



HAL
open science

Candidatus Nitrosocaldus cavascurensis, an Ammonia Oxidizing, Extremely Thermophilic Archaeon with a Highly Mobile Genome

Sophie S. Abby, Michael Melcher, Melina Kerou, Mart Krupovic, Michaela Stieglmeier, Claudia Rossel, Kevin Pfeifer, Christa Schleper

► **To cite this version:**

Sophie S. Abby, Michael Melcher, Melina Kerou, Mart Krupovic, Michaela Stieglmeier, et al.. Candidatus Nitrosocaldus cavascurensis, an Ammonia Oxidizing, Extremely Thermophilic Archaeon with a Highly Mobile Genome. *Frontiers in Microbiology*, 2018, 9, pp.28. 10.3389/fmicb.2018.00028 . hal-01702026

HAL Id: hal-01702026

<https://hal.science/hal-01702026>

Submitted on 6 Feb 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Candidatus Nitrosocaldus cavascurensis, an Ammonia Oxidizing, Extremely Thermophilic Archaeon with a Highly Mobile Genome

Sophie S. Abby^{1,2†}, Michael Melcher^{1†}, Melina Kerou¹, Mart Krupovic³, Michaela Stieglmeier¹, Claudia Rossel¹, Kevin Pfeifer¹ and Christa Schleper^{1*}

¹ Division of Archaea Biology and Ecogenomics, Department of Ecogenomics and Systems Biology, University of Vienna, Vienna, Austria, ² Laboratoire Techniques de l'Ingénierie Médicale et de la Complexité – Informatique, Mathématiques et Applications, Centre National de la Recherche Scientifique, Université Grenoble Alpes, Grenoble, France,

³ Unité Biologie Moléculaire du Gène chez les Extrémophiles, Institut Pasteur, Paris, France

OPEN ACCESS

Edited by:

Martin G. Klotz,
Washington State University Tri-Cities,
United States

Reviewed by:

Ivan A. Berg,
Universität Münster, Germany
Willm Martens-Habbena,
University of Florida, United States
Lisa Y. Stein,
University of Alberta, Canada

*Correspondence:

Christa Schleper
christa.schleper@univie.ac.at

[†]These authors have contributed
equally to this work.

Specialty section:

This article was submitted to
Evolutionary and Genomic
Microbiology,
a section of the journal
Frontiers in Microbiology

Received: 28 December 2017

Accepted: 08 January 2018

Published: 26 January 2018

Citation:

Abby SS, Melcher M, Kerou M,
Krupovic M, Stieglmeier M,
Rossel C, Pfeifer K and Schleper C
(2018) *Candidatus Nitrosocaldus*
cavascurensis, an Ammonia
Oxidizing, Extremely Thermophilic
Archaeon with a Highly Mobile
Genome. *Front. Microbiol.* 9:28.
doi: 10.3389/fmicb.2018.00028

Ammonia oxidizing archaea (AOA) of the phylum Thaumarchaeota are widespread in moderate environments but their occurrence and activity has also been demonstrated in hot springs. Here we present the first enrichment of a thermophilic representative with a sequenced genome, which facilitates the search for adaptive strategies and for traits that shape the evolution of Thaumarchaeota. *Candidatus Nitrosocaldus cavascurensis* has been enriched from a hot spring in Ischia, Italy. It grows optimally at 68°C under chemolithoautotrophic conditions on ammonia or urea converting ammonia stoichiometrically into nitrite with a generation time of approximately 23 h. Phylogenetic analyses based on ribosomal proteins place the organism as a sister group to all known mesophilic AOA. The 1.58 Mb genome of *Ca. N. cavascurensis* harbors an *amo*AXCB gene cluster encoding ammonia monooxygenase and genes for a 3-hydroxypropionate/4-hydroxybutyrate pathway for autotrophic carbon fixation, but also genes that indicate potential alternative energy metabolisms. Although a *bona fide* gene for nitrite reductase is missing, the organism is sensitive to NO-scavenging, underlining the potential importance of this compound for AOA metabolism. *Ca. N. cavascurensis* is distinct from all other AOA in its gene repertoire for replication, cell division and repair. Its genome has an impressive array of mobile genetic elements and other recently acquired gene sets, including conjugative systems, a provirus, transposons and cell appendages. Some of these elements indicate recent exchange with the environment, whereas others seem to have been domesticated and might convey crucial metabolic traits.

Keywords: ammonia oxidation, archaea, Thaumarchaeota, nitrification, mobile genetic elements, thermophilic adaptations, *amoA* gene, DNA repair enzymes

INTRODUCTION

Ammonia oxidizing archaea (AOA) now collectively classified as *Nitrososphaeria* within the phylum Thaumarchaeota (Brochier-Armanet et al., 2008; Spang et al., 2010; Stieglmeier et al., 2014; Kerou et al., 2016a) represent the sole archaeal group that is globally distributed in oxic environments efficiently competing with aerobic bacteria. Because of their large numbers in the

ocean plankton, in marine sediments, in lakes and in soils, AOA are considered one of the most abundant groups of prokaryotes on this planet (Schleper et al., 2005; Prosser and Nicol, 2008; Erguder et al., 2009; Schleper and Nicol, 2010; Pester et al., 2011; Hatzenpichler, 2012; Stahl and de la Torre, 2012; Offre et al., 2013). All currently cultivated AOA strains gain energy exclusively through the oxidation of ammonia to nitrite, i.e., they perform the first step in nitrification. They grow chemolithoautotrophically from inorganic carbon supply. Some strains show growth only in the presence of small organic acids (Tourna et al., 2011; Qin et al., 2014) which seem to catalyze degradation of reactive oxygen species (ROS) from the medium (Kim et al., 2016). Different from their bacterial ammonia oxidizing counterparts, AOA are often adapted to rather low levels of ammonia for growth, which seems to favor their activities in oligotrophic environments, such as the marine pelagic ocean and marine sediments, but also in acidic environments, where the concentration of ammonia decreases in favor of ammonium (Nicol et al., 2008). However, AOA occur also in large numbers in terrestrial environments, including fertilized soils and some waste water treatment plants, and several studies indicate alternative energy metabolisms (Mussmann et al., 2011; Alves et al., 2013). Although ammonia oxidation in archaea has not been biochemically resolved, the presence of genes for an ammonia monooxygenase in all AOA with remote similarity to methane and ammonia monooxygenases of bacteria implies involvement of this complex in the process (Konneke et al., 2005; Treusch et al., 2005; Nicol and Schleper, 2006). Hydroxylamine has been suggested to be the first product of ammonia oxidation (Vajjala et al., 2013), but further conversion to nitrite is performed in an unknown process, as no homolog of the bacterial hydroxylamine dehydrogenase has been found in the genomes of AOA. However, nitric oxide (NO) has been suggested to be involved in the process, because NO production and re-consumption have been observed (Martens-Habbena et al., 2015; Kozłowski et al., 2016b) and the NO scavenger PTIO was shown to inhibit AOA at very low concentrations (Yan et al., 2012; Shen et al., 2013; Martens-Habbena et al., 2015).

Ammonia oxidation by AOA has also been documented to occur in hot springs. With the help of molecular marker genes, diverse ammonia oxidizing thaumarchaea were found in hot terrestrial and marine environments (Reigstad et al., 2008; Zhang et al., 2008; Jiang et al., 2010; Cole et al., 2013). Furthermore, *in situ* nitrification activities up to 84°C were measured using isotopic techniques in AOA-containing habitats in Iceland (Reigstad et al., 2008), in Yellowstone National park (Dodsworth et al., 2011), in a Japanese geothermal water stream (Nishizawa et al., 2016), and in enrichment cultures from Tengchong Geothermal Field in China (Li et al., 2015). In addition, an enrichment of an ammonia oxidizing thaumarchaeon, *Ca. Nitrosocaldus yellowstonensis*, from a hot spring in Yellowstone National park grew at temperatures of up to 74°C confirming that archaeal ammonia oxidation, different from bacterial ammonia oxidation, indeed occurs at high temperatures (de la Torre et al., 2008).

A remarkable diversity of archaea with different metabolisms has been described from terrestrial and marine hot springs and hyperthermophilic organisms tend to be found at the base of almost all lineages of archaea. Therefore, it has often been proposed that the ancestor of all archaea and the ancestors of mesophilic lineages of archaea were hyperthermophilic or at least thermophiles (Barns et al., 1996; Lopez-Garcia et al., 2004; Gribaldo and Brochier-Armanet, 2006; Brochier-Armanet et al., 2012; Eme et al., 2013). This has been supported by sequence analyses which suggested a parallel adaptation from hot to moderate temperatures in several lineages of archaea (Groussin and Gouy, 2011; Williams et al., 2017). Indeed, the thermophilic strain *Ca. Nitrosocaldus yellowstonensis* emerged in 16S rRNA phylogeny as a sister group of all known mesophilic AOA (de la Torre et al., 2008) indicating that AOA also evolved in hot environments.

However, until now, no genome of an obligate thermophilic AOA was available to analyze in depth the phylogenetic placement of thermophilic AOA and to investigate adaptive features of these ammonia oxidizers that have a pivotal position in understanding the evolution of Thaumarchaeota.

In this work we present the physiology and first genome of an extremely thermophilic AOA of the *Nitrosocaldus* lineage that we cultivated from a hot spring in Southern Italy. We analyze its phylogeny based on ribosomal proteins and reconstruct metabolic and genomic features and adaptations.

MATERIALS AND METHODS

Sampling and Enrichment

About 500 mL of mud were sampled from a terrestrial hot spring on the Italian island Ischia at “Terme di Cavascuro” and stored at 4°C until it arrived at the laboratory in Vienna (after about one week, October 2013). A temperature of 77°C and pH 7–8 were measured *in situ* using a portable thermometer (HANNA HI935005) and pH stripes. Initial enrichments (20 mL final volume) were set up in 120 mL serum flasks (two times autoclaved with MilliQ water) containing 14 mL of autoclaved freshwater medium (FWM) (de la Torre et al., 2008; Tourna et al., 2011) amended with non-chelated trace element solution (MTE) (Konneke et al., 2005), vitamin solution, FeNaEDTA (20 µL each), 2 mM bicarbonate, 1 mM ammonium chloride and 2 mL of 0.2 µm filtrated hot spring water. Serum flasks were inoculated with 4 mL (20%) of sampled mud, sealed with gray rubber stoppers (6x boiled and autoclaved) and incubated aerobically at 78°C while rotating at 100 rpm.

Enrichment Strategies

The temperature was changed after one week to 70°C and medium amendment with filtrated hot spring water was discontinued after four transfers as there was no effect observed. Growth was monitored by microscopy, nitrite production and ammonia consumption. Cultures were transferred at a nitrite concentration of ~700 µM and in case ammonia was depleted before the desired nitrite concentration was reached, cultures were fed with 1 mM ammonium chloride. For a preliminary

growth temperature test, cultures were incubated at 55, 60, 65, 70, 72, 74, 76, and 80°C, with cultures at 70°C showing the highest nitrite production rate. The antibiotics Sulfamethazine, Rifampicin and Novobiocin (100 µg mL⁻¹) were used alone and in combination with pyruvate, glyoxylate and oxalacetate (0.1 mM), but with little effect on enrichment. The use of the mentioned organic acids or N,N'-Dimethylthiourea (DMTU, all 0.1 mM) as ROS scavengers increased the abundance of heterotrophic bacteria and even reduced nitrite production. Filtration of enrichments with 1.2 µm filters had no or even detrimental effect on AOA abundance when done repeatedly and 0.45 µm filtration sterilized the cultures.

As nitrite production rate increased over time, inoculum size was decreased from 20 to 10% and finally to 5%. Omitting vitamin solution from the medium led to an increase in nitrite production rate and AOA abundance. Most crucial for increasing thaumarchaeal abundance was keeping an exact timing on passaging cultures in late exponential phase (after 4 days) and setting up multiple cultures from which the best (based on microscopic observations) was used for further propagation. Based on cell counts maximal enrichments of up to 92% were achieved (see Results for more details).

Cultivation

Cultures are routinely grown at 68°C using 5% inoculum in 20 mL FWM amended with MTE and FeNaEDTA solutions, 1 mM NH₄Cl, 2 mM NaHCO₃ and are transferred every 4 days once they reach about 700 µM NO₂⁻. 1 mM NH₄Cl was substituted by 0.5 mM urea to confirm urease activity, while the rest of the medium components and conditions remained unchanged.

Inhibition Test with 2-phenyl-4, 4, 5, 5-tetramethylimidazole-1-oxyl 3-oxide (PTIO)

Cultures were grown in 20 mL FWM with 5% inoculum under standard conditions and different amounts of an aqueous PTIO stock solution were added when early to mid-exponential growth phase was reached (ca. 63 h after inoculation; 20, 100, and 400 µM final PTIO concentration). Ammonia oxidation ceased immediately at all applied PTIO concentrations, but cultures with 20 µM PTIO were able to resume nitrite production 72 h after the addition of PTIO.

DNA Extraction

DNA was extracted from cell pellets by bead beating in pre-warmed (65°C) extraction buffer [40.9 g L⁻¹ NaCl, 12.6 g L⁻¹ Na₂SO₃, 0.1 M Tris/HCl, 50 mM EDTA, 1% sodium dodecyl sulfate (SDS)] and phenol/chloroform/isoamylalcohol [25:24:1 (v/v/v), Fisher BioReagents], in a FastPrep-24 (MP) for 30 s. After centrifugation (10 min, 4°C, 16000 g) the aqueous phase was extracted with chloroform/isoamylalcohol [24:1 (v/v)], prior to DNA precipitation with 1 µL of glycogen (20 mg mL⁻¹) and 1 mL of PEG-solution (93.5 g L⁻¹ NaCl, 300 g L⁻¹ polyethylene glycol 6000) overnight at 4°C. Nucleic acid pellets were washed, dried, resuspended in nuclease-free water and stored at -20°C

until further use (adapted from Zhou et al., 1996; Griffiths et al., 2000).

