonia oxidation | proteomics | archaea | comparative genomics | biofilm

Ten years of extensive research on ammonia-oxidizing archaea (AOA) have provided important insights into their environmental distribution, ecophysiology, and general genome content. AOA belong to the archaeal phylum Thaumarchaeota (1, 2) and are a major component of many microbial communities, with abundances up to 10^7 cells per 1 mL in sea waters (3, 4) and even 100-fold more per gram of soil (5, 6). They perform a key role in global nitrogen cycling but also participate in carbon cycling and production of metabolites essential to many members of the microbial and eukaryotic community (7–9). Since their discovery in 2005 (10, 11), about a dozen strains have been enriched under aerobic and autotrophic conditions, and some were obtained in pure laboratory cultures. About an equal number of full or nearly complete genome sequences are now available, allowing the first comparative analyses of their core and variable gene repertoires (12, 13). Despite their vast distribution and large numbers in the environment, AOA are delicate organisms with which to work. Their small cell sizes, slow growth rates, and autotrophic (or mixotrophic) growth mode on low concentrations of ammonia as energy source (compared with their bacterial counterparts) have limited the ecophysiological and metabolic characterization of these oligotrophic organisms (12, 14, 15). *Nitrososphaera viennensis*, isolated from soil, represents the first deposited type strain of the class Nitrososphaeria within the phylum Thaumarchaeota that encompasses all known AOA (16, 17) and is one of the best studied AOA from soil. The growth characteristics of *N. viennensis* on ammonia and urea as energy sources have been described (16), as well as its dependence on small organic acids for growth (16), resistance to common nitrification inhibitors (18, 19), production of NO and contribution to greenhouse gas emissions (20, 21), thaumarchaeal ether–lipid components (22), and typical archaeal S-layer structure with p3 symmetry (17).

Despite this accumulated knowledge on *N. viennensis* and other AOA, many fundamental questions about their core and variable metabolism and niche differentiation remain unanswered. Although it is generally assumed that the archaeal version of ammonia monooxygenase encoded in all studied AOA is responsible for the first step of ammonia oxidation, the second step to nitrite production and the contributing cofactors and electron carriers have still not been identified. Moreover, although a unique and highly energy efficient aerobic pathway of carbon assimilation conserved in all AOA has been biochemically investigated in the marine strain "*Candidatus* (*Ca.*) Nitrosopumilus maritimus" (23), regulation

Significance

Ammonia-oxidizing archaea (AOA), key players in global biogeochemical cycles, represent a heterogeneous group with a broad environmental distribution. Understanding their activity and physiology is of great importance due to the impact of the overuse of agricultural fertilizers on the N cycle and the production of the greenhouse gas N_2O during nitrification. Despite their prominent ecological role, little is known about the fundamental metabolic processes of AOA. Here, we show that AOA of marine and terrestrial environments share unique and well-conserved pathways of carbon and nitrogen metabolism, and we raise hypotheses about missing steps in these pathways. Our approach also highlights the extensive environmental adaptations of the soil clade, including the capacity for cell surface modifications, carbohydrate conversions, detoxification, and biofilm formation.

Author contributions: M.K. and C.S. designed research; M.K., P.O., L.V., S.S.A., M.M., and M.N. performed research; M.K., P.O., L.V., S.S.A., M.N., W.W., and C.S. analyzed data; and M.K. and C.S. wrote the paper, with contributions from P.O. and W.W.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession no. CP007536), and the mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium, www.proteomexchange.org (identifier PXD005297).

[1]Present address: HGF MPG Joint Research Group for Deep-Sea Ecology and Technology, Max Planck Institute for Marine Microbiology, 28359 Bremen, Germany.

[2]Present address: Department of Organisms and Systems Biology, University of Oviedo, Catedratico Rodrigo Uria E-33006, Spain.

[3]To whom correspondence may be addressed. Email: wolfram.weckwerth@univie.ac.at or Christa.Schleper@univie.ac.at.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1601212113/-/DCSupplemental.

mechanisms that maintain the balance over the key metabolic processes of nitrogen and carbon assimilation have not been elucidated. In addition, adaptive features of different AOA to their respective ecological niches have not been investigated, which are needed to understand their metabolic flexibility and their reaction to changing environmental conditions. A better understanding of the ecosphysiology of terrestrial AOA compared with their ammonia-oxidizing bacteria (AOB) counterparts that coexist in the same habitats would help to elucidate the relative contribution of both AOA and AOB to nitrogen (and carbon) turnover in soils. This knowledge of AOA (eco)physiology is not only of fundamental ecological interest but also might eventually be of importance for mitigation strategies aimed at reducing N_2O emissions in agriculture, since AOB but not AOA can produce significant amounts of N_2O as a side reaction of ammonia oxidation or via the nitrifier denitrification pathway (21, 24, 25).

To gain deeper insights into the central metabolism of AOA and specific adaptations of the terrestrial group, we have performed a manual annotation of the genome of *N. viennensis*, the type strain of the phylum Thaumarchaeota, and determined its proteome under standard growth conditions. In a comparative genomic analysis with seven closed genomes of Thaumarchaeota (26–31), we have dissected the conserved gene repertoire of AOA and compared this with the expressed proteome of *N. viennensis* and "*Ca.* Nitrosopelagicus brevis" (30). This comparative analysis allows us to narrow down the minimal active gene set of AOA, reconstruct central metabolic pathways, including possible regulatory mechanisms, and identify candidates for potential missing steps in the core energy metabolism of AOA. In addition, we identify and discuss in the context of their ecophysiology the gene complement specific to *N. viennensis* and the terrestrial strains.

Results and Discussion

With 2.52 Mb and 3,123 predicted protein-coding genes, the *N. viennensis* genome is similar in size to the other genomes of the Nitrososphaerales [formerly group I.1b (17, 28, 29)] and larger than genomes of the other major order "*Ca.* Nitrosopumilales" (formerly group I.1a) that are mostly marine strains and range in size from 1.23 to 1.85 Mb (26, 27, 30, 32–38) [except marine symbiont "*Ca.* Cenarchaeum symbiosum", with 2.05 Mb (39)].

The proteome of late exponential-phase cultures of *N. viennensis* was analyzed on an LC-Orbitrap Mass Spectrometer. Total cellular lysates were fractionated in membrane-enriched and cytoplasmic fractions by ultracentrifugation to enrich and enable identification of otherwise poorly represented membrane-associated proteins. From the combined datasets, we recovered a total of 1,503 proteins, representing 48% of all predicted proteins encoded in the *N. viennensis* genome (Dataset S1). The absolute number of proteins identified in the proteome of *N. viennensis* was considerably higher than 1,012 proteins reported recently for (early stationary phase) cultures of the marine AOA *Ca.* N. brevis (30). The expressed proteome of *Ca*. N. brevis, however, represented a greater fraction (70%) of all proteins predicted from its genome (1,445 coding sequences) as expected from its more streamlined genome that results in a smaller size of its nonconserved (shell) gene complement. Assuming that the larger gene repertoire of *N. viennensis* is a consequence of its adaptations to a diverse range of environmental changes and stresses in soil, we do not expect expression of a higher protein complement of the genome under stable laboratory culture conditions.

The Core Gene Repertoire of AOA. Clusters of orthologous groups (COGs) of proteins were defined for seven closed genomes of free-living Thaumarchaeota representing the order Nitrososphaerales and the two candidate orders *Ca.* Nitrosopumilales and "*Ca.* Nitrosotaleales" (formerly clades I.1b, I.1a, and I.1a-associated, respectively) (features of the genomes are summarized in Table S1). This dataset included, in addition to *N. viennensis*, "*Ca.* Nitrososphaera gargensis" (28) and "*Ca.* Nitrososphaera evergladensis" (29) of the order Nitrososphaerales; *Ca.* N. maritimus (26), "*Ca.* Nitrosoarchaeum koreensis" (27), and *Ca.* N. brevis, a marine strain for which a proteome dataset is also available (30), from the order *Ca.* Nitrosopumilales; and "*Ca.* Nitrosotalea devanaterra", an acidophilic soil strain that, according to its divergence from other strains, probably represents a separate order, *Ca.* Nitrosotaleales, associated with *Ca.* Nitrosopumilales (31) (the *Candidatus* designation of the strains was omitted below to improve readability). This comparison enabled us to identify a conserved core set of 860 protein families (Fig. 1 and Dataset S2) (i.e., families shared by all analyzed genomes and likely representing the "core" gene repertoire of AOA; from here on, they are referred to as core COG families). Among these core COG families, we encountered all of the highly conserved information processing genes (involved in replication, transcription, translation, etc.) that are signatures of the phylum Thaumarchaeota (as outlined in ref. 2). A summary of the functional categories of the core COG families is depicted in Fig. S1.

Members of the core COG families represent a large fraction of the expressed proteome in both available proteomic datasets, composing 48.8% of the identified proteins in *N. viennensis* (733 in a total of 1,503) (Fig. 1) and 67.3% in *N. brevis* (681 in 1,012 identified proteins), whereas 611 were expressed in both (Dataset S1). Both proteomic datasets were generated from cultures grown in defined media [as defined for each organism (17, 30)] under ammonia-oxidizing conditions. Comparative analysis of the genomic and proteomic datasets highlights and confirms the common processes supporting the aerobic, autotrophic ammonia-oxidizing lifestyle in these representatives, although they stem from two distinct orders associated with radically different ecological habitats (12, 13, 40). This conservation among the core protein families enabled us to raise hypotheses about missing steps in shared metabolisms and dissect highly specialized distinct functions in the variable (accessory) genomes (see below).