Amplification of *amoA* Gene

Specific primers were designed for the amplification of the *amoA* gene from *Candidatus* N. cavascurensis: ThAOA-277F: CCA TAY GAC TTC ATG ATA GTC and ThAOA-632R: GCA GCC CAT CKR TAN GTC CA (Alves and Schleper, unpublished). They were based on an alignment of *amoA* gene sequences from available genomes, all PCR-amplified sequences related to Nitrosocaldus identified in GenBank through BLAST searches, and several reference sequences from the orders *Nitrososphaerales*, *Nitrosopumilales*, and *Nitrosotaleales*. ThAOA-277F targets a gene region conserved only among *Ca. Nitrosocaldus* spp., and primer ThAOA-632R was a modification of CrenamoA616r(48x) (Nicol et al., 2008) to match the *amoA* of *Nitrosocaldus* spp.; ThAOA-632R contains additional degenerated bases that account for the most common substitutions among all *amoA* genes, since only two sequences containing this region were available for *Ca. Nitrosocaldus* spp.

PCR conditions were 95°C for 10 min as initialization, followed by 35 cycles of 30 s denaturing at 95°C, 45 s primer annealing at 55°C, and elongation at 72°C for 45 s, finishing with 10 min final elongation at 72°C.

Quantitative PCR

Archaeal and bacterial 16S rRNA genes were quantified in triplicate 20 µL reactions containing 10 µL GoTaq[®] qPCR Master Mix 2x (Promega), 0.2 mg mL⁻¹ BSA and 1 µM of the appropriate primers: Arch931F [5'-AGG AAT TGG CGG GGG AGC A-3' (Jackson et al., 2001)] and Arch1100R [5'-BGG GTC TCG CTC GTT RCC-3' (Ovreas et al., 1997)] for the archaeal 16S rRNA gene, P1 (5'-CCT ACG GGA GGC AGC AG-3') and P2 (5'-ATT ACC GCG GCT GCT GG-3') (Muyzer et al., 1993) for the bacterial 16S rRNA gene. Reactions were performed in a realpex² (Mastercycler ep gradient S, Eppendorf AG) with the following cycling conditions: 95°C for 2 min, followed by 40 cycles of 30 s denaturing at 95°C, 30 s joint annealing-extension at 60°C, and extension with fluorescence measurement at 60°C for 30 s. Specificity of qPCR products was confirmed by melting curve analysis. Standards were prepared by cloning 16S rRNA genes from *Ca. N. cavascurensis* and *Escherichia coli* into pGEM[®]-T Easy Vectors (Promega). These were amplified using M13 primers and concentration was determined with Qubit[™] dsDNA BR Assay Kit (Thermo Fisher Scientific) before preparing serial dilutions. The efficiencies of archaeal and bacterial 16S rRNA standards were 95 and 99%, respectively.

Fluorescence *in Situ* Hybridization

2 mL samples were centrifuged at 4°C, 16000 g for 40 min, washed in PBS-buffer and fixed with 4% paraformaldehyde for 3 h using standard protocols (Amann et al., 1990). Cells were washed two times with 1 mL PBS and finally resuspended in 200 µL of 1:1 PBS:ethanol mix before storing them at -20°C. After dehydration in ethanol cells were hybridized overnight in hybridization buffer with 20% formamide using probes Eub338 (Fluos) 5'-GCT GCC TCC CGT AGG AGT-3' (Amann et al.,

1990) and Arch915 (dCy3) 5'-GTG CTC CCC CGC CAA TTC CT-3' (Stahl and Amann, 1991).

Scanning Electron Microscopy

For scanning electron microscopy, late exponential cells were harvested from 40 mL of culture by centrifugation (16000 g, 4°C, 30 min). The cells were resuspended and washed three times (0.02 mM sodium cacodylate) prior to prefixation (0.5% glutaraldehyde in 0.02 mM sodium cacodylate) overnight, after which the concentration of glutaraldehyde was increased to 2.5% for 2 h. Fixed cells were spotted onto 0.01%-poly-L-lysine coated glass slides (5 mm) and allowed to sediment for 15 min. Samples were dehydrated using an ethanol (30–100%, 15 min each) and acetone (2 × 100%) graduated series and dried using HMDS (2 × 50% HMDS in acetone, 2 × 100%). The slides were subsequently placed on conductive stubs, gold coated for 30 s (JEOL JFC-2300HR) and analyzed (JEOL JSM-IT200).

Genome Assembly

DNA was prepared from 480 mL of culture using standard procedures and sequenced using a PacBio Sequel sequencer at the VBCF (Vienna BioCenter Core Facilities GmbH) with a SMRT Cell 1Mv2 and Sequel Sequencing Kit 2.1. Insert size was 10 kb. Around 500000 reads were obtained with an average size of ~5 kb (N50: 6866 nt, maximal read length: 84504 nt).

The obtained PacBio reads were assembled using the CANU program (version 1.4, parameters “genomeSize = 20 m, corMhapSensitivity = normal, corOutCoverage = 1000, errorRate = 0.013”) (Koren et al., 2017), and then “polished” with the arrow program from the SMRT analysis software (Pacific Biosciences, United States). The *Ca. N. cavascurensis* genome consisted of two contiguous, overlapping contigs of 1580543 kb and 15533 kb, respectively. The resulting assembly was compared to a previous version of the genome bin we had obtained by using IDBA-UD and Newbler for assembly, and differential coverage binning on ~150 nt reads from five runs of IonTorrent sequencing (use of GC% and reads coverage from IonTorrent PGM reads) (Albertsen et al., 2013; Nurk et al., 2013). The new assembly confirmed the previous version as all nine contigs from the previous aligned with MUMMER to the largest of the two contigs, except for a ca. 21 kb region we had not manually selected using differential coverage binning because of a variable read coverage between runs. Interestingly, this region belongs to an integrated conjugative element (ICE1, see results) that might perhaps be excised occasionally. A repeated region between both extremities of the longest contig, and the 2nd small contig obtained was found, and analyzed using the nucmer program from the MUMMER package (see Supplementary Figure 1) (Kurtz et al., 2004). This region was merged using nucmer results, and the longest contig obtained was “circularized” using the information of the nucmer program, as sequence information from the 2nd contig was nearly identical (>99.5%) to that of the extremities of the long contig. It coincided with a repeat-rich adhesin (Supplementary Figure 1). The origin of replication was predicted using the Ori-Finder 2 webserver (Luo et al., 2014). By analogy with *Nitrosopumilus maritimus*, it was placed after the last annotated genomic object, before the ORB repeats and the

cdc gene. The annotated genome sequence has been deposited to the European Nucleotide Archive (ENA) with the study accession number: PRJEB24312.

Genome Dataset

Thaumarchaeota and Aigarchaeota complete or nearly complete genomes were downloaded from the NCBI, Genbank database, or IMG database when not available in Genbank. We included the genomes of 27 Thaumarchaeota: Thaumarchaeota archaeon BS4 (IMG_2519899514); Thaumarchaeota group 1.1c (bin Fn1) (IMG_2558309099); Thaumarchaeota NESA-775 (Offre and Schleper, unpublished); Thaumarchaeota archaeon DS1 (IMG_2263082000); *Cenarchaeum symbiosum* A (DP000238); Thaumarchaeota archaeon CSP1 (LDXL00000000); *Nitrosopumilus adriaticus* NF5 (CP011070); *Nitrosocosmicus arcticus* Kfb (Alves and Schleper unpublished); *Nitrosopelagicus brevis* CN25 (CP007026); *Nitrosotenuis chungbukensis* MY2 (AVSQ00000000); *Nitrosotalea devanaterterra* Nd1 (LN890280); *Nitrososphaera evergladensis* SR1 (CP007174); *Nitrosocosmicus exaquare* G61 (CP017922); *Nitrososphaera gargensis* Ga9.2 (CP002408); *Nitrosoarchaeum koreensis* MY1 (GCF_000220175); *Nitrosopumilus salaria* BD31 (GCF_000242875); *Nitrosopumilus koreensis* AR1 (CP003842); *Nitrosoarchaeum limnia* BG20 (GCF_000241145); *Nitrosoarchaeum limnia* SFB1 (CM001158); *Nitrosopumilus maritimus* SCM1 (CP000866); *Nitrosocosmicus oleophilus* MY3 (CP012850); *Nitrosopumilus piranensis* D3C (CP010868); *Nitrosopumilus sediminis* AR2 (CP003843); *Nitrosotenuis uzonensis* N4 (GCF_000723185); *Nitrososphaera viennensis* EN76 (CP007536); *Nitrosotenuis cloacae* SAT1 (CP011097); *Nitrosocaldus cavascurensis* SCU2 (this paper).

Two Aigarchaeota genomes were included: *Calditenuis aerorheumensis* (IMG_2545555825) and *Caldiarchaeum subterraneum* (BA000048). Additionally, we selected 11 Crenarchaeota genomes from Genbank to serve as an outgroup in our analysis: *Sulfolobus solfataricus* P2 (AE006641); *Pyrobaculum aerophilum* IM2 (AE009441); *Hyperthermus butylicus* DSM 5456 (CP000493); *Thermofilum pendens* Hrk 5 (CP000505); *Acidilobus saccharovorans* 345-15 (CP001742); *Desulfurococcus fermentans* DSM 16532 (CP003321); *Caldisphaera lagunensis* DSM 15908 (CP003378); *Fervidicoccus fontis* Kam940 (CP003423); *Metallosphaera sedula* DSM 5348 (CP000682); *Aeropyrum pernix* K1 (BA000002) and *Thermoproteus tenax* Kra 1 (FN869859).

Genome Annotation

The first annotation of the *Ca. N. cavascurensis* genome was obtained by the automatic annotation pipeline MicroScope (Vallenet et al., 2009, 2017; Medigue et al., 2017). Annotations from *Nitrososphaera viennensis* EN76 were carried over when aligned sequences were 35% identical (or 30% in syntenic regions) and covered 70% of the sequence. All subsequent manual annotations were held on the platform. Putative transporters were classified by screening against the Transporter Classification Database (Saier et al., 2014). To assist the annotation process, several sets of proteins of interest from Kerou et al. (2016b) (EPS, metabolism), Offre et al. (2014) (transporters), Makarova et al. (2016) (archaeal pili and flagella) and ribosomal proteins were screened specifically within our set of genomes using HMMER

and HMM protein profiles obtained from databases (TIGRFAM release 15, PFAM release 28, SUPERFAMILY version 1.75), or built from arCOG families (Gough et al., 2001; Haft et al., 2003; Finn et al., 2014; Makarova et al., 2015). In the former case, the trusted cutoff “—cut_tc” was used if defined in the profiles. In the latter case, for each family of interest, sequences were extracted from the arCOG2014 database, aligned using MAFFT (“linsi” algorithm) (Katoh and Standley, 2014), and automatically trimmed at both extremities of conserved sequence blocks using a home-made script relying on blocks built by BMGE v1.12 (BLOSUM 40) (Crisuolo and Gribaldo, 2010). HMMER was then used to screen genomes for different sets of specialized families (Eddy, 2011).

In order to visualize and compare the genetic organization of sets of genes (AMO, Urease – Ure, type IV pili), we built HMM profiles (as above except alignments were manually curated for AMO and Ure), based on the corresponding arCOG families and integrated the respective sets of profiles in the MacSyFinder framework that enables the detection of sets of co-localized genes in the genome (parameters `inter_gene_max_space = 10`, `min_nb_mandatory_genes = 1` for AMO and Ure, `inter_gene_max_space = 10`, `min_nb_mandatory_genes = 4` for type IV pili) (Abby et al., 2014). MacSyFinder was then run with the models and profiles on the genome dataset, and the resulting JSON files were used for visualization in the MacSyView web-browser application (Abby et al., 2014). The SVG files generated by MacSyView from the predicted operons were downloaded to create figures.

Reference Phylogeny

Ribosomal proteins were identified in genomes using a set of 73 HMM profiles built for 73 arCOG families of ribosomal proteins (see Genome Annotation). When multiple hits were obtained in a genome for a ribosomal protein, the best was selected. The sequences were extracted and each family was aligned using the MAFFT program (“linsi” algorithm) (Katoh and Standley, 2014). Alignments were curated with BMGE (version 1.12, default parameters and BLOSUM 45) (Crisuolo and Gribaldo, 2010) and then checked manually. In several cases, sequences identified as part of a given ribosomal protein family did not align well and were excluded from the analysis. In the end, 59 families with more than 35 genomes represented were selected to build a phylogeny. The corresponding family alignments were concatenated, and a tree was built using the IQ-Tree program version 1.5.5 (Nguyen et al., 2015) (model test, partitioned analysis with the best model selected per gene, 1000 UF-Boot and 1000 aLRT replicates). A tree drawing was obtained with the Scriptree program (Chevenet et al., 2010), and then modified with Inkscape.

Protein Families Reconstruction and Phyletic Patterns Analysis

Homologous protein families were built for the set of genomes selected. A blast similarity search was performed with all sequences against themselves, and the results used as input for the Silix and Hifix programs to cluster the sets of similar sequences into families (Miele et al., 2011, 2012). For sequences to be

clustered in a same Silix family, they had to share at least 30% of identity and the blast alignment cover at least 70% of the two sequences lengths. The distribution of protein families of interest was analyzed across the genome dataset.

Phylogenetic Analyses

Sequences from Silix families of interest were extracted in fasta files, and a similarity search was performed against a large database of sequences which consisted of 5750 bacterial and archaeal genomes (NCBI Refseq, last accessed in November 2016), and the genomes from our genome dataset. Sequences with a hit having an *e*-value below 10^{-10} were extracted, and the dataset was then de-replicated to remove identical sequences. An alignment was then obtained for the family using MAFFT (“linsi” algorithm), filtered with BMGE (BLOSUM 30) and a phylogenetic tree was reconstructed by maximum-likelihood using the IQ-Tree program (version 1.5.5, best evolutionary model selected, 1000 UF-Boot and 1000 aLRT replicates). A tree drawing was obtained with the Scriptree program (Chevenet et al., 2010), and modified with Inkscape.