Conserved Central Carbon Metabolism in AOA. All pathways of the central carbon metabolism including the genes (canonical and putative) participating in the hydroxypropionate/hydroxybutyrate CO_2 fixation pathway (23, 41), tricarboxylic acid (TCA) cycle, gluconeogenesis, and nonoxidative pentose phosphate pathways

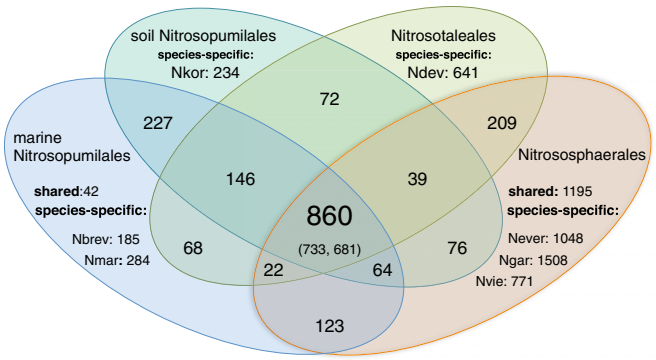

Fig. 1. Simplified Venn diagram illustrating the numbers of COGs shared between marine *Ca.* Nitrosopumilales (COGs shared between *N. maritimus* and *N. brevis* and between either of them and the other categories), soil *Ca.* Nitsosopumilales (COGs shared between *N. koreensis* and the other categories), *Ca.* Nitrosotaleales (COGs shared between *N. devanaterra* and the other categories), and Nitrososphaerales [COGs shared between *N. viennensis*, *N. gargensis*, and *N. evergladensis* (either among all three or pairwise) and between either of them and the other categories]. Proteins that were not grouped into COGs are represented as specific proteins for each organism. Not all combinations of shared COGs are represented in this Venn diagram. Numbers in parentheses indicate the numbers of clusters detected in the experimentally determined proteomes of *N. viennensis* and *N. brevis*, respectively. Genome abbreviations are Nbrev, *N. brevis*; Ndev, *N. devanaterra*; Never, *N. evergladensis*; Ngar, *N. gargensis*; Nkor, *N. koreensis*; Nmar, *N. maritimus*; and Nvie, *N. viennensis*.

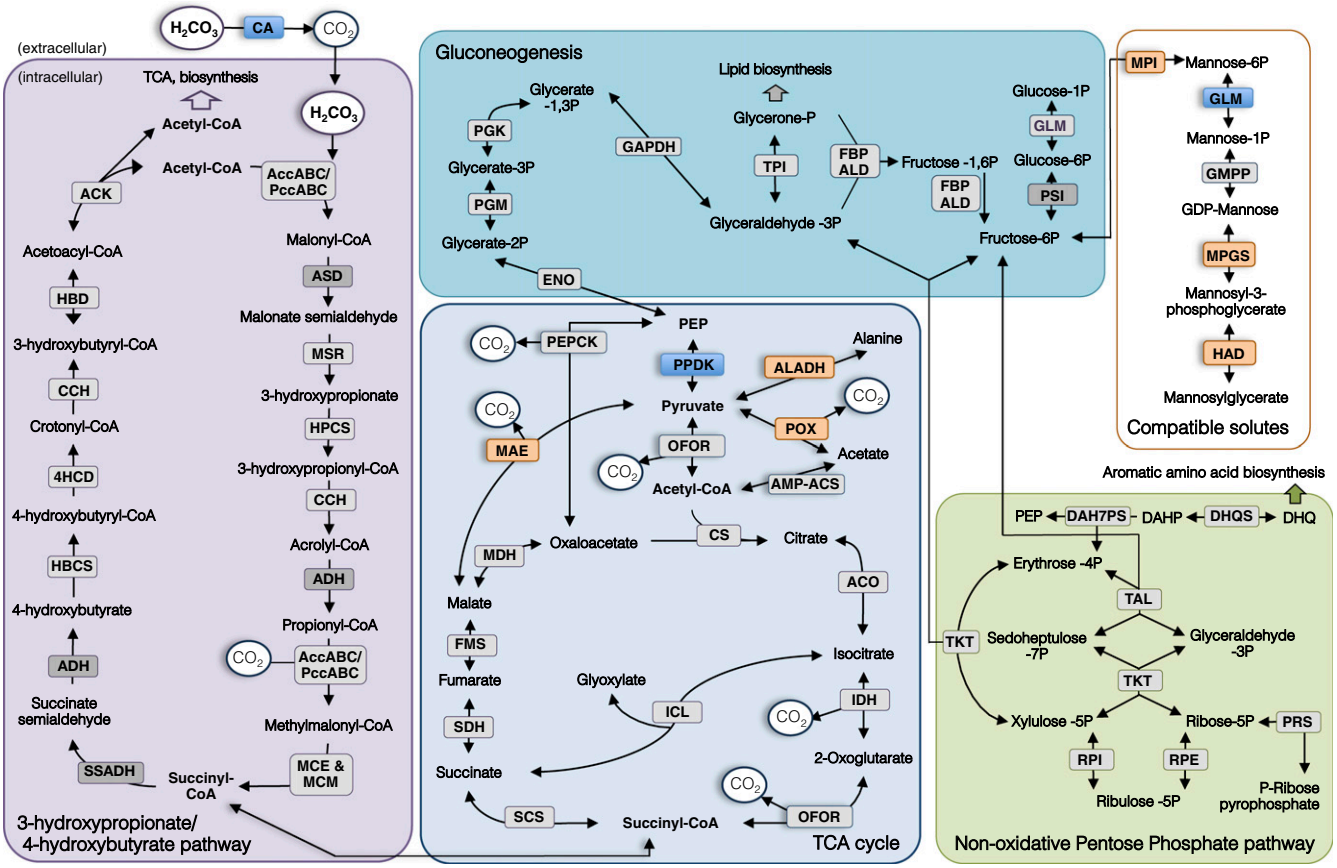

Fig. 2. Reconstruction of the predicted central carbon metabolism modules in AOA emphasizing the conservation of the core pathway enzymes and carriers and features of the Nitrososphaerales. Canonical enzymes belonging to the core COGs are depicted in light gray boxes, whereas candidate enzymes according to the work by Spang et al. (28) are in dark gray boxes. COGs conserved among Nitrososphaerales are depicted in orange boxes, and COGs present in some of the analyzed genomes are in blue (refer to the color key in Fig. 3). Proteins catalyzing all displayed reactions were detected in the proteome of *N. viennensis* by proteotypic peptides. Gene accession numbers and enzyme abbreviations are listed in Dataset S2.

are present in all analyzed AOA genomes and both proteomes, verifying the conserved metabolic features of the phylum (Datasets S1 and S2). A full reconstruction of the central carbon metabolism based on our comparative analysis is given in Fig. 2, distinguishing the core AOA gene repertoire (genes labeled in gray) from genes with a more limited distribution (orange and blue). Notably, Nitrososphaerales seem to be particularly versatile with respect to interconversions of the key metabolite pyruvate (28, 29), as indicated by the presence of a pyruvate/phosphate dikinase, converting pyruvate to phosphoenolpyruvate (PEP), a putative pyruvate dehydrogenase/oxidase converting it to acetate, and an alanine dehydrogenase (Dataset S2).

All Thaumarchaeota harbor candidate enzymes for all steps of an oxidative TCA cycle, the functionality of which in vivo is, however, debated. In particular, the substrate specificity of the 2-oxoacid:ferredoxin oxidoreductase (NVIE_029480–NVIE_029490), the candidate enzyme for conversion of both pyruvate and 2-oxoglutarate to acetyl-CoA and succinyl-CoA, respectively, remains unknown (42), and its functioning in both steps of the cycle might be unlikely. Moreover, classical TCA bypass routes seem to be incomplete or noncanonical, as only one candidate enzyme (isocitrate lyase) catalyzing the first step of the glyoxylate shunt could be identified in the genome and is also detected in the proteome (Datasets S1 and S2).

As shown for *Metallosphaera sedula* (43) and proposed by Könneke et al. (23) for the AOA *N. maritimus*, succinyl-CoA is probably the primary carbon fixation product in the hydroxypropionate/hydroxybutyrate pathway. Notably, *N. viennensis* and other Nitrososphaerales contain a putative malic enzyme apart from the PEP carboxykinase, which could connect the C4/C3 compound pool in the absence of a full oxidative TCA cycle under autotrophic growth. For a better understanding of AOA central carbon metabolism (and other metabolisms), it will be crucial to perform biochemical analyses to verify and/or determine substrate specificities of the predicted components.

Although this analysis and other genomic and physiological investigations support the assumption that AOA are, in principle, autotrophic ammonia oxidizers, the possibility of alternative metabolisms was indicated by studies showing growth of AOA without apparent ammonia oxidation activity (44, 45) and the finding of transporters for organics (13) as well as incorporation studies of amino acids (46). However, despite the large amount of transporters encoded in the genome of *N. viennensis* (13), we have so far not been able to identify alternate energy or carbon sources in growth experiments. In addition, Kim et al. (47) have convincingly shown that the dependency of *N. viennensis* and other strains on small amounts of organic acids (16, 17, 48) can be overcome by reducing reactive oxygen species in the medium. Consequently, based on growth experiments and genomic predictions, all AOA currently in pure culture can grow strictly autotrophically.

The AOA Core Genomic Repertoire Involved in Ammonia Oxidation. Based on the genomic and proteomic data, we have reconstructed the possible pathways of energy metabolism (via ammonia oxidation) and ammonia assimilation as they are found in all investigated thaumarchaeal genomes and translated in

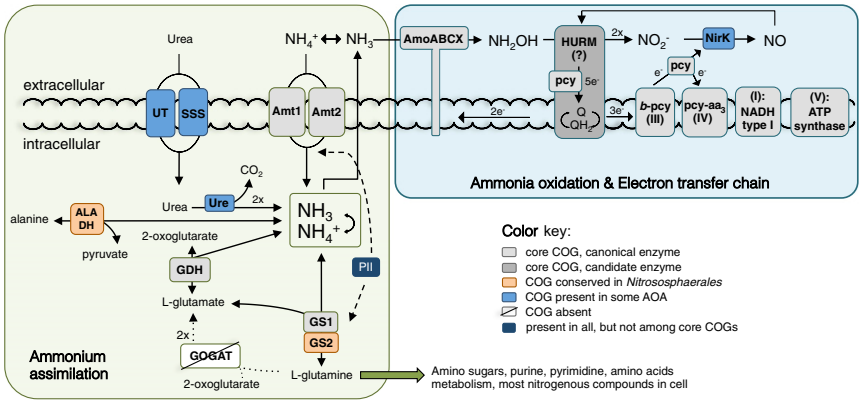

Fig. 3. Reconstruction of putative ammonia oxidation and primary nitrogen assimilation in AOA emphasizing the conservation of the core pathway enzymes and carriers and features of the Nitrososphaerales. The ammonia oxidation module was adapted from ref. 21. Dark gray indicates candidate enzyme [in this case, the unidentified HURM (21)]. Dashed lines indicate binding/regulation. Dotted lines indicate absent reaction, which is included here for clarity. All depicted proteins except the urea transporters were detected in the proteome of *N. viennensis*. Gene accession numbers are in Dataset S2. The presence and function of urea transporters (UT and SSS) and urease (Ure) as well as the two Amt ammonium transporters have been presented earlier (refs. 16 and 13, respectively) and are not discussed here.

N. viennensis (Fig. 3 and Dataset S2, gene accession numbers of *N. viennensis*). Among the core protein families involved in energy metabolism are the ammonia monooxygenase (AMO) subunits A, B, and C, which are homologs of the bacterial counterparts. A gene for a fourth potential membrane-associated subunit that we earlier referred to as ORF38 (11) or *amoX* (49) is linked to *amoA* in all sequenced genomes. The AmoC protein family includes a single homolog in the *Ca.* Nitrosopumilales genomes but is enlarged in the soil-dwelling Nitrososphaerales and AOB (28, 29, 50), with *N. viennensis* encoding six homologs (72–96% amino acid identity). The core AmoC COG includes two copies in Nitrososphaerales (AmoC3 and C6 in *N. viennensis*). The additional homologs either belong to COGs shared between the Nitrososphaerales genomes (AmoC4,C5) or cluster with the organism-specific COGs (AmoC1,C2), reflecting possible regulatory or stress response functions. Transcription studies in the AOB *Nitrosomonas europaea* implicate a role of additional *amoC* copies in the recovery from ammonia starvation or other types of stress by ensuring the stability of the AMO holoenzyme (51, 52). The A, B, and X subunits were detected in both proteomic datasets, whereas the core C subunit was only detected in the *N. brevis* proteome. AmoB was the second most abundant protein in the *N. viennensis* proteome (Dataset S1), supporting observations on the environmental abundance (or ease of detection) of this subunit from metaproteomic studies (53).

Although the identity of the enzyme(s) for the second step in ammonia oxidation remains unknown, hydroxylamine has been verified as an intermediate compound in *N. maritimus* (54). Currently, models for ammonia oxidation in AOA and AOB assume that the active site of the AMO is located on the periplasmic side of the membrane (21, 24, 26, 55). Consequently, it is suggested that the following oxidation of the intermediate hydroxylamine should also take place in the (pseudo)periplasmic space of AOA (i.e., in the space between the cytoplasmic membrane and the S-layer). An extracellular location of the oxidation of hydroxylamine into nitrite is bioenergetically favorable, because it contributes to the establishment of the proton gradient across the cell membrane via the release of scalar protons in the periplasmic space (24, 55). The intracellular oxidation of hydroxylamine would instead require the transport of the highly reactive and toxic hydroxylamine into the cytoplasm and the secretion of nitrite outside the cell, which seems, at least theoretically, unfavorable. The immediate implication of those considerations is that the enzyme(s) participating in the oxidation of hydroxylamine would be secreted as shown for the hydroxylamine oxidoreductase (HAO) in AOB.

We assume that all core and accessory enzymes in the energy metabolism would be among the core COGs in our analysis, and therefore, we searched for enzymes capable of performing oxidoreductase reactions among the core protein families. Naturally, this search is limited to the fraction of core COGs for which a function could be proposed, but we did not limit our search to putative candidates predicted to be membrane associated or secreted.

Five protein candidates were identified, which include two putative F420-dependent luciferase-like monooxygenase proteins (NVIE_004930 and NVIE_015060), with a domain found in flavin monooxygenases catalyzing mostly the oxidation of aldehydes in various pathways. Both proteins are present in high amounts in the proteomes of *N. viennensis* and *N. brevis*, suggesting an important role in the metabolism of those organisms. A putative NADPH-dependent F420 reductase (NVIE_027500) and two putative FAD- and FAD/NAD(P)-dependent oxidoreductases (NVIE_016850 and NVIE_007970) were also expressed in both proteome datasets. The predicted localization of these five enzymes is cytoplasmic (*Materials and Methods*), making them incompatible with the extracellular ammonia oxidation hypothesis. However, secreted proteins do not always possess a signal peptide, and studies of archaeal signal peptides are scarce; we cannot rule out the possibility that some of these enzymes are secreted via nonclassical protein secretion pathways (56).

The production of significant amounts of F420 by AOA was earlier verified by Spang et al. (28). The genes for the biosynthesis of this cofactor are among the core COGs, supporting the proposed important role of this cofactor in AOA metabolism (28). As noted by Spang et al. (28), reduced F420 generated by an F420 glucose-6-phosphate dehydrogenase in mycobacteria has a protective role against reactive nitrogen species (57). Thus, a similar F420-dependent mechanism might operate in AOA. Recently, a membrane-attached small cupredoxin from *N. maritimus* with a type 1 copper center (T1Cu) was shown to have reversible NO oxidation/NO_2^- reduction activity in vitro (58), indicating a need for detoxification of extracellularly produced reactive nitrogen species. Biochemical analyses will be needed to clarify whether F420 is involved in detoxification, is part of the archaeal hydroxylamine:ubiquinone redox module (HURM), catalyzing the oxidation of hydroxylamine (see below), or participates in both functions.

Conservation of Multicopper Proteins in AOA. It has been repeatedly proposed that a multicopper oxidase (MCO) could be the archaeal counterpart to the bacterial HAO or participate in the archaeal HURM (Fig. 3), as multiple MCOs that are predicted to be membrane associated or secreted were discovered in all AOA genomes (12, 21, 26, 55). However, none of the MCOs were found among the core COGs in this analysis. To examine the distribution of this protein family further, we conducted a thorough phylogenetic analysis on all available AOA genomes and representative genomes of other archaeal phyla and AOB/nitrite-oxidizing bacteria (details are in *SI Results and Discussion* and Table S2). The phylogenetic analysis reveals four well-supported major lineages of two-domain MCOs with type 1

(T1Cu) and trinuclear type 2/type 3 copper centers (labeled MCO1–4 in Fig. 4). MCO1 comprises only sequences from thaumarchaeal genomes, whereas MCO4 (with an additional blue copper domain) is present only in Nitrososphaerales and the genera *Nitrosopumilus* and *Nitrosoarchaeum*. MCO2 and MCO3 comprise nonthaumarchaeal sequences (with one exception) (Fig. S2 shows an uncollapsed tree). MCO1 orthologs are present in all described AOA, with the sole exception of *N. devanaterra*. All AOA (except *N. devanaterra*) encode a transporter of the Zinc-Iron Permease (ZIP) family immediately adjacent to MCO1 and on the same strand (except for *Nitrosotenuis chungbukensis* MY2 and *Nitrosotenuis uzonensis* N4, where the gene is unlinked). The co-occurrence pattern indicates a functional link and coregulation between these two proteins, which are also reflected in a shared evolutionary history (Fig. S3 shows the permease tree). Given that transporters of the ZIP family are involved in the transport of a range of divalent cations, including copper (Cu^{2+}), this pair of genes (MCO plus the ZIP family permease) could be involved in copper sequestration provided that the MCO is secreted. The absence of both genes in *N. devanaterra* further supports a functional link of the two and could be explained by the particularly low pH conditions that this archaeon encounters in the environment, which likely influences the bioavailability of copper in these soils (59) and might impose different copper uptake strategies. Alternatively, this MCO could still represent the missing module in ammonia oxidation (HURM in Fig. 3), whereas in the acidophile *N. devanaterra*, a substitution by a different enzyme might have occurred due to pH constraints. Such cases of nonorthologous gene replacements in acidophiles have already been reported. For example, a completely novel type of tetrathionate hydrolase essential for the dissimilatory oxidation of inorganic sulfur compounds has been exclusively found in acidophilic Bacteria and Archaea (60).