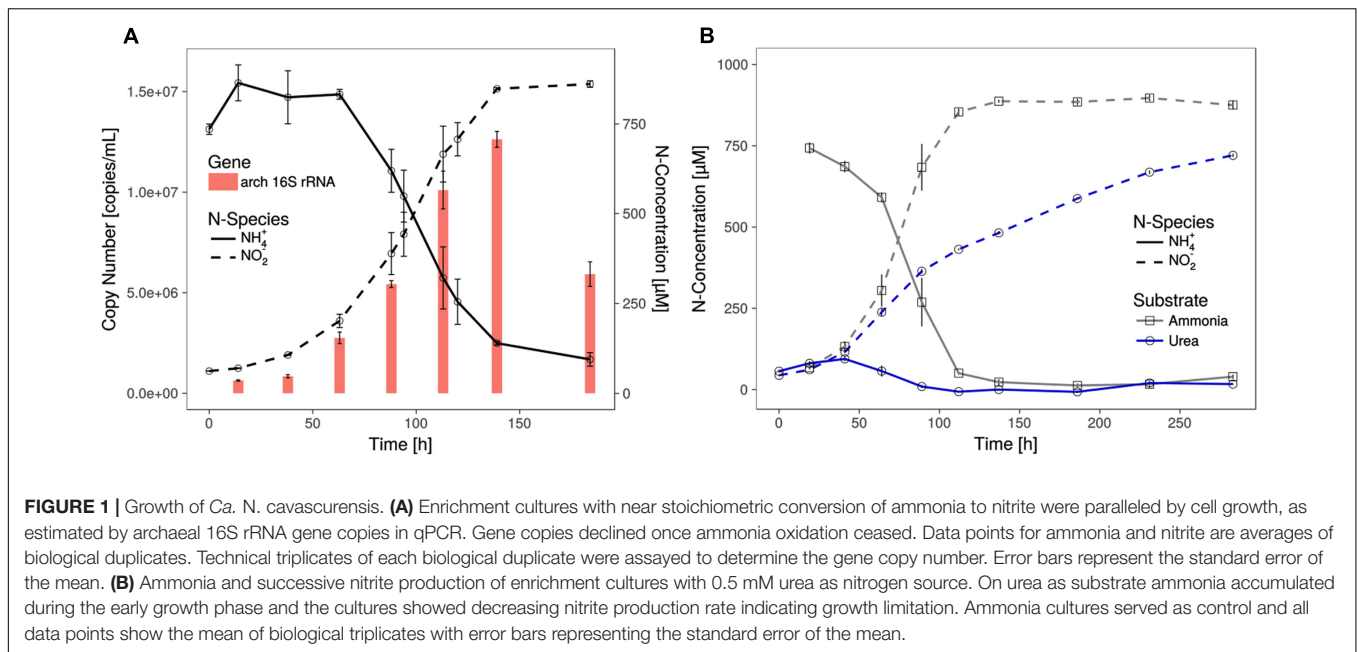
Identification and Annotation of Integrated MGE

Integrated mobile genetic elements (MGE) were identified as described previously (Kazlauskas et al., 2017). Briefly, the integrated MGE were identified by searching the *Ca. N. cavascurensis* genome for the presence of the hallmark genes specific to mobile genetic elements, such as integrases, casposases, VirB4 ATPases, large subunit of the terminase, major capsid proteins of archaeal viruses, etc. In the case of a positive hit, the corresponding genomic neighborhoods were analyzed further. IS elements were searched for using ISfinder (Siguier et al., 2006). The precise borders of integration were defined based on the presence of direct repeats corresponding to attachment sites. The repeats were searched for using Unipro UGENE (Okonechnikov et al., 2012). Genes of integrated MGE were annotated based on PSI-BLAST searches (Altschul et al., 1997) against the non-redundant protein database at NCBI and HHpred searches against CDD, Pfam, SCOPe and PDB databases (Soding, 2005).

RESULTS AND DISCUSSION

Enrichment of *Ca. Nitrosocaldus cavascurensis*

An ammonia oxidizing enrichment culture was obtained from thermal water outflow of the *Terme di Cavascura* on the Italian island Ischia. After repeated transfers into artificial freshwater medium supplemented with 1 mM ammonium and 2 mM bicarbonate, highly enriched cultures were obtained that exhibited almost stoichiometric conversion of ammonia into nitrite within 4–5 days (**Figure 1A**). A single phylotype of AOA but no AOB or Nitrospira/comammox was identified in the enrichment via amplification and sequencing of *amoA* and 16S rRNA gene fragments (data not shown) and the presence of a single AOA phylotype was confirmed by metagenomic



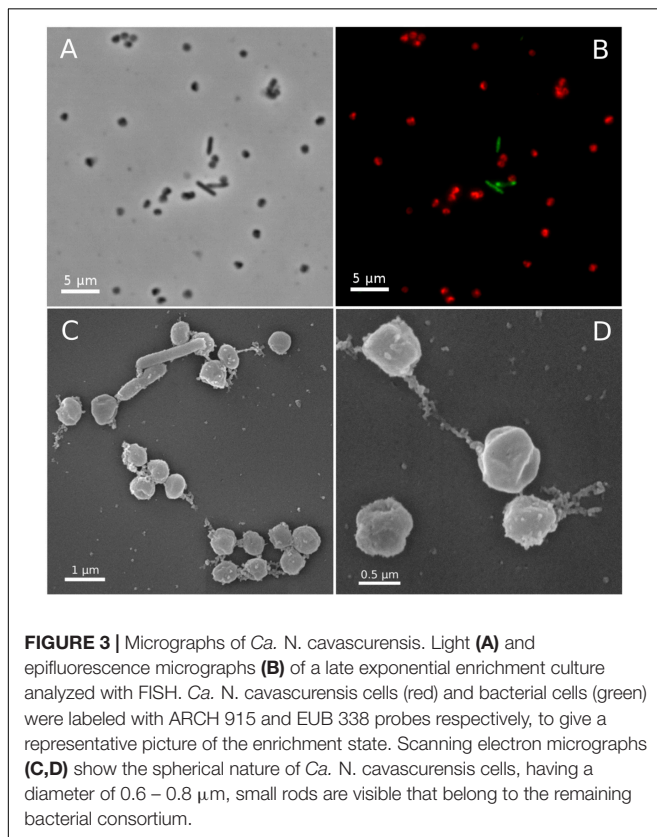
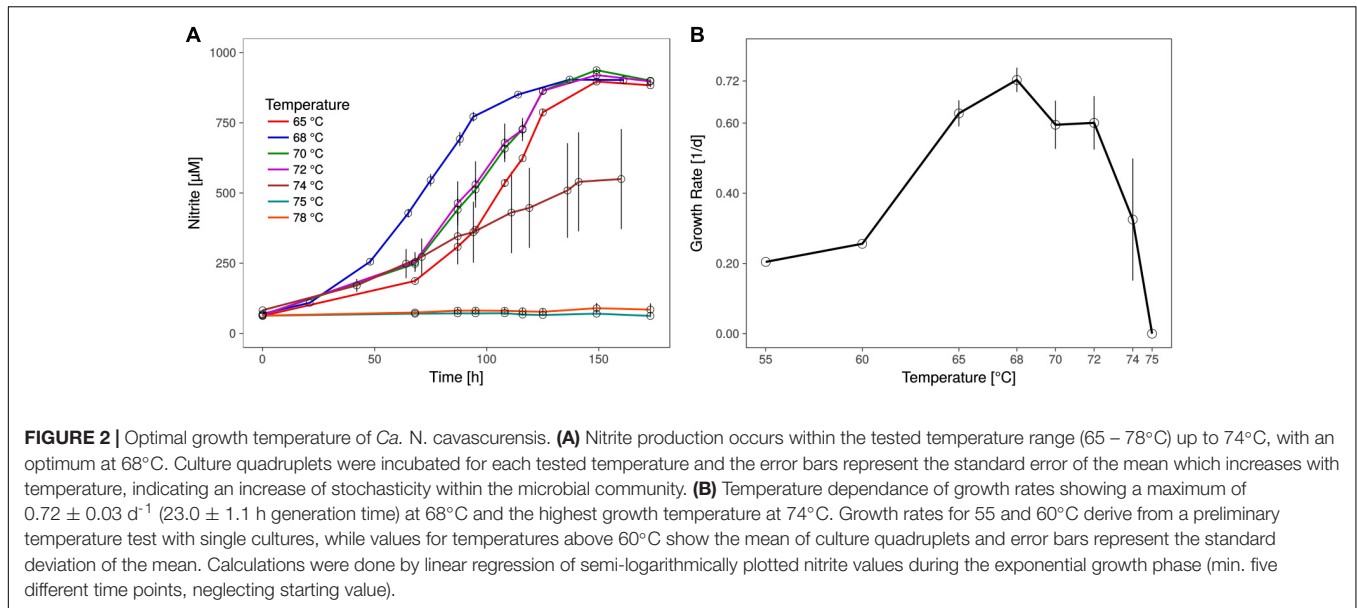
sequencing (see below). Its closest relative in the 16S rRNA database was fosmid 45H12 isolated from a gold mine metagenome with 99% identity and the next closest cultivated relative was *Ca. Nitrosocaldus yellowstonensis* with 96% identity (Nunoura et al., 2005; de la Torre et al., 2008). The AOA has been propagated for 4 years within a stable enrichment culture and its relative enrichment usually ranged between 75 and 90% based on qPCR or cell counts respectively. Different strategies were tested in order to obtain a better enrichment (heat treatment, antibiotics, dilution to extinction, filtration) which resulted in a maximal enrichment of 92% (based on cell counts) but not in a pure culture. In high throughput sequencing analyses using general prokaryotic primers to amplify the V2/V4 region of the 16S rRNA gene we identified phylotypes of the *Deinococcus/Thermus* group comprising up to 97% of the bacterial contaminating sequences (with 96% identity to sequences from the *Thermus* genus) and to minor extent sequences related to the Chloroflexi and Armatimonadetes (not shown). Although Aigarchaeota had been detected in earlier enrichments, the AOA was the only archaeon discovered in the high-level enrichments used in this study. *Ca. N. cavascurens* does not encode a catalase, and in the light of current knowledge on the auxiliary role of ROS scavengers in the growth of AOA (Kim et al., 2016), it is possible that a ROS detoxification mechanism is provided by the remaining bacteria in the enrichment culture, although supplementation with α -keto acids did not have a positive effect as it had on other AOA enrichments (see Materials and Methods).

The shortest generation time of the thermophilic AOA was approximately 23 h (± 1.1), as measured by nitrite production rates and in qPCR, which is comparable to the generation time of *Ca. N. yellowstonensis* (de la Torre et al., 2008). It grew in a temperature range from about 55 to 74°C with an apparent (but relatively wide) temperature optimum around 68°C (Figure 2).

In fluorescence *in situ* hybridizations (FISH) with archaea- and bacteria-specific probes, all coccoid-shaped cells of less than 1 μm in diameter were assigned to the AOA, while all shorter and longer rod-shaped morphotypes were clearly assigned to bacteria (Figures 3A,B). Scanning electron microscopy revealed the typical spherical and irregular shape of cocci, as seen, e.g., for the ammonia oxidizing archaeon *Nitrososphaera viennensis* (Tournai et al., 2011) or hyperthermophilic and halophilic coccoid archaea strains with a diameter of around 0.7 μm (Figures 3C,D). Based on its relationship with *Ca. Nitrosocaldus yellowstonensis* (de la Torre et al., 2008), and the location it was sampled from (*Terme di cavascura*, Italy) we named this organism provisionally *Candidatus Nitrosocaldus cavascurens*.

***Ca. Nitrosocaldus cavascurens* Represents a Deeply Branching Lineage of AOA**

We gathered all complete or nearly complete 27 genome sequences available for Thaumarchaeota. These included 23 genomes of closely related cultivated or uncultivated AOA all harboring *amo* genes, and four genomes obtained from metagenomes that do not have genes for ammonia oxidation. Among the latter were two assembled genomes from moderate environments (Fn1 and NESA-775) and two from hot environments (BS4 and DS1) (Beam et al., 2014; Lin et al., 2015). In addition, two Aigarchaeota genomes, and a selection of representative genomes of Crenarchaeota (11) were included to serve as outgroups (Petitjean et al., 2014; Raymann et al., 2015; Adam et al., 2017; Williams et al., 2017). This resulted in a 40 genomes dataset (see Materials and Methods). A maximum likelihood analysis based on 59 concatenated ribosomal proteins (in one copy in at least 35 of the 40 genomes) resulted in a highly supported tree: UF-boot and aLRT supports were maximal or



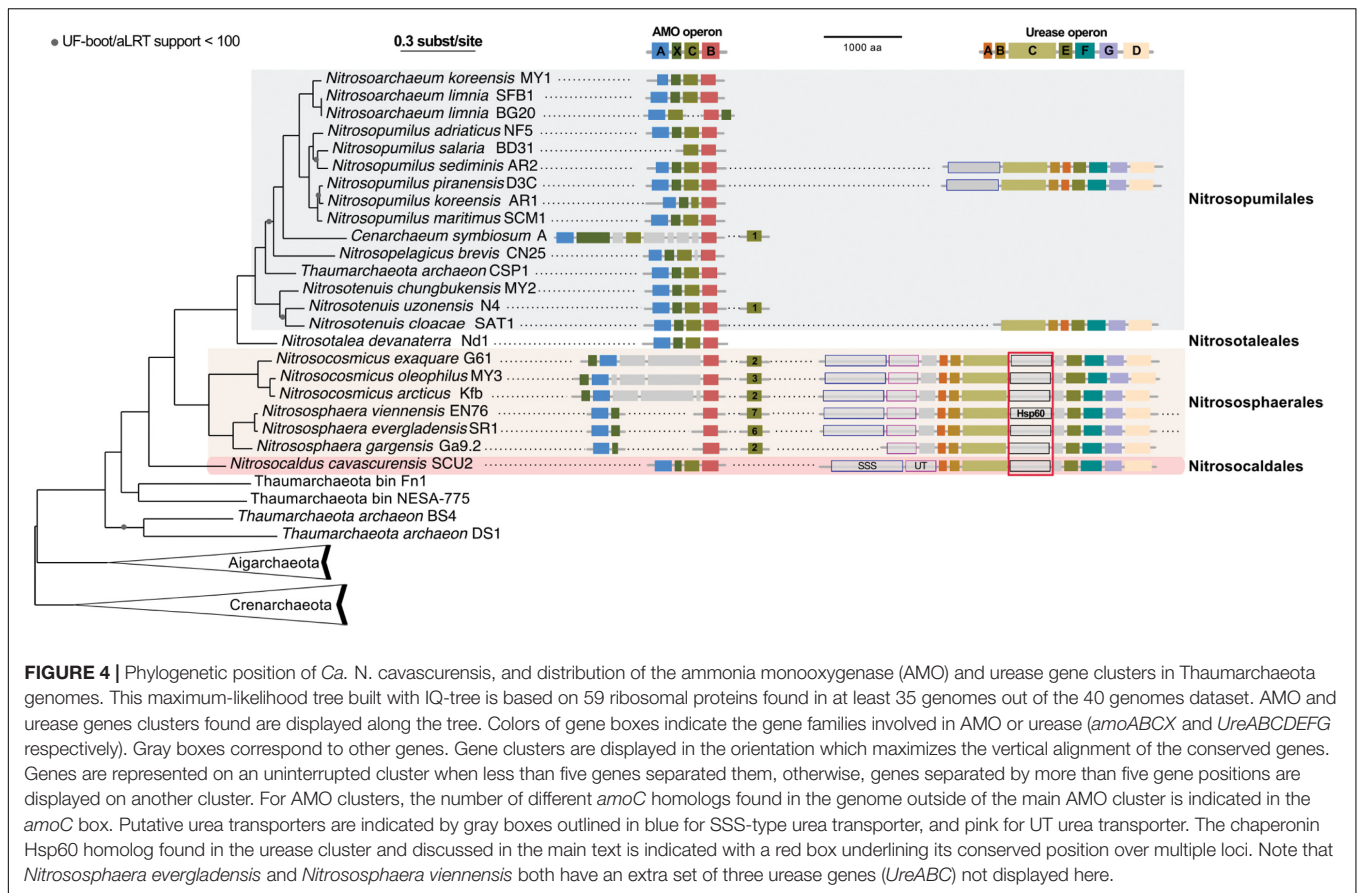
above 99, except for two nodes (>90) in the Thaumarchaeota subtree and one node (>70) in the Crenarchaeota subtree (**Figure 4**). The tree confirms the monophyly of Thaumarchaeota and Nitrososphaeria (AOA), and the deeply branching position of *Ca. N. cavascurensis* as a sister-group of all other (mesophilic) AOA, as initially indicated with a single gene (16S rRNA

phylogeny of *Ca. Nitrosocaldus yellowstonensis* (de la Torre et al., 2008).