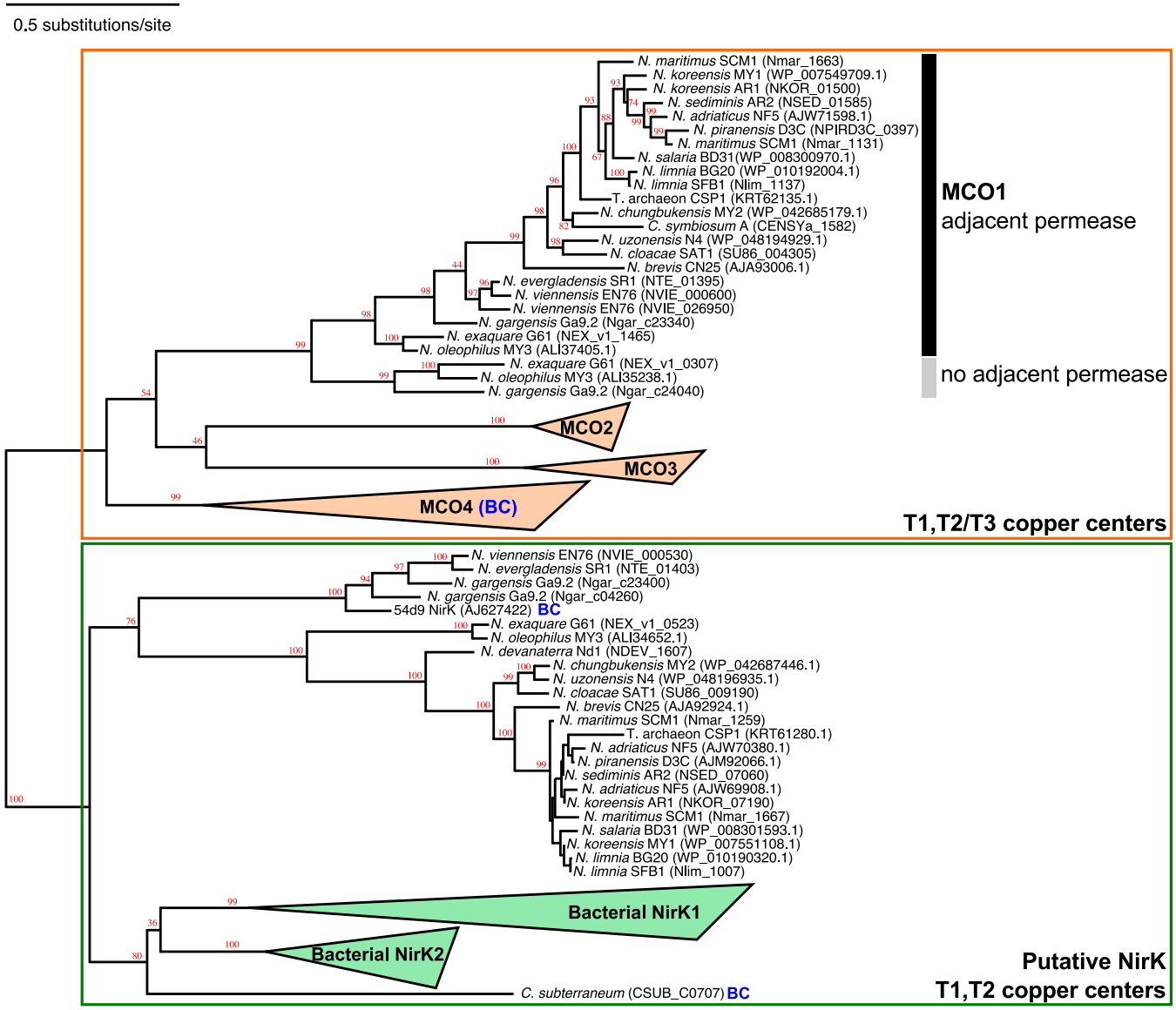

Fig. 4. Maximum likelihood tree (after automatic model selection with IQ-Tree, version 1.4.2) of multicopper proteins found in all available AOA genomes (25 in total) and Aigarchaeota genomes and selected genomes from AOB and nitrite-oxidizing bacteria. NirK is as identified in ref. 61. The two uncollapsed clades contain exclusively genes from AOA and are represented in almost all 25 species (MCO1 is not in *N. devanaterra*, and NirK is not in *C. symbiosum* and *N. yellowstonii*). More details and strain abbreviations are in the text and *SI Materials and Methods*. Values at nodes represent ultrafast bootstraps («UF-boot »). BC, additional blue copper domain.

The second major lineage represents the putative NO-forming nitrite reductase (NirK) family enzymes in both AOA and AOB (61), containing sequences with only two cupredoxin domains and T1, T2 Cu centers. NirK is broadly distributed in AOA with the exceptions of the symbiont *C. symbiosum* (39), for which evidence for capability of chemolithoautotrophic growth is still missing, and *Nitrosocaldus yellowstonii* (12), for which the full genome is not published. Nevertheless, NirK (or a noncanonical NirK) has been suggested to assist in the process of ammonia oxidation (14, 21, 26), because NO (the product of the NirK reaction) is essential for ammonia oxidation in AOA (18, 21, 62). In support of this hypothesis, NirK is found highly expressed on the mRNA level in cultivated AOA and environmental samples (63–65) and was identified in the *N. viennensis* proteome (NVIE_000530).

The AOA Core Proteome Involved in Ammonia Assimilation. Components of the ammonia assimilation machinery are mostly conserved among the analyzed genomes, but some genomic features differ between the three orders of AOA being compared (Fig. 3). All sequenced AOA encode two copies of glutamine synthetase (GS), but only one is among the core COGs. Glutamate dehydrogenase (GDH), the other key enzyme involved in the regulation of the glutamine/glutamate pool in the cell, also clusters within the core COGs, and both proteins (GDH and GS) are found in the proteomics datasets from *N. viennensis* and *N. brevis*. A homolog of glutamate synthase (GOGAT) was not identified in any of the analyzed genomes. The GS/GOGAT pathway usually exhibits higher affinity for ammonia, which might reflect an adjustment of AOA to higher intracellular substrate concentrations (discussion is in *SI Results and Discussion*).

Only genomes of Nitrososphaerales encode alanine dehydrogenases (Fig. 3) (28, 29). The amination of pyruvate could be a possible mechanism of ammonia storage or alleviation of ammonia toxicity or could also represent another primary nitrogen assimilation process as in AOB and other bacteria (66, 67).

Members of the PII superfamily of signal transduction proteins, key players in nitrogen regulation, are overrepresented in all analyzed AOA genomes with five to six homologs. These proteins have been shown to integrate information regarding the carbon/nitrogen ratio and energy status of the cell and regulate the ammonia uptake and use in response to changes in the extracellular nitrogen availability (68). This function is achieved by binding the key molecules 2-oxoglutarate and ATP/ADP/AMP and subsequently regulating a broad collection of target proteins (e.g., Amt family transporters, GS, and others) (68). The overrepresentation of these regulatory proteins in thaumarchaeal genomes, their relatively conserved genetic linkage to *amt* in a number of AOA, and the high relative abundance in the proteome of *N. viennensis* (five homologs detected of six encoded in the genome and three among the top 50 most abundant proteins) (Dataset S1) reinforce the putative central role of these molecules in maintaining a tight regulation and balance over the key metabolic processes of nitrogen and carbon assimilation.

Environmental Adaptations of the Terrestrial Organisms. A closer look at 1,195 COGs shared between the Nitrososphaerales genomes (present in at least two of three) (Fig. 1) enabled us to identify distinguishing characteristics of this mostly terrestrial lineage. The consistently larger genome sizes compared to the marine organisms, together with an apparently conserved core metabolism suggest a variety of environmental adaptations (see below). A total of 789 of the shared COGs are of unknown function, and proteins assigned to 411 of them are identified in the proteome, making up one-third of the dataset.

An additional 324 COGs are shared between members of the Nitrososphaerales and the soil-dwelling organisms of the *Ca.* Nitrosopumilales (i.e., *N. koreensis*) and *Ca.* Nitrosotaleales (*N. devanaterra*) (sum of 209, 39, and 76 shared COGs, respectively, in Fig. 1). Among these COGs, we encounter protein families related to detoxification, motility, chemotaxis, and cell surface modifications, the importance of which as soil adaptations is elaborated in the subsequent paragraphs.

Osmotic Regulation, CO_2 Uptake, and Detoxification in Terrestrial AOA. The presence of a putative mannosyl-3-phosphoglycerate synthase in the Nitrososphaerales-specific protein families indicates their capability to synthesize the compatible solute mannosylglycerate from mannose-6-phosphate (Fig. 2). This enzyme was detected in the *N. viennensis* proteome, suggesting the production of this osmotic regulator, which also functions as a thermoprotectant in hyperthermophiles (69).

N. viennensis, *N. evergladensis*, and *N. devanaterra* encode a γ-class carbonic anhydrase (CA) homolog (NVIE_017960) (Fig. 2), which contains an N-terminal signal peptide indicating its extracellular localization. This enzyme is responsible for the reversible hydration of CO_2 to bicarbonate and widely distributed in both autotrophic and heterotrophic organisms. It mainly fulfills the role of CO_2/bicarbonate transport between tissues/compartments and maintains the CO_2 partial pressure in a range favorable for the activity of carboxylating enzymes (70). The presence and localization of this enzyme in *N. viennensis* suggest a role in facilitating carbon transfer into the cell by extracellularly converting bicarbonate to CO_2, which can subsequently diffuse through the cell membrane. At a stable intracellular pH of 7, equilibrium favors rehydration to bicarbonate, which can then be used by carbon fixation enzymes, such as the acetyl-CoA/propionyl-CoA carboxylase. An analogous role has been proposed for CA homologs (of mostly the α and β classes) in AOBs, where under carbon-limiting conditions the activity of CA correlated with the rate of CO_2 fixation (71). In the case of terrestrial AOA, CA might constitute an adaptation to the discontinuous soil environment, where local and frequent changes in CO_2 pressure and pH can lead to low CO_2 availability. In contrast, marine AOA encode a sodium-bicarbonate family transporter (13) and would not require a CA homolog, since bicarbonate is the predominant form of inorganic carbon in marine environments. As this gene does not occur in *N. gargensis* (isolated from a biofilm), it might constitute a bona fide adaptation to the terrestrial lifestyle of soil AOA, which remains to be validated upon availability of more genomes.