Genome and Energy Metabolism of *Ca. Nitrosocaldus cavascurensis*

The genome of *Ca. N. cavascurensis* contains 1.58 Mbp, with 1748 predicted coding sequences, one 16S/23S rRNA operon and 29 tRNA genes. It has a G + C content of 41.6%.

Similar to the genomes of most marine strains (*Nitrosopumilales*) and to *Ca. Nitrosocaldus yellowstonensis*, but different from those of the terrestrial organisms (*Nitrososphaerales*) it encodes all putative subunits of the ammonia monooxygenase in a single gene cluster of the order *amoA*, *amoX*, *amoC*, *amoB* (**Figure 4**) indicating that this might represent an ancestral gene order. Different from several other AOA it has a single copy of *amoC*. The genome contains a cluster of genes for the degradation of urea, including the urease subunits and two urea transporters (**Figures 4, 5**) with a similar structure to the urease locus of *Nitrososphaerales*. Accordingly, growth on urea, albeit slower than on ammonia, could be demonstrated (**Figure 1B**). Urease loci were so far found in all *Nitrososphaerales* genomes, and in some *Nitrosopumilales* (Hallam et al., 2006; Park et al., 2012; Spang et al., 2012; Bayer et al., 2016). No urease cluster could be found in non-AOA Thaumarchaeota, nor in Aigarchaeota. In Crenarchaeota, only one *Sulfolobales* genome harbored a urease (*Metallosphaera sedula* DSM 5348). A specific protein family of a putative chaperonin (Hsp60 homolog, arCOG01257) was found to be conserved within the urease loci of *Ca. N. cavascurensis* and the *Nitrososphaerales* (“soil-group” of AOA) (**Figure 4**) as also observed in some bacteria (e.g., in *Haemophilus influenzae*). Unlike the conserved “thermosome” Hsp60 (also part of arCOG01257), this particular Hsp60 homolog was not found in any *Nitrosopumilales* genome. Given the conservation of the urease loci between *Ca. Nitrosocaldus* and *Nitrososphaerales*, it



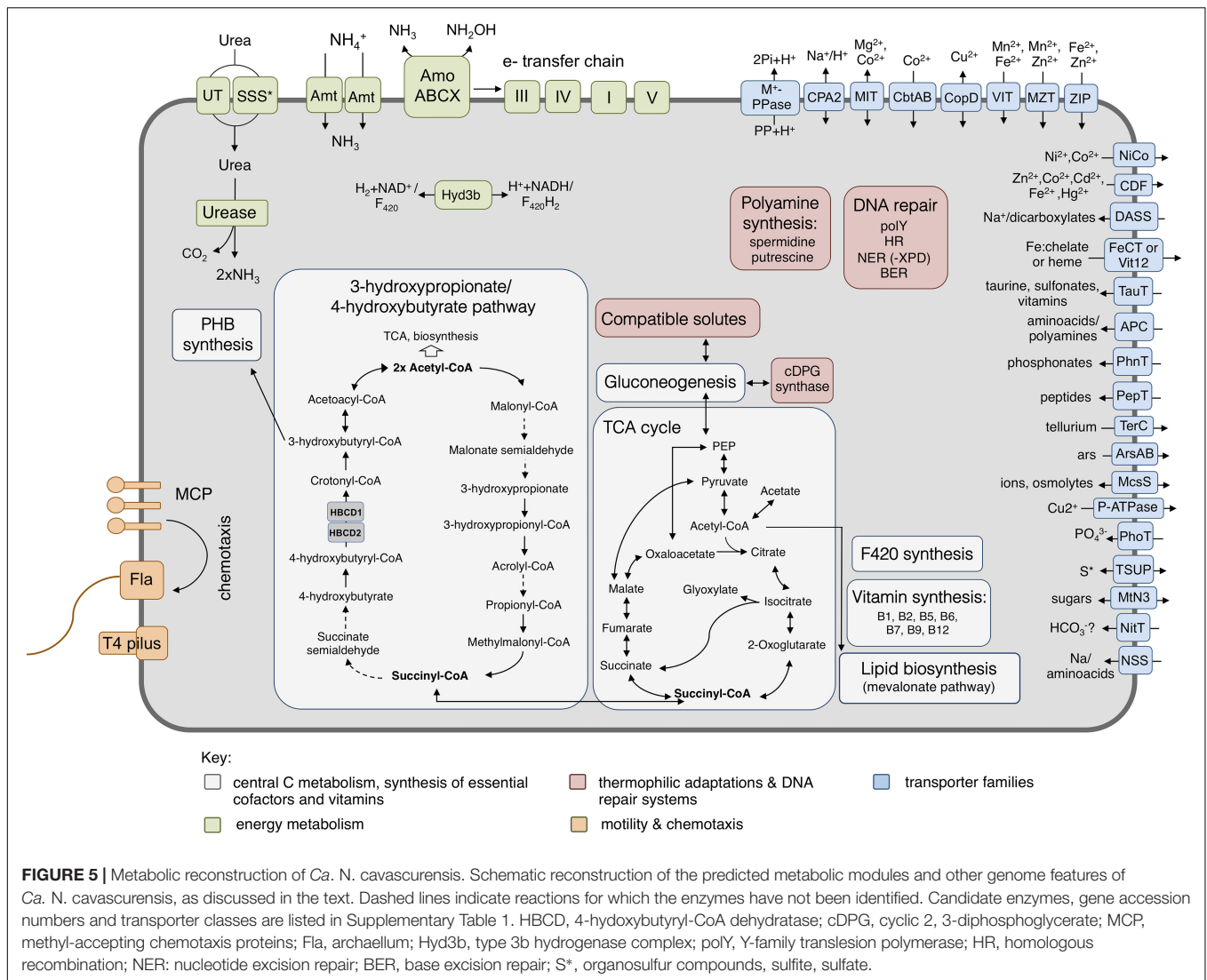
is possible that these genes have a common origin, and were acquired *en bloc* from bacteria by the ancestor of AOA. It is likely that the Hsp60 homolog is involved in stabilization of the urease, as demonstrated in the deltaproteobacterium *Helicobacter pylori* (Evans et al., 1992).

Intriguingly, a gene encoding a nitrite reductase that is present in all genomes of AOA (except the sponge symbiont, *Ca. Cenarchaeum symbiosum*, Bartossek et al., 2010) and which is highly transcribed during ammonia oxidation in *Nitrososphaera viennensis* (Kerou et al., 2016b) and in meta-transcriptomic datasets (Shi et al., 2011) is missing from the genome. Nitrite reductases have been postulated to be involved in ammonia oxidation by providing nitric oxide (NO) for the oxidation step to nitrite (Kozłowski et al., 2016a). We therefore tested, if the organism was affected by the NO-scavenger PTIO which inhibited at low concentrations all AOA tested so far (Martens-Habbena et al., 2015; Kozłowski et al., 2016a). Ammonia oxidation of *Ca. N. cavascurensis* was fully inhibited at concentrations as low as 20 μ M, similar to those affecting *Nitrososphaera viennensis* and *Nitrosopumilus maritimus* (Figure 6). This indicates that NO is also an important intermediate in the ammonia oxidation of this thermophile and it might be produced by an unknown nitrite reductase or perhaps by an unidentified hydroxylamine dehydrogenase (HAO), as recently shown for the HAO of the ammonia oxidizing bacterium *Nitrosomonas europaea* (Caranto and Lancaster,

2017). Alternatively, NO might be supplied through the activity of other organisms in the enrichment culture. Indeed we found a *nirK* gene in the genome of the *Deinococcus* species of our enrichment culture. The sensitivity to PTIO reinforces the previously raised hypothesis that NO represents an important intermediate or cofactor in ammonia oxidation in archaea (Schleper and Nicol, 2010; Simon and Klotz, 2013; Kozłowski et al., 2016b). However, the full understanding of ammonia oxidation in AOA and the role of NO awaits further biochemical investigations.

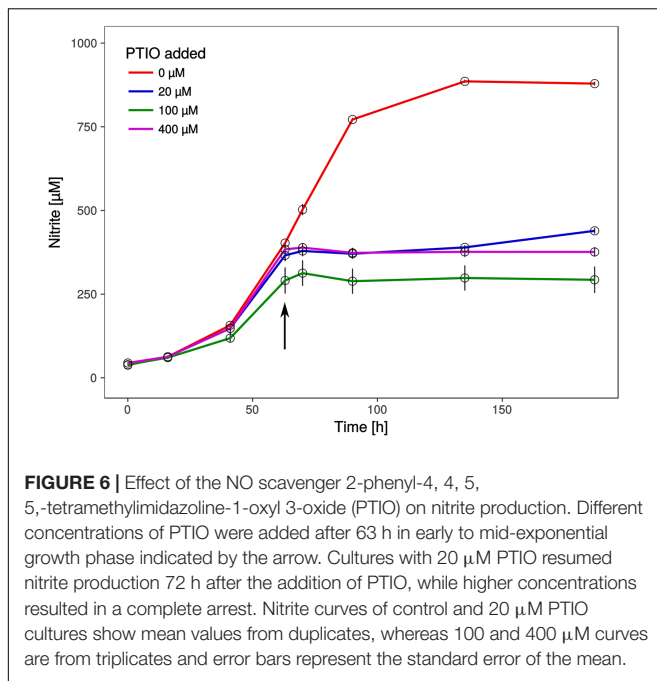
Pathways of central carbon metabolism and -fixation of *Ca. N. cavascurensis* were found to be generally similar to those of other AOA, underlining the strikingly high similarity and conservation of metabolism within the Nitrososphaeria (Figure 5) (Kerou et al., 2016b).

Key enzymes of the 3-hydroxypropionate/4-hydroxybutyrate carbon fixation pathway were identified in the genome of *Ca. N. cavascurensis* (Supplementary Table 1, Könneke et al., 2014). Intriguingly, and similar to *Sulfolobales*, we find two copies of the gene encoding for 4-hydroxybutyryl-CoA dehydratase located in tandem. The first copy (NCAV_0127) is homologous to the other AOA sequences which are more closely related to fermenting Clostridia rather than their crenarchaeal counterparts (“Cren Type-1,” Supplementary Figure 2), while the second copy (NCAV_0126) exhibits high similarity to genes from bacterial candidate division NC10 and



Handelsmanbacteria (49% identity). In a phylogenetic tree of the protein family, it clusters together with the second group of crenarchaeal genes (“Cren Type-2,” Supplementary Figure 2, see Materials and Methods) which lack essential catalytic residues and whose function remains unknown (Ramos-Vera et al., 2011). Könneke et al. (2014) suggested an independent emergence of the cycle in Thaumarchaeota and autotrophic Crenarchaeota based on the unrelatedness of their respective enzymes (Könneke et al., 2014). But the existence of the second gene in this deep-branching lineage rather indicates that both genes could have been present in a common ancestor of Crenarchaeota and Thaumarchaeota and that their CO₂ fixation pathways could have a common origin. No homolog of the experimentally characterized malonic semialdehyde reductase from *Nitrosopumilus maritimus* (Otte et al., 2015), or any other iron-containing alcohol dehydrogenase protein family member, was found in the *Ca. N. cavascurensis* genome. Since this enzyme (Nmar_1110) is so far the only characterized malonic semialdehyde reductase active in mesophilic conditions (Otte

et al., 2015), we searched for homologs of the corresponding (but unrelated) thermophilic enzymes from Crenarchaea (*Metallosphaera sedula*) (Kockelkorn and Fuchs, 2009) and Chloroflexi (Hügler et al., 2002). The best hits of Msed_1993, a member of the 3-hydroxyacyl-CoA dehydrogenase family, were the 3-hydroxybutyryl-CoA dehydrogenase homologs of the thaumarchaeal genomes, as observed already by Könneke et al. (2014). The reduction of malonic semialdehyde in the 3-hydroxypropionate bi-cycle of *Chloroflexus aurantiacus* is carried out by a bifunctional protein (Caur_2614) with an N-terminal short-chain alcohol dehydrogenase domain and a C-terminal aldehyde dehydrogenase domain (Hügler et al., 2002). No homolog of this bifunctional enzyme was identified in the genome of *Ca. N. cavascurensis*, but a number of short-chain alcohol dehydrogenases with low identities (~30%) to Caur_2614 exist (Supplementary Table 1) that could serve as candidates in the absence of a canonical malonic semialdehyde reductase. Whether temperature adaptation leads to non-orthologous gene replacement



we observe in mesophilic AOA remains to be seen upon experimental characterisation of the thermophilic thaumarchaeal enzyme.

A full oxidative TCA cycle is present in *Ca. N. cavascurensis*, including a malic enzyme and a pyruvate/phosphate dikinase connecting the cycle to gluconeogenesis. Additionally, and like most mesophilic AOA, *Ca. N. cavascurensis* encodes a class III polyhydroxyalkanoate synthase (phaEC) allowing for the production of polyhydroxyalkanoates, carbon polyester compounds that form during unbalanced growth and serve as carbon and energy reserves (Poli et al., 2011; Spang et al., 2012) (Figure 5 and Supplementary Table 1).

The presence of a full set of genes encoding for the four subunits (and maturation factors) of a soluble type 3b [NiFe]-hydrogenase, uniquely in *Ca. N. cavascurensis*, indicates the ability to catalyze hydrogen oxidation potentially as part of the energy metabolism. This group of hydrogenases is typically found among thermophilic archaea, is oxygen-tolerant and bidirectional, and can couple oxidation of H_2 to reduction of NAD(P) in order to provide reducing equivalents for biosynthesis, while some have been proposed to have sulfhydrogenase activity (Kanai et al., 2011; Peters et al., 2015; Greening et al., 2016) (and references therein). Although classified by sequence analysis (Sondergaard et al., 2016) and subunit composition as a type 3b hydrogenase, the alpha and delta subunits belong to the arCOG01549 and arCOG02472 families respectively, which contain coenzyme F_{420} -reducing hydrogenase subunits, so far exclusively found in methanogenic archaea. Given the fact that Thaumarchaeota can synthesize this cofactor and encode a number of F_{420} -dependent oxidoreductases with a yet unknown function, it is interesting to speculate whether oxidized F_{420} could also be a potential substrate for the hydrogenase. Expression of

the hydrogenase is likely regulated through a cAMP-activated transcriptional regulator encoded within the hydrogenase gene cluster (Supplementary Table 1).

Putative S-layer subunits were identified in the genome of *Ca. N. cavascurensis* (NCAV_0187 and NCAV_0188, Supplementary Table 1). Although the overall identity among thaumarchaeal S-layer subunits is very low (<30% between different genera), investigations of the genomic neighborhood reveals conserved synteny of this region.

Adaptations to Thermophilic Life

The molecular adaptations that enable survival and the maintenance of cell integrity at high temperatures have been the subject of intense studies since the discovery of thermophilic organisms. The issue of extensive DNA damage occurring at high temperature has led to the study of systems of DNA stabilization and repair in thermophilic and hyperthermophilic archaea. Among them, the reverse gyrase, a type IA DNA topoisomerase shown to stabilize and protect DNA from heat-induced damage, is often (but not always (Brochier-Armanet and Forterre, 2007)) found in thermophiles, and is even considered a hallmark of hyperthermophilic organisms growing optimally above 80°C (Bouthier de la Tour et al., 1990; Forterre, 2002). However, the gene might not always be essential for survival at high temperature in the laboratory (Atomi et al., 2004; Brochier-Armanet and Forterre, 2007). Interestingly, we could not identify a gene encoding for reverse gyrase in the genome of *Ca. N. cavascurensis*. This might either reflect that its growth optimum is at the lower end of extreme thermophiles or that there is a separate evolutionary line of adaptation to thermophily possible without reverse gyrase.

The following DNA repair mechanisms were identified in the genome of *Ca. N. cavascurensis*, in agreement with the general distribution of these systems in thermophiles (for reviews see Rouillon and White, 2011; Grasso and Tell, 2014; Ishino and Narumi, 2015):

(a) Homologous recombination repair (HR): Homologs of RadA and RadB recombinase, Mre11, Rad50, the HerA-NurA helicase/nuclease pair, and Holliday junction resolvase Hjc are encoded in the genome. These genes have been shown to be essential in other archaeal (hyper)thermophiles, leading to hypotheses regarding their putative role in replication and more generally the tight integration of repair, recombination and replication processes in (hyper)thermophilic archaea (Grogan, 2015).

(b) Base excision repair (BER): the machinery responsible for the repair of deaminated bases was identified in the genome, including uracil DNA glycosylases (UDG) and putative apurinic/apyrimidinic lyases. Deletion of UDGs was shown to impair growth of (hyper)thermophilic archaea (Grogan, 2015).

(c) Nucleotide excision repair (NER): Homologs of the putative DNA repair helicase: nuclease pair XPB-Bax1 and nuclease XPF were found in the genome, but repair nuclease XPD could not be identified. XPD is present in all so far analyzed mesophilic AOA, but it is also absent in other thermophiles (Kelman and White, 2005). It should be noted that NER functionality in archaea is still unclear, and deletions of the

respective genes were shown to have no observable phenotype (Grogan, 2015).

(d) Translesion polymerase: A Y-family polymerase with low-fidelity able to perform translesion DNA synthesis is encoded in *Ca. N. cavascurensis*.

Key enzymes of all the above-mentioned systems are also found in mesophilic AOA. Given their extensive study in the crenarchaeal (hyper)thermophiles, it would be interesting to characterize their respective functions and regulation in both mesophilic and thermophilic Thaumarchaeota.

(e) Bacterial-type UvrABC excision repair: In contrast to mesophilic AOA and in agreement with the known distribution of the system among mesophilic archaea and bacteria but its absence in (hyper)thermophiles, *Ca. N. cavascurensis* does not encode homologs of this repair machinery.

Ca. N. cavascurensis could potentially produce the polyamines putrescine and spermidine, which have been shown to bind and stabilize compacted DNA from thermal denaturation, acting synergistically with histone molecules, also present in AOA (Higashibata et al., 2000; Oshima, 2007). Although a putative spermidine synthase is also found in mesophilic AOA, the gene encoding for the previous step, S-adenosylmethionine decarboxylase (NCAV_0959), is only found in *Ca. N. cavascurensis*, located in tandem with the former. The biosynthesis of putrescine (a substrate for spermidine synthase) is unclear, since we could not identify a pyruvoyl-dependent arginine decarboxylase (ADC, catalyzing the first of the two-step biosynthesis of putrescine). However, it was shown that the crenarchaeal arginine decarboxylase evolved from an S-adenosylmethionine decarboxylase enzyme, raising the possibility of a promiscuous enzyme (Giles and Graham, 2008).

The production of thermoprotectant compounds with a role in stabilizing proteins from heat denaturation seems to be a preferred strategy of heat adaptation in *Ca. N. cavascurensis*. Firstly, the presence of mannosyl-3-phosphoglycerate synthase (NCAV_1295) indicates the ability to synthesize this compatible solute, shown to be involved in heat stress response and protection of proteins from heat denaturation in *Thermococcales* (Neves et al., 2005). Homologous genes are also present in mesophilic members of the order *Nitrososphaerales* and other (hyper)thermophiles (Spang et al., 2012; Kerou et al., 2016b). Secondly, only *Ca. N. cavascurensis*, but no other AOA encodes a cyclic 2, 3-diphosphoglycerate (cDPG) synthetase (NCAV_0908), an ATP-dependent enzyme which can synthesize cDPG from 2,3-biphosphoglycerate, an intermediate in gluconeogenesis. High intracellular concentrations of cDPG accumulate in hyperthermophilic methanogens, where it is required for the activity and thermostability of important metabolic enzymes (Shima et al., 1998).

Notable Features of the DNA Replication and Cell Division Systems in *Ca. N. cavascurensis*

Strikingly, only one family B replicative polymerase PolB was identified in the genome of *Ca. N. cavascurensis* (NCAV_1300),

making it the only archaeon known so far to encode a single subunit of the replicative family B polymerase, as in Crenarchaeota multiple paralogs with distinct functions coexist (Makarova et al., 2014). Both subunits of the D-family polymerases PolD present in all other AOA and shown to be responsible for DNA replication in *Thermococcus kodakarensis* (also encoding both polD and polB families) (Cubonova et al., 2013) were absent from the genome, raising intriguing questions about the role of the polB family homolog in mesophilic Thaumarchaeota. Sequence analysis indicated that the *Ca. N. cavascurensis* homolog belongs to the polB1 group present exclusively in the TACK superphylum and shown recently by Yan and colleagues to be responsible for both leading and lagging strand synthesis in the crenarchaeon *Sulfolobus solfataricus* (Yan et al., 2017). However, the activity of *Sulfolobus solfataricus* PolB1 is dependent on the presence and binding of two additional proteins, PBP1 and PBP2, mitigating the strand-displacement activity during lagging strand synthesis and enhancing DNA synthesis and thermal stability of the holoenzyme, respectively. No homologs of these two additional subunits were identified in *Ca. N. cavascurensis*, raising the question of enzymatic thermal stability and efficiency of DNA synthesis on both strands.

Homologs of genes for the Cdv cell division system proteins CdvB (3 paralogs) and CdvC, but not CdvA, were identified in *Ca. N. cavascurensis*. This is surprising given that all three proteins were detected by specific immuno-labeling in the mesophilic AOA *Nitrosopumilus maritimus*, where CdvA, CdvC and two CdvB paralogs were shown to localize mid-cell in cells with segregated nucleoids, indicating that this system mediates cell division in Thaumarchaeota (Pelve et al., 2011). The *Ca. N. cavascurensis* CdvB paralogs (as all other thaumarchaeal CdvBs) all share the core ESCRT-III with the crenarchaeal CdvB sequences, contain a putative (but rather unconvincing) MIM2 motif necessary for interacting with CdvC in Crenarchaeota, while they all lack the C-terminal winged-helix domain responsible for interacting with CdvA (Samson et al., 2011; Ng et al., 2013). Interestingly, one of the *Ca. N. cavascurensis* paralogs (NCAV_0805) possesses a 40 amino-acids serine-rich C-terminal extension right after the putative MIM2 motif absent from other thaumarchaea. It is worth noting that CdvA is also absent from the published Aigarchaeota genomes *Ca. Caldiarchaeum subterraneum* (Nunoura et al., 2011) and *Ca. Calditenuis aerorheumensis* (Beam et al., 2016), while both phyla (Thaumarchaeota and Aigarchaeota) encode an atypical FtsZ homolog. *Thermococcales*, albeit they presumably divide with the FtsZ system, also encode CdvB and CdvC homologs, while no CdvA homolog is detectable (Makarova et al., 2010). Given the emerging differences in the molecular and regulatory aspects of the Cdv system between Crenarchaeota and Thaumarchaeota (Pelve et al., 2011; Ng et al., 2013), the intriguing additional roles of the system (Samson et al., 2017) and the fact that only CdvB and CdvC are homologous to the eukaryotic ESCRT-III system and therefore seem to have fixed roles in evolutionary terms, this observation raises interesting questions regarding the versatility of different players of the cell division apparatus.

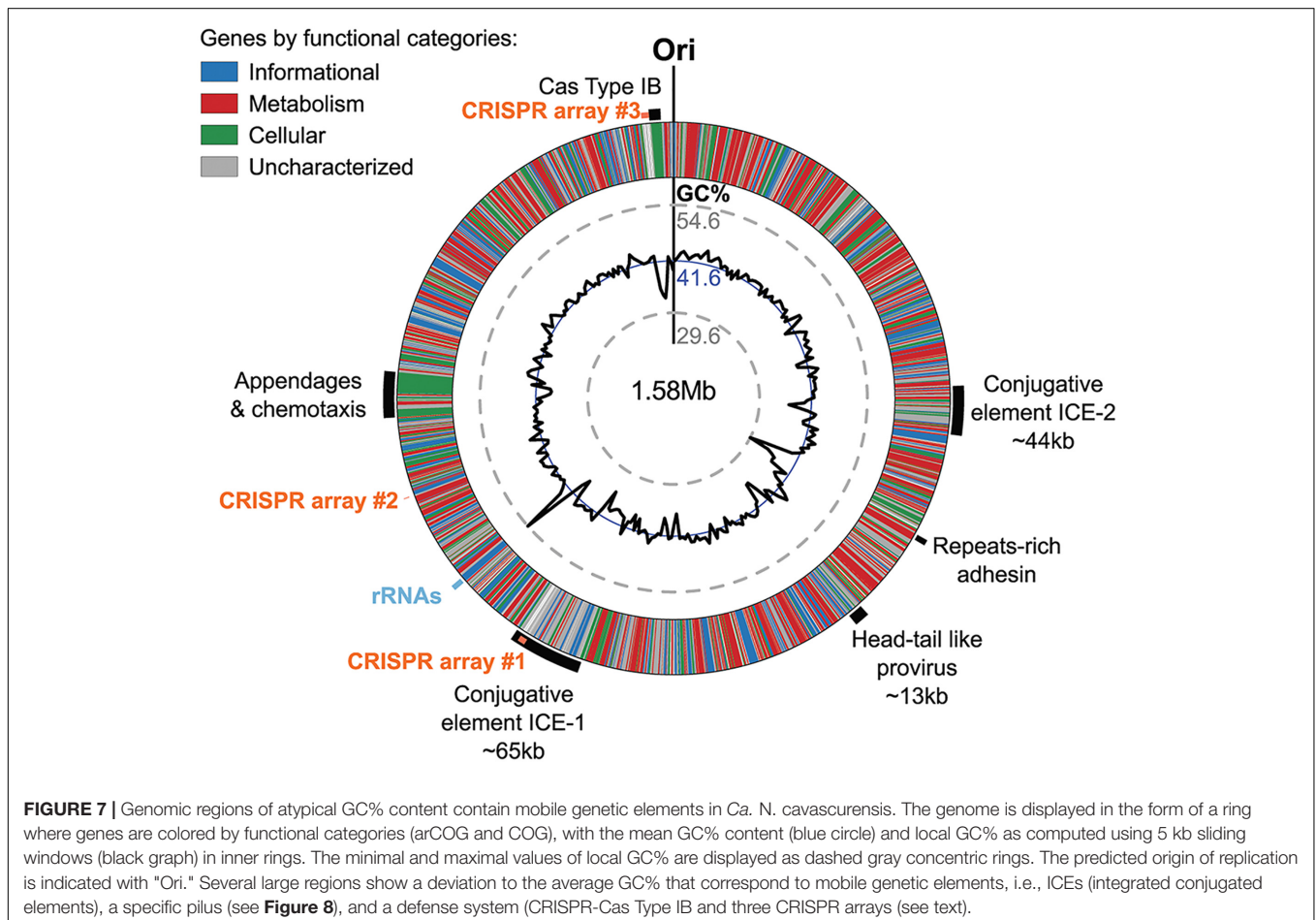
Ca. *N. cavascurensis* Has a Dynamic Genome

The *Ca. N. cavascurensis* genome contains large clusters of genes that showed deviations from the average G+C content of the genome (i.e., 41.6%), indicating that these regions might have been acquired by lateral gene transfer (Figure 7).