Genomes of soil-associated Nitrososphaerales encode a higher number of genes belonging to the glyoxalase/bleomycin resistance/dioxygenase superfamily (SSF54593; 9 in *N. viennensis* and 14 in *N. evergladensis* as opposed to 3–6 in the *Ca.* Nitrosopumilales genomes and 6 in both *N. gargensis* and *Nitrosotalea*), an observation that stands true for some soil AOB genomes as well (e.g., 11 in *Nitrosospira multiformis*). This superfamily contains proteins involved in antibiotic resistance and degradation of aromatic compounds, an essential trait for soil organisms (72) and a likely function of at least some of the thaumarchaeal homologs. Six homologs are present in the proteome. An additional member of this superfamily is glyoxalase I, an enzyme involved in the glyoxal pathway for detoxification of methylglyoxal, which converts this toxic product of central carbon metabolism [also shown to be produced in archaea by dephosphorylation of glyceraldehyde-3-phosphate catalyzed by triose-P-isomerase (73)] to lactate, a precursor of F420 biosynthesis (74). Interestingly, glyoxalase II (the second enzyme of the glyoxal pathway) is a metallo-beta-lactamase fold-containing protein (PF00753), a number of which are also encoded in the genomes of Thaumarchaeota, suggesting the possible existence of this pathway.

Evidence for Exopolysaccharide Production and Cell Surface Modification Pathways in Nitrososphaerales. Nitrososphaerales encode an extensive gene repertoire for exopolysaccharide production

and cell surface modification pathways, which is a prerequisite of biofilm formation (Fig. 5 and Dataset S2). Compared with the *Ca.* Nitrosopumilales, they have expanded gene families of glycosyl transferases (GTs) represented by the CAZy families GT1 (PF04101), GT2 (PF00535), and GT4 (PF00534). GTs are involved in the modification of lipids and proteins by the synthesis and covalent attachment of polysaccharides and the biosynthesis of high-molecular weight heteropolymeric exopolysaccharides (75, 76). A number of the GTs in *N. viennensis* are related to the putative Vectorial Glycosyl Polymerization transporters (4.D.1), a family of transporters that couples polysaccharide biosynthesis with membrane translocation. The carbohydrate esterase family 4 (CE4; 13–15 in *N. viennensis* and *N. evergladensis* compared with 1–4 in all others) (Fig. 5 and Dataset S2) includes enzymes responsible for the hydrolysis of *N*- or *O*-linked acetyl groups from *N*-acetylglucosamine or acetylxylose residues, respectively, which comprise the subunits of polymeric compounds such as chitin, peptidoglycan, and xylan (77). The majority of the thaumarchaeal homologs encode signal peptides and are predicted to be extracellular or membrane attached (Dataset S2). Additionally, all sequenced AOA possess an E-type oligosaccharyl-transferase AglB (NVIE_025270), indicating the ability to perform *N*-linked glycosylation (75, 76), and a Multidrug/Oligosaccharidyl-lipid/Polysaccharide (MOP) flippase. The above-mentioned protein families (GT, CE4, MOP, and AglB) feature prominently in the cell surface modification and extracellular polymeric substances (EPS) production pathways of biofilm-forming bacteria and archaea (75, 76, 78–80). Genes involved in these pathways are often organized in operons (76, 78–80), and in *N. viennensis* we can identify at least six gene clusters encoding genes with related functions (Fig. S4), a number of which are detected in the proteome (Fig. 5).

Although the molecular mechanisms of biofilm formation among the archaea remain largely unknown (81), production and modification of various types of EPS are essential processes in biofilm development and the establishment of cell–cell and cell–surface interactions (82). These complex phenotypes entail the existence of pathways for acetamidosugar biosynthesis (UDP-Glu, UDP-Gal, GDP-Man, UDP-GlcNAc, and UDP-ManNAc), oligosaccharide assembly on an appropriate carrier (usually dolichol), membrane translocation, polymerization, and modification (Fig. 5 and Dataset S2). Distribution patterns of key enzymes for the pathways involved reveal distinct capacities for the different thaumarchaeal

Domain description	Nvie	Never	Ngar	Ndev	Nkor	Nmar	Nbrev	Nvie Proteome
Extracellular Polymeric Substances (EPS) biosynthesis and Cell Surface Modification								
Oligosaccharyl transferase STT3 subunit (AglB)								1/1
Lysylphosphatidylglycerol synthetase/glycosyltransferase AglD								1/1
MOP flippase (TCDB 2.A.66.2) superfamily and other families involved in cell wall biogenesis and export of polysaccharides								0/1
Glycosyl transferase families (GT) 1, 2, 4, 28, family 2-like								7/15
Six-hairpin glycosidases								3/4
Glycoside hydrolase families 5, 14								1/2
Sialidase (neuraminidase) - family proteins								2/6
Polysaccharide deacetylase (CE4)								6/13
Galactose, PG binding-like & PG-binding lysin domain								1/5
Dolichyl-phosphate-mannose-protein mannosyltransferase (GT family 39)								1/1
Polysaccharide biosynthesis protein, CapD-like domain & Capsule polysac. biosynthesis protein								-
N-acetylneuraminic acid synthase, N-terminal								-
CDP-glycerol glycerophosphotransferase								-
Acetamidosugar biosynthesis genes: biosynthesis of UDP-Glu, UDP-Gal, GDP-Man, UDP-GlcNAc, UDP-ManNAc, dTDP-L-rhamnose								
glutamine-fructose-6-phosphate transaminase (glmS)								1/2
phosphoglucomutase, first 3 domains (glmM)								2/2
UDP-GlcNAc diphosphorylase/GlcNAc-1-P N-acetyltransferase (GlmU) bact. type								1/1
UDP-N-acetylglucosamine 2-epimerase (wecB)								2/2
UDP-N-acetyl-D-mannosaminuronate dehydrogenase (wecC)								4/4
Galactose-1-phosphate uridyl transferase, N-terminal (galT)								1/1
UDP-glucose 4-epimerase (galE)								0/1
Other NAD-dependent putative sugar epimerases/dehydratases								-
Mannose-6-phosphate isomerase, type II								2/2
Nucleotide-diphospho-sugar transferases								7/15
Bifunctional glucose-6-phosphate/mannose-6-phosphate isomerase, C-terminal								-
dTDP-L-rhamnose biosynthesis genes (4 in total)								-
Motility & Adhesion								
autotransporter adhesin								3/4
coagulation factor 5/8 type domain								0/3
PKD domains								2/3
Large exoproteins involved in heme utilisation and adhesion								1/5
hyalin								3/4
YVTN beta-propeller repeats								1/2
fibronectin								1/1
S-layer domain								1/2
type IV related pili cluster								2/3
Archaellum: archaeal flagelin								2/3

Fig. 5. Distribution of polysaccharide biosynthesis and adhesion-related domains in the analyzed AOA genomes. Shading reflects the expansion of the respective gene family/domain: darker shades of red represent an increasing number of homologs, and white represents absence of the respective category in the genome. The number of homologs detected in the proteome of *N. viennensis*, out of the total number of homologs in the *N. viennensis* genome, are listed in the last column. An extended version of this table, including the numbers and the protein family accession numbers, can be found in Dataset S2. Nbrev, *N. brevis*; Ndev, *N. devanaterra*; Never, *N. evergladensis*; Ngar, *N. gargensis*; Nkor, *N. koreensis*; Nmar, *N. maritimus*; and Nvie, *N. viennensis*.

orders [e.g., a deoxythymidine diphosphate (dTDP)-rhamnose biosynthesis pathway in *Ca.* Nitrosopumilales] (Fig. 5).

Targets and Role of Cell Surface Modifications in Terrestrial AOA. Little is known regarding the targets for cell surface modifications in Thaumarchaeota. The S-layer proteins and archaella components (archaellins) have been shown to be heavily glycosylated in archaea, with the degree and type of glycosylation varying greatly among species (83). In *N. viennensis* and *N. evergladensis*, the two putative S-layer subunits present in all Thaumarchaeota and identified in high abundance in the proteomics dataset (NVIE_016730 and NVIE_016740) are encoded adjacent to a series of genes involved in polysaccharide production and modification. Both S-layer proteins contain four and five canonical *N*-glycosylation sequons (N-X-S/T motifs) respectively, suggesting that they might be heavily modified.

Interestingly, the modification of S-layer proteins seems to be of a highly dynamic nature in Archaea and can be modified reversibly in response to altering environmental conditions (84). Glycosylation seems to have an adaptive role, enabling the alteration of cell surface properties (e.g., charge distribution, hydration layer, or interaction between cell surface proteins) to provide protection, promote survival against rapidly changing environmental conditions, and establish inter- or intraspecies interactions (82, 85). Moreover, the large S-layer subunit conserved in all Thaumarchaeota clusters in two different COGs representing the two orders *Ca.* Nitrosopumilales and Nitrososphaerales, suggesting adaptations to the different environmental habitats at a single-protein level (Fig. 5 and Dataset S2).

Additional evidence for cell surface modification mechanisms comes from the fact that all sequenced AOA encode an archaeosortase/exosortase family transpeptidase (also detected in the proteome). This family is found in Bacteria and Euryarchaea, where it correlates with biofilm-forming strains and has been implicated in biofilm formation (86, 87). Distinct subfamilies recognize specific C-terminal sorting signals in target proteins (which include the S-layer proteins) and mediate their membrane translocation, proteolysis, and covalent attachment to the membrane, usually accompanied by extensive posttranslational modifications (86, 87).