Two of the larger regions were integrative and conjugative elements (ICE-1 and ICE-2 in Figures 7, 8) of 65.5 and 43.6 kb, respectively. Both are integrated into tRNA genes and flanked by characteristic 21 and 24 bp-long direct repeats corresponding to the attachment sites, a typical sequence feature generated upon site-specific recombination between the cellular chromosome and a mobile genetic element (Grindley et al., 2006). ICE-1 and ICE-2 encode major proteins required for conjugation (colored in red in Figure 8A), including VirD4-like and VirB4-like ATPases, VirB6-like transfer protein and VirB2-like pilus protein (in ICE-2). The two elements also share homologs of CopG-like proteins containing the ribbon-helix-helix DNA-binding motifs and beta-propeller-fold sialidases (the latter appears to be truncated in ICE-1). The sialidases of ICE-1 and ICE-2 are most closely related to each other (37% identity over 273 aa alignment). It is not excluded that ICE-1 and ICE-2 have inherited the conjugation machinery as

well as the genes for the sialidase and the CopG-like protein from a common ancestor, but have subsequently diversified by accreting functionally diverse gene complements. Indeed, most of the genes carried by ICE-1 and ICE-2 are unrelated. ICE-1, besides encoding the conjugation proteins, carries many genes for proteins involved in DNA metabolism, including an archaeo-eukaryotic primase-superfamily 3 helicase fusion protein, Cdc6-like ATPase, various nucleases (HEPN, PD-(D/E)XK, PIN), DNA methyltransferases, diverse DNA-binding proteins and an integrase of the tyrosine recombinase superfamily (yellow in Figure 8A). The conservation of the attachment sites and the integrase gene as well as of the conjugative genes indicates that this element is likely to be a still-active Integrative and Conjugative Element (ICE) able to integrate into the chromosome and excise from it as a conjugative plasmid.

ICE-2 is shorter and encodes a distinct set of DNA metabolism proteins, including topoisomerase IA, superfamily 1 helicase, ParB-like partitioning protein, GIY-YIG and FLAP nucleases as well as several DNA-binding proteins. More importantly, this element also encodes a range of proteins involved in various metabolic activities as well as a Co/Zn/Cd efflux system that might provide relevant functions to the host (green in Figure 8A). In bacteria, ICE elements often contain cargo genes that are not



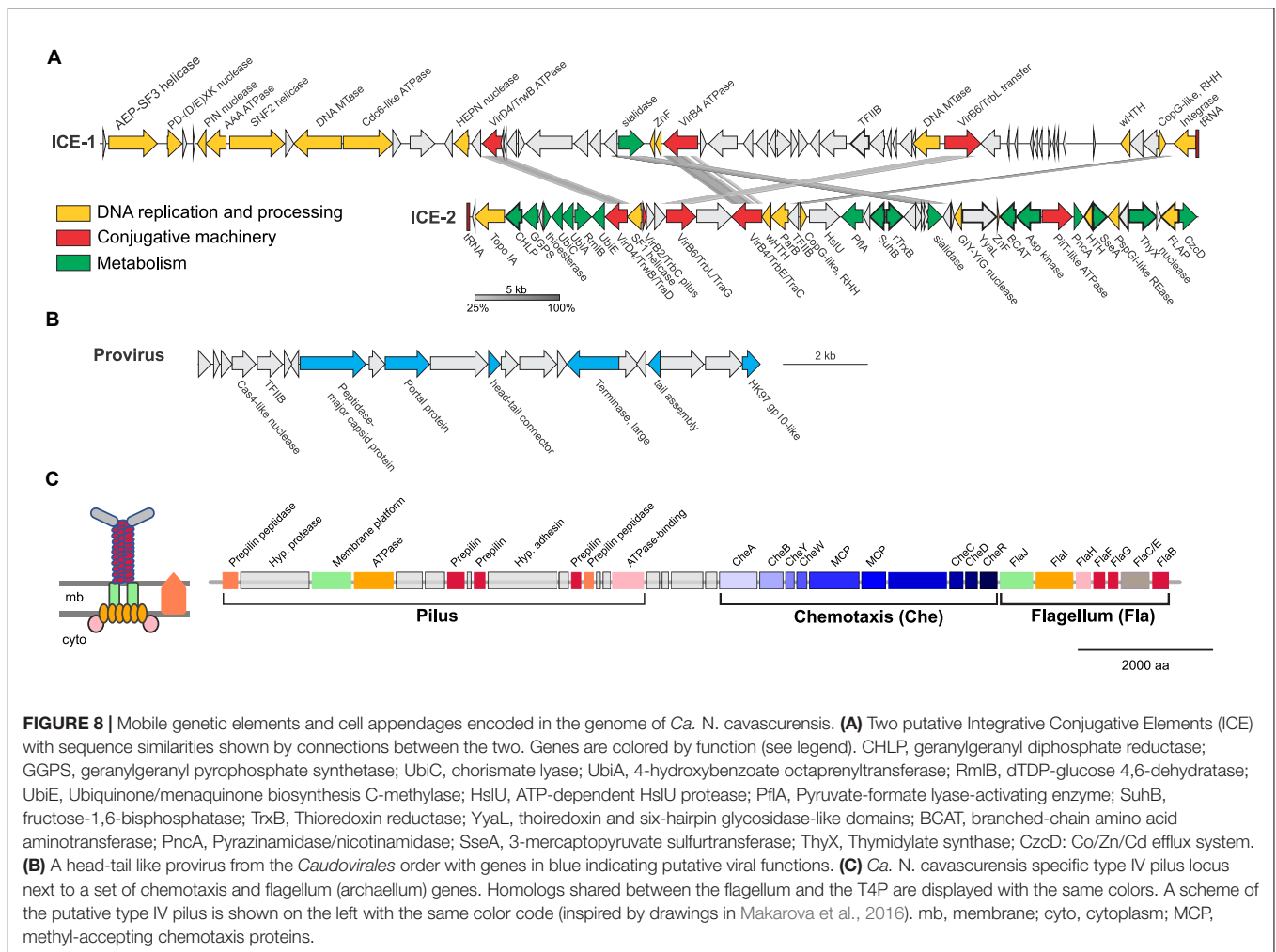


FIGURE 8 | Mobile genetic elements and cell appendages encoded in the genome of *Ca. N. cavascurens*. **(A)** Two putative Integrative Conjugative Elements (ICE) with sequence similarities shown by connections between the two. Genes are colored by function (see legend). CHLP, geranylgeranyl diphosphate reductase; GGPS, geranylgeranyl pyrophosphate synthetase; UbiC, chorismate lyase; UbiA, 4-hydroxybenzoate octaprenyltransferase; RmIB, dTDP-glucose 4,6-dehydratase; UbiE, Ubiquinone/menaquinone biosynthesis C-methylase; HslU, ATP-dependent HslU protease; PflA, Pyruvate-formate lyase-activating enzyme; SuhB, fructose-1,6-bisphosphatase; TrxB, Thioredoxin reductase; YyaL, thio redoxin and six-hairpin glycosidase-like domains; BCAT, branched-chain amino acid aminotransferase; PncA, Pyrazinamidase/nicotinamidase; SseA, 3-mercaptopyruvate sulfurtransferase; ThyX, Thymidylate synthase; CzcD: Co/Zn/Cd efflux system. **(B)** A head-tail like provirus from the *Caudovirales* order with genes in blue indicating putative viral functions. **(C)** *Ca. N. cavascurens* specific type IV pilus locus next to a set of chemotaxis and flagellum (archaellum) genes. Homologs shared between the flagellum and the T4P are displayed with the same colors. A scheme of the putative type IV pilus is shown on the left with the same color code (inspired by drawings in Makarova et al., 2016). mb, membrane; cyto, cytoplasm; MCP, methyl-accepting chemotaxis proteins.

related to the ICE life cycle and that confer novel phenotypes to host cells (Johnson and Grossman, 2015). It is possible that under certain conditions, genes carried by ICE-1 and ICE-2, and in particular the metabolic genes of ICE-2, improve the fitness of *Ca. N. cavascurens*. Notably, only ICE-1 encodes an integrase, whereas ICE-2 does not, suggesting that ICE-2 is an immobilized conjugative element that can be vertically inherited in AOA. Given that the attachment sites of the two elements do not share significant sequence similarity, the possibility that ICE-2 is mobilized *in trans* by the integrase of ICE-1 appears unlikely.

The third mobile genetic element is derived from a virus, related to members of the viral order *Caudovirales* (Figure 8B) (Prangishvili et al., 2017). It encodes several signature proteins of this virus group, most notably the large subunit of the terminase, the portal protein and the HK97-like major capsid protein (and several other viral homologs). All these proteins with homologs in viruses are involved in virion assembly and morphogenesis. However, no proteins involved in genome replication seem to be present. The element does not contain an integrase gene, nor is it flanked by attachment sites, which indicates that it is immobilized. Interestingly, a similar observation has been made

with the potential provirus-derived element in *Nitrososphaera viennensis* (Krupovic et al., 2011). Given that the morphogenesis genes of the virus appear to be intact, one could speculate that these elements represent domesticated viruses akin to gene transfer agents, as observed in certain methanogenic euryarchaea (Eislerling et al., 1999; Krupovic et al., 2010), or killer particles (Bobay et al., 2014), rather than deteriorating proviruses. Notably, ICE-1, ICE-2 and the provirus-derived element all encode divergent homologs of TFIIB, a transcription factor that could alter the promoter specificity to the RNA polymerase.

A fourth set of unique, potentially transferred genes encodes a putative pilus (Figure 8C). All genes required for the assembly of a type IV pilus (T4P) are present, including an ATPase and an ATPase binding protein, a membrane platform protein, a prepilin peptidase, and several prepilins (Makarova et al., 2016). Based on the arCOG families of the broadly conserved ATPase, ATPase-binding protein, membrane-platform protein, and prepilin peptidase (arCOG001817, -4148, -1808, and -2298), this pilus seems unique in family composition, but more similar to the archetypes of pili defined as clades 4 (A, H, I, J) by Makarova and colleagues, which are mostly found in *Sulfolobales* and *Desulfurococcales* (Figure 4 from Makarova et al., 2016).

Yet, the genes associated to this putative pilus seem to be more numerous, as the locus consists of approximately 16 genes (versus ~5 genes for the *aap* and *ups* in *Sulfolobus*). Such a combination of families is not found either in the T4P or flagellar loci found in analyzed Thaumarchaeota genomes. Prepilins are part of the core machinery of T4P, but display a high level of sequence diversity. The three prepilins found in *Ca. N. cavascurens* T4P locus correspond to families that are not found in any other genome from our dataset (arCOG003872, -5987, -7276). We found a putative adhesin right in between the genes encoding the prepilins, and therefore propose that this *Ca. N. cavascurens*-specific type IV pilus is involved in adhesion (**Figure 8C**). This pilus thus appears to be unique in protein families composition when compared to experimentally validated T4P homologs in archaea (flagellum, *ups*, *aap*, bindosome (Szabo et al., 2007; Zolghadr et al., 2007; Frols et al., 2008; Tripepi et al., 2010; Henche et al., 2012), bioinformatically predicted ones (Makarova et al., 2016), and the pili found in other Thaumarchaeota, which correspond to different types. Interestingly, this pilus gene cluster lies directly next to a conserved chemotaxis/archaeallum cluster as the one found in *Nitrososphaera gargensis* or *Nitrosoarchaeum limnia* (four predicted genes separate the T4P genes and *cheA*, **Figure 8C**) (Spang et al., 2012). This suggests that this pilus might be controlled by exterior stimuli through chemotaxis. The interplay between the archaeallum and pilus expression would be interesting to study in order to comprehend their respective roles.

The genome of *Ca. N. cavascurens* also carries traces of inactivated integrase genes as well as transposons related to bacterial and archaeal IS elements, suggesting that several other types of mobile genetic elements have been colonizing the genome. Collectively, these observations illuminate the flexibility of the *Ca. N. cavascurens* genome, prone to lateral gene transfer and invasion by alien elements. Accordingly, we found a CRISPR-Cas adaptive immunity system among the sets of genes specific to *Ca. N. cavascurens* that we could assign to the subtype I-B (Abby et al., 2014). We detected using the CRISPRFinder website (Grissa et al., 2007) at least three CRISPR arrays containing between 4 and 101 spacers presumably targeting mobile genetic elements associated with *Ca. N. cavascurens*, reinforcing the idea of a very dynamic genome. Interestingly, the second biggest CRISPR array (96 spacers) lies within the integrated conjugative element ICE-1, which we hypothesize to be still active. This suggests that ICE-1 may serve as a vehicle for the horizontal transfer of the CRISPR spacers between *Ca. N. cavascurens* and other organisms present in the same environment through conjugation, thus spreading the acquired immunity conferred by these spacers against common enemies.

CONCLUSION

We present an obligately thermophilic ammonia oxidizing archaeon from a hot spring on the Italian island of Ischia that is related to, but also clearly distinct from *Ca. Nitrosocaldus yellowstonensis*. It contains most of the genes that have been

found to be conserved among AOA and are implicated in energy and central carbon metabolism, except *nirK* encoding a nitrite reductase. Its genome gives indications for alternative energy metabolism and exhibits adaptations to (high temperature) extreme environments. However, it lacks an identifiable reverse gyrase, which is found in most thermophiles with optimal growth temperatures above 65°C and apparently harbors a provirus of head-and tail structure that is not usually found at high temperatures. *Ca. N. cavascurens* differs also in its gene sets for replication and cell division, which has implications for the function and evolution of these systems in archaea. In addition, its extensive mobilome and the defense system indicate that thermophilic AOA are in constant exchange with the environment and with neighboring organisms as discussed for other thermophiles (van Wolferen et al., 2013). This might have shaped and continues to shape the evolution of Thaumarchaeota in hot springs. The pivotal phylogenetic position of *Ca. N. cavascurens* will enable the reconstruction of the last common ancestor of AOA and provide further insights into the evolution of this ecologically widespread group of archaea.