Motility and Cell Adhesion Mechanisms in Terrestrial AOA. Motility-associated genes are encoded in all investigated genomes of the genus *Nitrososphaera* as well as in *N. devanaterra* (17, 28, 29, 31) (Fig. 5). These organisms possess gene clusters dedicated to chemotaxis, biosynthesis, and assembly of an archaellum and a type IV-related pili cluster. Cell appendages have been implicated in biofilm formation and surface adhesion (reviewed in ref. 88), and it is interesting to speculate an equivalent function in AOA given the genomic repertoire in support of the capacity for biofilm formation. We observed both types of cell appendages during morphological characterization of the strain using transmission EM and SEM (17).

All AOA isolated from soil encode a large number of predicted membrane proteins with diverse cell adhesion-related domains, a number of which are detected in the proteome of *N. viennensis* (Fig. 5). The presence of these protein families suggests an increased capacity for recognition of and adhesion to an array of different surfaces.

Biofilm Production in Nitrososphaerales. Our analysis suggests that *N. viennensis*, all analyzed Nitrososphaerales, and to a much lesser extent *Ca.* Nitrosopumilales, seem to encode the genomic inventory for both EPS production and cell surface modification, with an apparent versatility in the types of polysaccharides produced. Moreover, the strains have the potential for surface adhesion and to switch between a sessile and motile lifestyle. These observations indicate an increased ability to adapt to their extracellular environment in order to achieve the optimal conditions for cell–surface and interspecies interactions, and to form synergistic microconsortia in the context of a biofilm community (89).

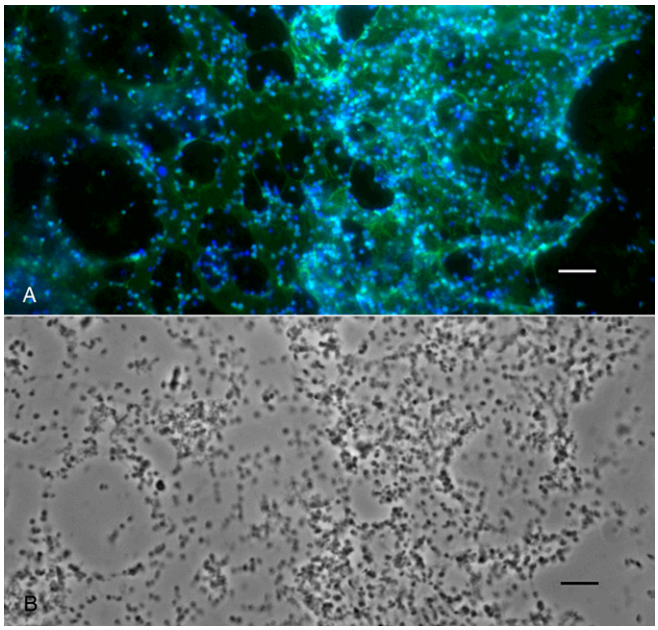

Fig. 6. *N. viennensis* biofilm grown on a glass coverslip: (*A*) stained with DAPI (blue) and FITC-conjugated WGA (green), the latter binding to *N*-acetylglucosamine and/or *N*-acetylneuraminic acid residues; (*B*) phase-contrast image of the biofilm. (Scale bar: 5 μm.)

Indeed, we describe here extensive biofilm formation in *N. viennensis* continuous batch and fed batch cultures provided with appropriate attachment surfaces. In particular, aggregates and EPS-like structures were observed during microscopic investigations of stationary cultures. Fig. 6 shows an example of biofilm formation in a fully grown culture stained with the lectin wheat germ agglutinin (WGA), indicating the presence of *N*-acetylglucosamine and/or *N*-acetylneuraminic acid residues composing the EPS.

Conclusions and Outlook

With the increasing number of genomes of Thaumarchaeota becoming available, it is now possible to obtain a better picture of conserved and specific adaptations of AOA. Therefore, we discuss in this study the genome and proteome of *N. viennensis* in the context of six other fully closed genomes of free-living AOA to identify commonalities and predict specific adaptive features of the terrestrial organisms. From this study, it becomes clear that AOA share an impressive number of conserved genes and that their central energy and carbon metabolism is highly conserved and evolved independently from their bacterial counterparts. AOA contain an ammonia monooxygenase that is only remotely related to that of AOB and contains four rather than three subunits. Most probably, it catalyzes the first step in ammonia oxidation, whereas the enzymes for the subsequent reactions remain enigmatic, in particular since earlier suggested candidate proteins (MCOs) do not belong to the conserved gene repertoire. Nevertheless, we suggest that multicopper proteins play an important role for AOA metabolism and we also suggest candidates for the oxidation of hydroxylamine that should be considered in future biochemical analyses, including F420-dependent enzymes.

The enlarged genomes of the Nitrososphaerales reveal the extensive environmental adaptation strategies of these soil

organisms, including detoxification, biofilm formation, adhesion, and cell–cell recognition, which indicate the potential for extensive interactions with other microorganisms in soil. Studying these features in an environmental context will be important to identify microbial interaction partners of AOA (for example, specific associations with nitrite oxidizers or other groups that could participate in the process of nitrogen species interconversions connected through the substrate–product chain) (90). It will be equally interesting to elucidate those adaptations that are the basis for the enormous ecological success of AOA in terrestrial environments.

Materials and Methods

Genome Annotation and Determination of COGs. Preparation of genomic DNA and sequencing have been described (16). Genome was assembled using single-end and mate-pair 454 pyrosequencing reads and remaining gaps were closed using Sanger sequencing of PCR-amplified gap regions. Prediction and annotation of ORFs were performed on the Microscope platform (91) with subsequent manual curation as in ref. 28. The annotated genome sequence has been deposited into GenBank under accession number CP007536. Predicted protein coding sequences from six thaumarchaeal genomes were retrieved from the National Center for Biotechnology Information RefSeq database (March of 2016). Bidirectional best hits (BBHs) between all genomes were calculated and clustered into COGs using the COGtriangles algorithm in the COGsoft package (92) with cutoff e values of $1e^{-10}$ and 70% minimal coverage for both proteins in each protein pair. Because of the small number of involved genomes, BBHs not part of any triangle were represented as specific proteins for each organism.

Phylogenetic Analysis. MCO and ZIP permease sequences were retrieved by using hidden Markov model (HMM) profiles based on manually curated alignments. From the filtered alignments, phylogenetic trees were built by maximum likelihood after automatic model selection with IQ-Tree, version 1.4.2 (93).

Cultivation and Biofilm Analysis. The strain was cultivated as described (17) in 20 mL with glass coverslips to allow adherence. The glass coverslips were dried, fixed, and stained with DAPI and FITC-labeled WGA (Sigma).

Proteome Analysis. Late exponential-phase cells were lysed by homogenization and sonication, and the lysate was fractionated into membrane and cytoplasmic fractions by ultracentrifugation in 60% (wt/vol) sucrose gradient. Proteins were extracted and processed as in ref. 94, and generated peptides were analyzed on an LC-LTQ-Orbitrap Mass Spectrometer (Thermo Scientific Inc.) coupled to ultrahigh-pressure nanoLC according to ref. 94. Data are available via ProteomeXchange with identifier PXD005297.

Detailed information on the bioinformatic tools and experimental procedures can be found in *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank all current and former members of the Archaea Biology and Ecogenomics Division for continuous discussions on the *N. viennensis* genome, in particular Anja Spang, Michaela Stieglmeier, Simon Rittmann, and Filipa Sousa. We also thank Thomas Weinmaier, Romana Bittner, and Tobias Reinelt for technical assistance and the reviewers for their valuable comments on the manuscript. This work was financed through Austrian Science Fund Grants P27017 and W1257 (to C.S.) and P26342 (to W.W.); European Research Council Advanced Grant ERC-ADG 695192 (to C.S.); and Marie Skłodowska-Curie Grants FP7-PEOPLE-2009-IEF-255109 (to L.V.) and H2020-MSCA-IF-2017:701981 (to S.S.A.).

1. Brochier-Armanet C, Boussau B, Gribaldo S, Forterre P (2008) Mesophilic Crenarchaeota: Proposal for a third archaeal phylum, the Thaumarchaeota. *Nat Rev Microbiol* 6(3):245–252.
2. Spang A, et al. (2010) Distinct gene set in two different lineages of ammonia-oxidizing archaea supports the phylum Thaumarchaeota. *Trends Microbiol* 18(8): 331–340.
3. Karner MB, DeLong EF, Karl DM (2001) Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* 409(6819):507–510.
4. Lin X, et al. (2006) Comparison of vertical distributions of prokaryotic assemblages in the anoxic Cariaco Basin and Black Sea by use of fluorescence in situ hybridization. *Appl Environ Microbiol* 72(4):2679–2690.
5. Leininger S, et al. (2006) Archaea predominate among ammonia-oxidizing prokaryotes in soils. *Nature* 442(7104):806–809.
6. Erguder TH, Boon N, Wittebolle L, Marzorati M, Verstraete W (2009) Environmental factors shaping the ecological niches of ammonia-oxidizing archaea. *FEMS Microbiol Rev* 33(5):855–869.
7. Metcalf WW, et al. (2012) Synthesis of methylphosphonic acid by marine microbes: A source for methane in the aerobic ocean. *Science* 337(6098):1104–1107.
8. Offre P, Spang A, Schleper C (2013) Archaea in biogeochemical cycles. *Annu Rev Microbiol* 67(1):437–457.
9. Doxey AC, Kurtz DA, Lynch MD, Sauder LA, Neufeld JD (2015) Aquatic metagenomes implicate Thaumarchaeota in global cobalamin production. *ISME J* 9(2):461–471.
10. Könneke M, et al. (2005) Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* 437(7058):543–546.
11. Treusch AH, et al. (2005) Novel genes for nitrite reductase and Amo-related proteins indicate a role of uncultivated mesophilic crenarchaeota in nitrogen cycling. *Environ Microbiol* 7(12):1985–1995.
12. Stahl DA, de la Torre JR (2012) Physiology and diversity of ammonia-oxidizing archaea. *Annu Rev Microbiol* 66:83–101.
13. Offre P, Kerou M, Spang A, Schleper C (2014) Variability of the transporter gene complement in ammonia-oxidizing archaea. *Trends Microbiol* 22(12):665–675.
14. Schleper C, Nicol GW (2010) Ammonia-oxidising archaea–physiology, ecology and evolution. *Adv Microb Physiol* 57:1–41.
15. Stieglmeier M, Alves R, Schleper C (2014) Thaumarchaeota. *The Prokaryotes—Other Major Lineages of Bacteria and the Archaea*, eds Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F (Springer, Berlin), 4th Ed, pp 1–40.
16. Tourna M, et al. (2011) *Nitrososphaera viennensis*, an ammonia oxidizing archaeon from soil. *Proc Natl Acad Sci USA* 108(20):8420–8425.
17. Stieglmeier M, et al. (2014) *Nitrososphaera viennensis* gen. nov., sp. nov., an aerobic and mesophilic, ammonia-oxidizing archaeon from soil and a member of the archaeal phylum Thaumarchaeota. *Int J Syst Evol Microbiol* 64:2738–2752.
18. Shen T, Stieglmeier M, Dai J, Urich T, Schleper C (2013) Responses of the terrestrial ammonia-oxidizing archaeon *Ca.* Nitrososphaera viennensis and the ammonia-oxidizing bacterium *Nitrosospira multiformis* to nitrification inhibitors. *FEMS Microbiol Lett* 344(2):121–129.
19. Taylor AE, et al. (2015) Inhibitory effects of C2 to C10 1-alkynes on ammonia oxidation in two *Nitrososphaera* species. *Appl Environ Microbiol* 81(6):1942–1948.
20. Stieglmeier M, et al. (2014) Aerobic nitrous oxide production through N-nitrosating hybrid formation in ammonia-oxidizing archaea. *ISME J* 8(5):1135–1146.
21. Kozlowski JA, Stieglmeier M, Schleper C, Klotz MG, Stein LY (2016) Pathways and key intermediates required for obligate aerobic ammonia-dependent chemolithotrophy in bacteria and Thaumarchaeota. *ISME J* 10(8):1836–1845.
22. Sinninghe Damsté JS, et al. (2012) Intact polar and core glycerol dibiphytanyl glycerol tetraether lipids of group I.1a and I.1b thaumarchaeota in soil. *Appl Environ Microbiol* 78(19):6866–6874.
23. Könneke M, et al. (2014) Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO2 fixation. *Proc Natl Acad Sci USA* 111(22):8239–8244.
24. Arp DJ, Stein LY (2003) Metabolism of inorganic N compounds by ammonia-oxidizing bacteria. *Crit Rev Biochem Mol Biol* 38(6):471–495.
25. Stein LY (2011) Surveying N2O-producing pathways in bacteria. *Methods in Enzymology: Research on Nitrification and Related Processes Part A*, eds Klotz MG, Stein LY (Elsevier, Amsterdam), Vol. 486, pp 131–152.
26. Walker CB, et al. (2010) *Nitrosopumilus maritimus* genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. *Proc Natl Acad Sci USA* 107(19):8818–8823.
27. Kim BK, et al. (2011) Genome sequence of an ammonia-oxidizing soil archaeon, "*Candidatus* Nitrosoarchaeum koreensis" MY1. *J Bacteriol* 193(19):5539–5540.
28. Spang A, et al. (2012) The genome of the ammonia-oxidizing *Candidatus* Nitrososphaera gargensis: Insights into metabolic versatility and environmental adaptations. *Environ Microbiol* 14(12):3122–3145.
29. Zhalnina KV, et al. (2014) Genome sequence of *Candidatus* Nitrososphaera evergladensis from group I.1b enriched from Everglades soil reveals novel genomic features of the ammonia-oxidizing archaea. *PLoS One* 9(7):e101648.
30. Santoro AE, et al. (2015) Genomic and proteomic characterization of "*Candidatus* Nitrosopelagicus brevis": An ammonia-oxidizing archaeon from the open ocean. *Proc Natl Acad Sci USA* 112(4):1173–1178.
31. Lehtovirta-Morley LE, et al. (2016) Identifying potential mechanisms enabling acidophily in the ammonia-oxidising archaeon " *Candidatus* Nitrosotalea devanaterra." *Appl Environ Microbiol* 82(9):2608–2619.
32. Blainey PC, Mosier AC, Potanina A, Francis CA, Quake SR (2011) Genome of a low-salinity ammonia-oxidizing archaeon determined by single-cell and metagenomic analysis. *PLoS One* 6(2):e16626.
33. Mosier AC, Allen EE, Kim M, Ferriera S, Francis CA (2012) Genome sequence of "*Candidatus* Nitrosopumilus salaria" BD31, an ammonia-oxidizing archaeon from the San Francisco Bay estuary. *J Bacteriol* 194(8):2121–2122.
34. Lebedeva EV, et al. (2013) Enrichment and genome sequence of the group I.1a ammonia-oxidizing Archaeon "*Ca.* Nitrosotenuis uzonensis" representing a clade globally distributed in thermal habitats. *PLoS One* 8(11):e80835.
35. Jung M-Y, et al. (2014) A mesophilic, autotrophic, ammonia-oxidizing archaeon of thaumarchaeal group I.1a cultivated from a deep oligotrophic soil horizon. *Appl Environ Microbiol* 80(12):3645–3655.
36. Park S-J, et al. (2014) Genomes of two new ammonia-oxidizing archaea enriched from deep marine sediments. *PLoS One* 9(5):e96449.
37. Bayer B, et al. (2016) Physiological and genomic characterization of two novel marine thaumarchaeal strains indicates niche differentiation. *ISME J* 10(5):1051–1063.
38. Li Y, et al. (2016) A novel ammonia-oxidizing archaeon from wastewater treatment plant: Its enrichment, physiological and genomic characteristics. *Sci Rep* 6:23747.