For our enriched strain we propose a candidate status with the following taxonomic assignment:

Nitrosocaldales order

Nitrosocaldaceae fam. and

Candidatus Nitrosocaldus cavascurens sp. nov.

Etymology: L. adj. nitrosus, “full of natron,” here intended to mean nitrous (nitrite producer); L. masc.n. caldus, hot; cavascurensis (L.masc. gen) describes origin of sample (Terme di Cavascure, Ischia)

Locality: hot mud, outflow from hot underground spring, 77°C

Diagnosis: an ammonia oxidizing archaeon growing optimally around 68°C at neutral pH

under chemolithoautotrophic conditions with ammonia or urea,

spherically shaped with a diameter of approximately 0.6–0.8 μm, 4% sequence divergence in 16S rRNA gene from its next cultivated relative *Ca. Nitrosocaldus yellowstonensis*.

AUTHOR'S NOTE

Another report on the enrichment and genome analysis of a thermophilic Thaumarchaeota was submitted to *Frontiers in Microbiology*, Daebeler et al. (2018, in review), shortly before submission of this manuscript. The core genomes of both organisms share many of the above-discussed features, but it remains to be seen if the observed similarities extend to their shell genome.

AUTHOR CONTRIBUTIONS

CS conceived the study; MS did the first enrichments; MM and CR made the growth characterizations; KP did the electron microscopy; SSA assembled the genome and performed the phylogenetic and genomic analyses for annotation; MKE annotated all metabolic and information processing genes;

MKr analyzed the mobile genetic elements; CS, SSA, and MKE wrote the manuscript with contributions from MKr and MM.

FUNDING

This project was supported by ERC Advanced Grant TACKLE (695192) and Doktoratskolleg W1257 of the Austrian Science Fund (FWF). SSA was funded by a “Marie-Curie Action” fellowship, grant number THAUMECOPHYL 701981.

ACKNOWLEDGMENTS

We are grateful to Lucilla Monti for sharing her geological knowledge and for guiding us to the hot springs of Ischia in spring and fall 2013 and to Silvia Bulgheresi for crucial help in organizing the fall expedition. We also thank Romana Bittner for excellent technical assistance and continuous cultivation and

the 12 participants of the Bachelor practical class (University of Vienna) for help in sampling hot springs on Ischia and for setting up initial enrichments. We are grateful to Ricardo Alves for primer design, Thomas Pribasnik for qPCR and to Thomas Rattei and Florian Goldenberg for help in running computations and for providing access to the CUBE servers in Vienna. Special thanks go to LABGeM, the National Infrastructure “France Génomique” and the MicroScope team of the Genoscope (Evry, France) for providing the pipeline of annotation, and for their help during this study. We thank Eduardo Rocha and Marie Touchon for kindly providing access to their curated genome database.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2018.00028/full#supplementary-material>

REFERENCES

- Abby, S. S., Neron, B., Menager, H., Touchon, M., and Rocha, E. P. (2014). MacSyFinder: a program to mine genomes for molecular systems with an application to CRISPR-Cas systems. *PLOS ONE* 9:e110726. doi: 10.1371/journal.pone.0110726
- Adam, P. S., Borrel, G., Brochier-Armanet, C., and Gribaldo, S. (2017). The growing tree of Archaea: new perspectives on their diversity, evolution and ecology. *ISME J.* 11, 2407–2425. doi: 10.1038/ismej.2017.122
- Albertsen, M., Hugenholtz, P., Skarshewski, A., Nielsen, K. L., Tyson, G. W., and Nielsen, P. H. (2013). Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. *Nat. Biotechnol.* 31, 533–538. doi: 10.1038/nbt.2579
- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389
- Alves, R. J., Wanek, W., Zappe, A., Richter, A., Svenning, M. M., Schleper, C., et al. (2013). Nitrification rates in Arctic soils are associated with functionally distinct populations of ammonia-oxidizing archaea. *ISME J.* 7, 1620–1631. doi: 10.1038/ismej.2013.35
- Amann, R. L., Krumholz, L., and Stahl, D. A. (1990). Fluorescent-oligonucleotide probing of whole cells for determinative, phylogenetic, and environmental studies in microbiology. *J. Bacteriol.* 172, 762–770. doi: 10.1128/jb.172.2.762-770.1990
- Atomi, H., Matsumi, R., and Imanaka, T. (2004). Reverse gyrase is not a prerequisite for hyperthermophilic life. *J. Bacteriol.* 186, 4829–4833. doi: 10.1128/JB.186.14.4829-4833.2004
- Barns, S. M., Delwiche, C. F., Palmer, J. D., and Pace, N. R. (1996). Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *Proc. Natl. Acad. Sci. U.S.A.* 93, 9188–9193. doi: 10.1073/pnas.93.17.9188
- Bartossek, R., Nicol, G. W., Lanzen, A., Klenk, H. P., and Schleper, C. (2010). Homologues of nitrite reductases in ammonia-oxidizing archaea: diversity and genomic context. *Environ. Microbiol.* 12, 1075–1088. doi: 10.1111/j.1462-2920.2010.02153.x
- Bayer, B., Vojvoda, J., Offre, P., Alves, R. J., Elisabeth, N. H., Garcia, J. A., et al. (2016). Physiological and genomic characterization of two novel marine thaumarchaeal strains indicates niche differentiation. *ISME J.* 10, 1051–1063. doi: 10.1038/ismej.2015.200
- Beam, J. P., Jay, Z. J., Kozubal, M. A., and Inskeep, W. P. (2014). Niche specialization of novel Thaumarchaeota to oxic and hypoxic acidic geothermal springs of Yellowstone National Park. *ISME J.* 8, 938–951. doi: 10.1038/ismej.2013.193
- Beam, J. P., Jay, Z. J., Schmid, M. C., Rusch, D. B., Romine, M. F., Jennings Rde, M., et al. (2016). Ecophysiology of an uncultivated lineage of Aigarchaeota from an oxic, hot spring filamentous ‘streamer’ community. *ISME J.* 10, 210–224. doi: 10.1038/ismej.2015.83
- Bobay, L. M., Touchon, M., and Rocha, E. P. (2014). Pervasive domestication of defective prophages by bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 111, 12127–12132. doi: 10.1073/pnas.1405336111
- Bouthier de la Tour, C., Portemer, C., Nadal, M., Stetter, K. O., Forterre, P., and Duguet, M. (1990). Reverse gyrase, a hallmark of the hyperthermophilic archaeobacteria. *J. Bacteriol.* 172, 6803–6808. doi: 10.1128/jb.172.12.6803-6808.1990
- Brochier-Armanet, C., Boussau, B., Gribaldo, S., and Forterre, P. (2008). Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat. Rev. Microbiol.* 6, 245–252. doi: 10.1038/nrmicro1852
- Brochier-Armanet, C., and Forterre, P. (2007). Widespread distribution of archaeal reverse gyrase in thermophilic bacteria suggests a complex history of vertical inheritance and lateral gene transfers. *Archaea* 2, 83–93. doi: 10.1155/2006/582916
- Brochier-Armanet, C., Gribaldo, S., and Forterre, P. (2012). Spotlight on the Thaumarchaeota. *ISME J.* 6, 227–230. doi: 10.1038/ismej.2011.145
- Caranto, J. D., and Lancaster, K. M. (2017). Nitric oxide is an obligate bacterial nitrification intermediate produced by hydroxylamine oxidoreductase. *Proc. Natl. Acad. Sci. U.S.A.* 114, 8217–8222. doi: 10.1073/pnas.1704504114
- Chevenet, F., Croce, O., Hebrard, M., Christen, R., and Berry, V. (2010). ScriptTree: scripting phylogenetic graphics. *Bioinformatics* 26, 1125–1126. doi: 10.1093/bioinformatics/btq086
- Cole, J. K., Peacock, J. P., Dodsworth, J. A., Williams, A. J., Thompson, D. B., Dong, H., et al. (2013). Sediment microbial communities in Great Boiling Spring are controlled by temperature and distinct from water communities. *ISME J.* 7, 718–729. doi: 10.1038/ismej.2012.157
- Crisuolo, A., and Gribaldo, S. (2010). BMGE (block mapping and gathering with entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol. Biol.* 10:210. doi: 10.1186/1471-2148-10-210
- Cubonova, L., Richardson, T., Burkhart, B. W., Kelman, Z., Connolly, B. A., Reeve, J. N., et al. (2013). Archaeal DNA polymerase D but not DNA polymerase B is required for genome replication in *Thermococcus kodakarensis*. *J. Bacteriol.* 195, 2322–2328. doi: 10.1128/JB.02037-12
- de la Torre, J. R., Walker, C. B., Ingalls, A. E., Konneke, M., and Stahl, D. A. (2008). Cultivation of a thermophilic ammonia oxidizing archaeon synthesizing