39. Hallam SJ, et al. (2006) Genomic analysis of the uncultivated marine crenarchaeote Cenarchaeum symbiosum. *Proc Natl Acad Sci USA* 103(48):18296–18301.
40. Prosser JI, Nicol GW (2012) Archaeal and bacterial ammonia-oxidisers in soil: The quest for niche specialisation and differentiation. *Trends Microbiol* 20(11):523–531.
41. Fuchs G (2011) Alternative pathways of carbon dioxide fixation: Insights into the early evolution of life? *Annu Rev Microbiol* 65:631–658.
42. Park YJ, Yoo CB, Choi SY, Lee HB (2006) Purifications and characterizations of a ferredoxin and its related 2-oxoacid:ferredoxin oxidoreductase from the hyperthermophilic archaeon, *Sulfolobus solfataricus* P1. *J Biochem Mol Biol* 39(1):46–54.
43. Estelmann S, et al. (2011) Labeling and enzyme studies of the central carbon metabolism in Metallosphaera sedula. *J Bacteriol* 193(5):1191–1200.
44. Alves RJE, et al. (2013) Nitrification rates in Arctic soils are associated with functionally distinct populations of ammonia-oxidizing archaea. *ISME J* 7(8):1620–1631.
45. Mussmann M, et al. (2011) Thaumarchaeotes abundant in refinery nitrifying sludges express *amoA* but are not obligate autotrophic ammonia oxidizers. *Proc Natl Acad Sci USA* 108(40):16771–16776.
46. Ouverney CC, Fuhrman JA (2000) Marine planktonic archaea take up amino acids. *Appl Environ Microbiol* 66(11):4829–4833.
47. Kim J-G, et al. (2016) Hydrogen peroxide detoxification is a key mechanism for growth of ammonia-oxidizing archaea. *Proc Natl Acad Sci USA* 113(28):7888–7893.
48. Qin W, et al. (2014) Marine ammonia-oxidizing archaeal isolates display obligate mixotrophy and wide ecotypic variation. *Proc Natl Acad Sci USA* 111(34):12504–12509.
49. Bartossek R, Spang A, Weidler G, Lanzen A, Schleper C (2012) Metagenomic analysis of ammonia-oxidizing archaea affiliated with the soil group. *Front Microbiol* 3:208–222.
50. Norton JM, Alzerreca JJ, Suwa Y, Klotz MG (2002) Diversity of ammonia monooxygenase operon in autotrophic ammonia-oxidizing bacteria. *Arch Microbiol* 177(2):139–149.
51. Berube PM, Samudrala R, Stahl DA (2007) Transcription of all amoC copies is associated with recovery of *Nitrosomonas europaea* from ammonia starvation. *J Bacteriol* 189(11):3935–3944.
52. Berube PM, Stahl DA (2012) The divergent AmoC3 subunit of ammonia monooxygenase functions as part of a stress response system in *Nitrosomonas europaea*. *J Bacteriol* 194(13):3448–3456.
53. Morris RM, et al. (2010) Comparative metaproteomics reveals ocean-scale shifts in microbial nutrient utilization and energy transduction. *ISME J* 4(5):673–685.
54. Vajrala N, et al. (2013) Hydroxylamine as an intermediate in ammonia oxidation by globally abundant marine archaea. *Proc Natl Acad Sci USA* 110(3):1006–1011.
55. Simon J, Klotz MG (2013) Diversity and evolution of bioenergetic systems involved in microbial nitrogen compound transformations. *Biochim Biophys Acta* 1827(2):114–135.
56. Bendtsen JD, Kiemer L, Fausbøll A, Brunak S (2005) Non-classical protein secretion in bacteria. *BMC Microbiol* 5:58.
57. Purwantini E, Mukhopadhyay B (2009) Conversion of NO_2 to NO by reduced coenzyme F_{420} protects mycobacteria from nitrosative damage. *Proc Natl Acad Sci USA* 106(15):6333–6338.
58. Hosseinzadeh P, et al. (2016) A purple cupredoxin from *Nitrosopumilus maritimus* containing a mononuclear type 1 copper center with an open binding site. *J Am Chem Soc* 138(20):6324–6327.
59. Reddy KJ, Wang L, Gloss SP (1995) Solubility and mobility of copper, zinc and lead in acidic environments. *Plant Soil* 171(1):53–58.
60. Protze J, et al. (2011) An extracellular tetrathionate hydrolase from the thermoacidophilic archaeon *Acidianus ambivalens* with an activity optimum at pH 1. *Front Microbiol* 2:68.
61. Bartossek R, Nicol GW, Lanzen A, Klenk HP, Schleper C (2010) Homologues of nitrite reductases in ammonia-oxidizing archaea: Diversity and genomic context. *Environ Microbiol* 12(4):1075–1088.
62. Sauder LA, Ross AA, Neufeld JD (2016) Nitric oxide scavengers differentially inhibit ammonia oxidation in ammonia-oxidizing archaea and bacteria. *FEMS Microbiol Lett* 363(7):1–8.
63. Shi Y, Tyson GW, Eppley JM, DeLong EF (2011) Integrated metatranscriptomic and metagenomic analyses of stratified microbial assemblages in the open ocean. *ISME J* 5(6):999–1013.
64. Hollibaugh JT, Gifford S, Sharma S, Bano N, Moran MA (2011) Metatranscriptomic analysis of ammonia-oxidizing organisms in an estuarine bacterioplankton assemblage. *ISME J* 5(5):866–878.
65. Lund MB, Smith JM, Francis CA (2012) Diversity, abundance and expression of nitrite reductase (nirK)-like genes in marine thaumarchaea. *ISME J* 6(10):1966–1977.
66. Caballero FJ, Cárdenas J, Castillo F (1989) Purification and properties of L-alanine dehydrogenase of the phototrophic bacterium *Rhodobacter capsulatus* E1F1. *J Bacteriol* 171(6):3205–3210.
67. Stein LY, et al. (2007) Whole-genome analysis of the ammonia-oxidizing bacterium, *Nitrosomonas eutropha* C91: Implications for niche adaptation. *Environ Microbiol* 9(12):2993–3007.
68. Huergo LF, Chandra G, Merrick M (2013) P(II) signal transduction proteins: Nitrogen regulation and beyond. *FEMS Microbiol Rev* 37(2):251–283.
69. Empadinhas N, da Costa MS (2006) Diversity and biosynthesis of compatible solutes in hyper/thermophiles. *Int Microbiol* 9(3):199–206.
70. Ferry JG (2010) The γ class of carbonic anhydrases. *Biochim Biophys Acta* 1804(2): 374–381.
71. Jahnke L, Lyman C, Hooper AB (1984) Carbonic anhydrase, carbondioxide levels and growth of *Nitrosomonas*. *Arch Microbiol* 140(2-3):291–293.
72. Tropel D, van der Meer JR (2004) Bacterial transcriptional regulators for degradation pathways of aromatic compounds. *Microbiol Mol Biol Rev* 68(3):474–500.
73. White RH, Xu H (2006) Methylglyoxal is an intermediate in the biosynthesis of 6-deoxy-5-ketofructose-1-phosphate: A precursor for aromatic amino acid biosynthesis in *Methanocaldococcus jannaschii*. *Biochemistry* 45(40):12366–12379.
74. Grochowski LL, Xu H, White RH (2006) Identification of lactaldehyde dehydrogenase in *Methanocaldococcus jannaschii* and its involvement in production of lactate for F420 biosynthesis. *J Bacteriol* 188(8):2836–2844.
75. Magidovich H, Eichler J (2009) Glycosyltransferases and oligosaccharyltransferases in Archaea: Putative components of the N-glycosylation pathway in the third domain of life. *FEMS Microbiol Lett* 300(1):122–130.
76. Jarrell KF, et al. (2014) N-linked glycosylation in archaea: A structural, functional, and genetic analysis. Microbiol Mol Biol Rev 78:304–341.
77. Caufrier F, Martinou A, Dupont C, Bouriotis V (2003) Carbohydrate esterase family 4 enzymes: Substrate specificity. *Carbohydr Res* 338(7):687–692.
78. Péant B, et al. (2005) Comparative analysis of the exopolysaccharide biosynthesis gene clusters from four strains of *Lactobacillus rhamnosus*. *Microbiology* 151:1839–1851.
79. Zhao D, et al. (2013) Improving polyhydroxyalkanoate production by knocking out the genes involved in exopolysaccharide biosynthesis in *Haloferax mediterranei*. *Appl Microbiol Biotechnol* 97(7):3027–3036.
80. Chimileski S, Franklin MJ, Papke RT (2014) Biofilms formed by the archaeon *Haloferax volcanii* exhibit cellular differentiation and social motility, and facilitate horizontal gene transfer. *BMC Biol* 12:65.
81. Orell A, Fröls S, Albers S (2013) Archaeal biofilms: The great unexplored. *Annu Rev Microbiol* 67(1):337–354.
82. Flemming H, Wingender J (2010) The biofilm matrix. *Nat Rev Microbiol* 8(9):623–633.
83. Jarrell KF, Jones GM, Kandiba L, Nair DB, Eichler J (2010) S-layer glycoproteins and flagellins: Reporters of archaeal posttranslational modifications. *Archaea* 2010:612948.
84. Guan Z, Naparstek S, Calo D, Eichler J (2012) Protein glycosylation as an adaptive response in Archaea: Growth at different salt concentrations leads to alterations in *Haloferax volcanii* S-layer glycoprotein N-glycosylation. *Environ Microbiol* 14(3): 743–753.
85. Kaminski L, et al. (2013) Add salt, add sugar: N-glycosylation in *Haloferax volcanii*. *Biochem Soc Trans* 41(1):432–435.
86. Haft DH, Payne SH, Selengut JD (2012) Archaeosortases and exosortases are widely distributed systems linking membrane transit with posttranslational modification. *J Bacteriol* 194(1):36–48.
87. Abdul Halim MF, et al. (2013) *Haloferax volcanii* archaeosortase is required for motility, mating, and C-terminal processing of the S-layer glycoprotein. *Mol Microbiol* 88(6):1164–1175.
88. Lassak K, Ghosh A, Albers SV (2012) Diversity, assembly and regulation of archaeal type IV pili-like and non-type-IV pili-like surface structures. *Res Microbiol* 163(9-10): 630–644.
89. Comolli LR, Banfield JF (2014) Inter-species interconnections in acid mine drainage microbial communities. *Front Microbiol* 5:367.
90. Daebeler A, et al. (2014) Interactions between Thaumarchaea, *Nitrospira* and methanotrophs modulate autotrophic nitrification in volcanic grassland soil. *ISME J* 8(12):2397–2410.
91. Vallenet D, et al. (2009) MicroScope: A platform for microbial genome annotation and comparative genomics. *Database (Oxford)* 2009:bap021.
92. Kristensen DM, et al. (2010) A low-polynomial algorithm for assembling clusters of orthologous groups from intergenomic symmetric best matches. *Bioinformatics* 26(12):1481–1487.
93. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ (2015) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 32(1):268–274.
94. Valledor L, Weckwerth W (2014) An improved detergent-compatible gel-fractionation LC-LTQ-Orbitrap-MS workflow for plant and microbial proteomics. *Methods in Molecular Biology*, ed Walker JM (Humana, Clifton, NJ), pp 347–358.
95. Rouillon C, White MF (2011) The evolution and mechanisms of nucleotide excision repair proteins. *Res Microbiol* 162(1):19–26.
96. Palatinszky M, et al. (2015) Cyanate as an energy source for nitrifiers. *Nature* 524(7563):105–108.
97. Eddy SR (2011) Accelerated profile HMM searches. *PLOS Comput Biol* 7(10): e1002195.
98. Bagos PG, Tsirigos KD, Plessas SK, Liakopoulos TD, Hamodrakas SJ (2009) Prediction of signal peptides in archaea. *Protein Eng Des Sel* 22(1):27–35.
99. Bendtsen JD, Nielsen H, Widdick D, Palmer T, Brunak S (2005) Prediction of twin-arginine signal peptides. *BMC Bioinformatics* 6:167.
100. Käll L, Krogh A, Sonnhammer EL (2007) Advantages of combined transmembrane topology and signal peptide prediction–the Phobius web server. *Nucleic Acids Res* 35:W429–W432.
101. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol Biol Evol* 30(4):772–780.
102. Zybailov B, et al. (2006) Statistical analysis of membrane proteome expression changes in *Saccharomyces cerevisiae*. *J Proteome Res* 5(9):2339–2347.
103. May P, et al. (2008) Metabolomics- and proteomics-assisted genome annotation and analysis of the draft metabolic network of *Chlamydomonas reinhardtii*. *Genetics* 179(1):157–166.
104. Hatzenpichler R, et al. (2008) A moderately thermophilic ammonia-oxidizing crenarchaeote from a hot spring. *Proc Natl Acad Sci USA* 105(6):2134–2139.
105. Santoro AE, Casciotti KL (2011) Enrichment and characterization of ammonia-oxidizing archaea from the open ocean: Phylogeny, physiology and stable isotope fractionation. *ISME J* 5(11):1796–1808.
106. Lehtovirta-Morley LE, Stoecker K, Vilcinskas A, Prosser JI, Nicol GW (2011) Cultivation of an obligate acidophilic ammonia oxidizer from a nitrifying acid soil. *Proc Natl Acad Sci USA* 108(38):15892–15897.