- crenarchaeol. *Environ. Microbiol.* 10, 810–818. doi: 10.1111/j.1462-2920.2007.01506.x
- Dodsworth, J. A., Hungate, B. A., and Hedlund, B. P. (2011). Ammonia oxidation, denitrification and dissimilatory nitrate reduction to ammonium in two US Great Basin hot springs with abundant ammonia-oxidizing archaea. *Environ. Microbiol.* 13, 2371–2386. doi: 10.1111/j.1462-2920.2011.02508.x
- Eddy, S. R. (2011). Accelerated profile HMM searches. *PLOS Comput. Biol.* 7:e1002195. doi: 10.1371/journal.pcbi.1002195
- Eiserling, F., Pushkin, A., Gingery, M., and Bertani, G. (1999). Bacteriophage-like particles associated with the gene transfer agent of methanococcus voltae PS. *J. Gen. Virol.* 80(Pt 12), 3305–3308. doi: 10.1099/0022-1317-80-12-3305
- Eme, L., Reigstad, L. J., Spang, A., Lanzen, A., Weinmaier, T., Rattei, T., et al. (2013). Metagenomics of Kamchatkan hot spring filaments reveal two new major (hyper)thermophilic lineages related to Thaumarchaeota. *Res. Microbiol.* 164, 425–438. doi: 10.1016/j.resmic.2013.02.006
- Erguder, T. H., Boon, N., Wittebolle, L., Marzorati, M., and Verstraete, W. (2009). Environmental factors shaping the ecological niches of ammonia-oxidizing archaea. *FEMS Microbiol. Rev.* 33, 855–869. doi: 10.1111/j.1574-6976.2009.00179.x
- Evans, D. J. Jr., Evans, D. G., Engstrand, L., and Graham, D. Y. (1992). Urease-associated heat shock protein of *Helicobacter pylori*. *Infect. Immun.* 60, 2125–2127.
- Finn, R. D., Bateman, A., Clements, J., Coghill, P., Eberhardt, R. Y., Eddy, S. R., et al. (2014). Pfam: the protein families database. *Nucleic Acids Res.* 42, D222–D230. doi: 10.1093/nar/gkt1223
- Forterre, P. (2002). A hot story from comparative genomics: reverse gyrase is the only hyperthermophile-specific protein. *Trends Genet.* 18, 236–237. doi: 10.1016/S0168-9525(02)02650-1
- Frols, S., Ajon, M., Wagner, M., Teichmann, D., Zolghadr, B., Folea, M., et al. (2008). UV-inducible cellular aggregation of the hyperthermophilic archaeon *Sulfolobus solfataricus* is mediated by pili formation. *Mol. Microbiol.* 70, 938–952. doi: 10.1111/j.1365-2958.2008.06459.x
- Giles, T. N., and Graham, D. E. (2008). Crenarchaeal arginine decarboxylase evolved from an S-adenosylmethionine decarboxylase enzyme. *J. Biol. Chem.* 283, 25829–25838. doi: 10.1074/jbc.M802674200
- Gough, J., Karplus, K., Hughey, R., and Chothia, C. (2001). Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. *J. Mol. Biol.* 313, 903–919. doi: 10.1006/jmbi.2001.5080
- Grasso, S., and Tell, G. (2014). Base excision repair in Archaea: back to the future in DNA repair. *DNA Repair* 21, 148–157. doi: 10.1016/j.dnarep.2014.05.006
- Greening, C., Biswas, A., Carere, C. R., Jackson, C. J., Taylor, M. C., Stott, M. B., et al. (2016). Genomic and metagenomic surveys of hydrogenase distribution indicate H₂ is a widely utilised energy source for microbial growth and survival. *ISME J.* 10, 761–777. doi: 10.1038/ismej.2015.153
- Gribaldo, S., and Brochier-Armanet, C. (2006). The origin and evolution of Archaea: a state of the art. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 361, 1007–1022. doi: 10.1098/rstb.2006.1841
- Griffiths, R. I., Whiteley, A. S., O'donnell, A. G., and Bailey, M. J. (2000). Rapid method for coextraction of DNA and RNA from natural environments for analysis of ribosomal DNA- and rRNA-based microbial community composition. *Appl. Environ. Microbiol.* 66, 5488–5491. doi: 10.1128/AEM.66.12.5488-5491.2000
- Grindley, N. D., Whiteson, K. L., and Rice, P. A. (2006). Mechanisms of site-specific recombination. *Annu. Rev. Biochem.* 75, 567–605. doi: 10.1146/annurev.biochem.73.011303.073908
- Grissa, I., Vergnaud, G., and Pourcel, C. (2007). CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* 35, W52–W57. doi: 10.1093/nar/gkm360
- Grogan, D. W. (2015). Understanding DNA repair in hyperthermophilic Archaea: persistent gaps and other reasons to focus on the fork. *Archaea* 2015:942605. doi: 10.1155/2015/942605
- Groussin, M., and Gouy, M. (2011). Adaptation to environmental temperature is a major determinant of molecular evolutionary rates in archaea. *Mol. Biol. Evol.* 28, 2661–2674. doi: 10.1093/molbev/msr098
- Haft, D. H., Selengut, J. D., and White, O. (2003). The TIGRFAMs database of protein families. *Nucleic Acids Res.* 31, 371–373. doi: 10.1093/nar/gkg128
- Hallam, S. J., Mincer, T. J., Schleper, C., Preston, C. M., Roberts, K., Richardson, P. M., et al. (2006). Pathways of carbon assimilation and ammonia oxidation suggested by environmental genomic analyses of marine Crenarchaeota. *PLOS Biol.* 4:e95. doi: 10.1371/journal.pbio.0040095
- Hatzenpichler, R. (2012). Diversity, physiology, and niche differentiation of ammonia-oxidizing archaea. *Appl. Environ. Microbiol.* 78, 7501–7510. doi: 10.1128/AEM.01960-12
- Henche, A. L., Ghosh, A., Yu, X., Jeske, T., Egelman, E., and Albers, S. V. (2012). Structure and function of the adhesive type IV pilus of *Sulfolobus acidocaldarius*. *Environ. Microbiol.* 14, 3188–3202. doi: 10.1111/j.1462-2920.2012.02898.x
- Higashibata, H., Fujiwara, S., Ezaki, S., Takagi, M., Fukui, K., and Imanaka, T. (2000). Effect of polyamines on histone-induced DNA compaction of hyperthermophilic archaea. *J. Biosci. Bioeng.* 89, 103–106. doi: 10.1016/S1389-1723(00)88061-5
- Hügler, M., Menendez, C., Schägger, H., and Fuchs, G. (2002). Malonyl-coenzyme A reductase from *Chloroflexus aurantiacus*, a key enzyme of the 3-hydroxypropionate cycle. (for) autotrophic CO₂ fixation. *J. Bacteriol.* 184, 2404–2410. doi: 10.1128/JB.184.9.2404-2410.2002
- Ishino, Y., and Narumi, I. (2015). DNA repair in hyperthermophilic and hyperradioresistant microorganisms. *Curr. Opin. Microbiol.* 25, 103–112. doi: 10.1016/j.mib.2015.05.010
- Jackson, C. R., Langner, H. W., Donahoe-Christiansen, J., Inskip, W. P., and McDermott, T. R. (2001). Molecular analysis of microbial community structure in an arsenite-oxidizing acidic thermal spring. *Environ. Microbiol.* 3, 532–542. doi: 10.1046/j.1462-2920.2001.00221.x
- Jiang, H., Huang, Q., Dong, H., Wang, P., Wang, F., Li, W., et al. (2010). RNA-based investigation of ammonia-oxidizing archaea in hot springs of Yunnan Province, China. *Appl. Environ. Microbiol.* 76, 4538–4541. doi: 10.1128/AEM.00143-10
- Johnson, C. M., and Grossman, A. D. (2015). Integrative and conjugative elements (ICEs): what they do and how they work. *Annu. Rev. Genet.* 49, 577–601. doi: 10.1146/annurev-genet-112414-055018
- Kanai, T., Matsuoka, R., Beppu, H., Nakajima, A., Okada, Y., Atomi, H., et al. (2011). Distinct physiological roles of the three [NiFe]-hydrogenase orthologs in the hyperthermophilic archaeon *Thermococcus kodakarensis*. *J. Bacteriol.* 193, 3109–3116. doi: 10.1128/JB.01072-10
- Katoh, K., and Standley, D. M. (2014). MAFFT: iterative refinement and additional methods. *Methods Mol. Biol.* 1079, 131–146. doi: 10.1007/978-1-62703-646-7_8
- Kazlauskas, D., Sezonov, G., Charpin, N., Venclovas, C., Forterre, P., and Krupovic, M. (2017). Novel families of archaeo-eukaryotic primases associated with mobile genetic elements of Bacteria and Archaea. *J. Mol. Biol.* doi: 10.1016/j.jmb.2017.11.014 [Epub ahead of print].
- Kelman, Z., and White, M. F. (2005). Archaeal DNA replication and repair. *Curr. Opin. Microbiol.* 8, 669–676. doi: 10.1016/j.mib.2005.10.001
- Kerou, M., Eloy Alves, R. J., and Schleper, C. (2016a). “Nitrososphaeria,” in *Bergey's Manual of Systematics of Archaea and Bacteria*, ed. Bergey's Manual Trust (Hoboken, NJ: John Wiley & Sons), 1–8. doi: 10.1002/9781118960608.cbm00055
- Kerou, M., Offre, P., Villedor, L., Abby, S. S., Melcher, M., Nagler, M., et al. (2016b). Proteomics and comparative genomics of *Nitrososphaera viennensis* reveal the core genome and adaptations of archaeal ammonia oxidizers. *Proc. Natl. Acad. Sci. U.S.A.* 113, E7937–E7946.
- Kim, J. G., Park, S. J., Sinnighe Damste, J. S., Schouten, S., Rijpstra, W. I., Jung, M. Y., et al. (2016). Hydrogen peroxide detoxification is a key mechanism for growth of ammonia-oxidizing archaea. *Proc. Natl. Acad. Sci. U.S.A.* 113, 7888–7893. doi: 10.1073/pnas.1605501113
- Kockelkorn, D., and Fuchs, G. (2009). Malonic semialdehyde reductase, succinic semialdehyde reductase, and succinyl-coenzyme A reductase from *Metallosphaera sedula*: enzymes of the autotrophic 3-hydroxypropionate/4-hydroxybutyrate cycle in Sulfolobales. *J. Bacteriol.* 191, 6352–6362. doi: 10.1128/JB.00794-09
- Konneke, M., Bernhard, A. E., De La Torre, J. R., Walker, C. B., Waterbury, J. B., and Stahl, D. A. (2005). Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* 437, 543–546. doi: 10.1038/nature03911

- Könneke, M., Schubert, D. M., Brown, P. C., Hugler, M., Standfest, S., Schwander, T., et al. (2014). Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO₂ fixation. *Proc. Natl. Acad. Sci. U.S.A.* 111, 8239–8244. doi: 10.1073/pnas.1402028111
- Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., and Phillippy, A. M. (2017). Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* 27, 722–736. doi: 10.1101/gr.215087.116
- Kozlowski, J. A., Kits, K. D., and Stein, L. Y. (2016a). Comparison of nitrogen oxide metabolism among diverse ammonia-oxidizing bacteria. *Front. Microbiol.* 7:1090. doi: 10.3389/fmicb.2016.01090
- Kozlowski, J. A., Stieglmeier, M., Schleper, C., Klotz, M. G., and Stein, L. Y. (2016b). Pathways and key intermediates required for obligate aerobic ammonia-dependent chemolithotrophy in bacteria and Thaumarchaeota. *ISME J.* 10, 1836–1845. doi: 10.1038/ismej.2016.2
- Krupovic, M., Forterre, P., and Bamford, D. H. (2010). Comparative analysis of the mosaic genomes of tailed archaeal viruses and proviruses suggests common themes for virion architecture and assembly with tailed viruses of bacteria. *J. Mol. Biol.* 397, 144–160. doi: 10.1016/j.jmb.2010.01.037
- Krupovic, M., Spang, A., Gribaldo, S., Forterre, P., and Schleper, C. (2011). A thaumarchaeal provirus testifies for an ancient association of tailed viruses with archaea. *Biochem. Soc. Trans.* 39, 82–88. doi: 10.1042/BST0390082
- Kurtz, S., Phillippy, A., Delcher, A. L., Smoot, M., Shumway, M., Antonescu, C., et al. (2004). Versatile and open software for comparing large genomes. *Genome Biol.* 5:R12. doi: 10.1186/gb-2004-5-2-r12
- Li, H., Yang, Q., Li, J., Gao, H., Li, P., and Zhou, H. (2015). The impact of temperature on microbial diversity and AOA activity in the Tengchong Geothermal Field, China. *Sci. Rep.* 5:17056. doi: 10.1038/srep17056
- Lin, X., Handley, K. M., Gilbert, J. A., and Kostka, J. E. (2015). Metabolic potential of fatty acid oxidation and anaerobic respiration by abundant members of Thaumarchaeota and Thermoplasmata in deep anoxic peat. *ISME J.* 9, 2740–2744. doi: 10.1038/ismej.2015.77
- Lopez-Garcia, P., Brochier, C., Moreira, D., and Rodriguez-Valera, F. (2004). Comparative analysis of a genome fragment of an uncultivated mesopelagic crenarchaeote reveals multiple horizontal gene transfers. *Environ. Microbiol.* 6, 19–34. doi: 10.1046/j.1462-2920.2003.00533.x
- Luo, H., Zhang, C. T., and Gao, F. (2014). Ori-Finder 2, an integrated tool to predict replication origins in the archaeal genomes. *Front. Microbiol.* 5:482. doi: 10.3389/fmicb.2014.00482
- Makarova, K. S., Koonin, E. V., and Albers, S. V. (2016). Diversity and evolution of type IV pili systems in archaea. *Front. Microbiol.* 7:667. doi: 10.3389/fmicb.2016.00667
- Makarova, K. S., Krupovic, M., and Koonin, E. V. (2014). Evolution of replicative DNA polymerases in archaea and their contributions to the eukaryotic replication machinery. *Front. Microbiol.* 5:354. doi: 10.3389/fmicb.2014.00354
- Makarova, K. S., Wolf, Y. I., and Koonin, E. V. (2015). Archaeal clusters of orthologous genes (arCOGs): an update and application for analysis of shared features between Thermococcales, Methanococcales, and Methanobacteriales. *Life* 5, 818–840. doi: 10.3390/life5010818
- Makarova, K. S., Yutin, N., Bell, S. D., and Koonin, E. V. (2010). Evolution of diverse cell division and vesicle formation systems in Archaea. *Nat. Rev. Microbiol.* 8, 731–741. doi: 10.1038/nrmicro2406
- Martens-Habbena, W., Qin, W., Horak, R. E., Urakawa, H., Schauer, A. J., Moffett, J. W., et al. (2015). The production of nitric oxide by marine ammonia-oxidizing archaea and inhibition of archaeal ammonia oxidation by a nitric oxide scavenger. *Environ. Microbiol.* 17, 2261–2274. doi: 10.1111/1462-2920.12677
- Medigue, C., Calteau, A., Cruveiller, S., Gachet, M., Gautreau, G., Josso, A., et al. (2017). MicroScope-an integrated resource for community expertise of gene functions and comparative analysis of microbial genomic and metabolic data. *Brief. Bioinform.* doi: 10.1093/bib/bbx113 [Epub ahead of print].
- Miele, V., Penel, S., Daubin, V., Picard, F., Kahn, D., and Duret, L. (2012). High-quality sequence clustering guided by network topology and multiple alignment likelihood. *Bioinformatics* 28, 1078–1085. doi: 10.1093/bioinformatics/bts098
- Miele, V., Penel, S., and Duret, L. (2011). Ultra-fast sequence clustering from similarity networks with SiLiX. *BMC Bioinformatics* 12:116. doi: 10.1186/1471-2105-12-116
- Musmann, M., Brito, I., Pitcher, A., Sinnighe Damste, J. S., Hatzenpichler, R., Richter, A., et al. (2011). Thaumarchaeotes abundant in refinery nitrifying sludges express amoA but are not obligate autotrophic ammonia oxidizers. *Proc. Natl. Acad. Sci. U.S.A.* 108, 16771–16776. doi: 10.1073/pnas.1106427108
- Muyzer, G., De Waal, E. C., and Uitterlinden, A. G. (1993). Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl. Environ. Microbiol.* 59, 695–700.
- Neves, C., Da Costa, M. S., and Santos, H. (2005). Compatible solutes of the hyperthermophile *Palaeococcus ferrophilus*: osmoadaptation and thermoadaptation in the order *Thermococcales*. *Appl. Environ. Microbiol.* 71, 8091–8098. doi: 10.1128/AEM.71.12.8091-8098.2005
- Ng, K. H., Srinivas, V., Srinivasan, R., and Balasubramanian, M. (2013). The *Nitrosopumilus maritimus* CdvB, but not FtsZ, assembles into polymers. *Archaea* 2013:104147. doi: 10.1155/2013/104147
- Nguyen, L. T., Schmidt, H. A., Von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300
- Nicol, G. W., Leininger, S., Schleper, C., and Prosser, J. I. (2008). The influence of soil pH on the diversity, abundance and transcriptional activity of ammonia oxidizing archaea and bacteria. *Environ. Microbiol.* 10, 2966–2978. doi: 10.1111/j.1462-2920.2008.01701.x
- Nicol, G. W., and Schleper, C. (2006). Ammonia-oxidising Crenarchaeota: important players in the nitrogen cycle? *Trends Microbiol.* 14, 207–212.
- Nishizawa, M., Sakai, S., Konno, U., Nakahara, N., Takaki, Y., Saito, Y., et al. (2016). Nitrogen and oxygen isotope effects of ammonia oxidation by thermophilic *Thaumarchaeota* from a geothermal water stream. *Appl. Environ. Microbiol.* 82, 4492–4504. doi: 10.1128/AEM.00250-16
- Nunoura, T., Hirayama, H., Takami, H., Oida, H., Nishi, S., Shimamura, S., et al. (2005). Genetic and functional properties of uncultivated thermophilic crenarchaeotes from a subsurface gold mine as revealed by analysis of genome fragments. *Environ. Microbiol.* 7, 1967–1984. doi: 10.1111/j.1462-2920.2005.00881.x
- Nunoura, T., Takaki, Y., Kakuta, J., Nishi, S., Sugahara, J., Kazama, H., et al. (2011). Insights into the evolution of Archaea and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. *Nucleic Acids Res.* 39, 3204–3223. doi: 10.1093/nar/gkq1228
- Nurk, S., Bankevich, A., Antipov, D., Gurevich, A. A., Korobeynikov, A., Lapidus, A., et al. (2013). Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. *J. Comput. Biol.* 20, 714–737. doi: 10.1089/cmb.2013.0084
- Offre, P., Kerou, M., Spang, A., and Schleper, C. (2014). Variability of the transporter gene complement in ammonia-oxidizing archaea. *Trends Microbiol.* 22, 665–675. doi: 10.1016/j.tim.2014.07.007
- Offre, P., Spang, A., and Schleper, C. (2013). Archaea in biogeochemical cycles. *Annu. Rev. Microbiol.* 67, 437–457. doi: 10.1146/annurev-micro-092412-155614
- Okonechnikov, K., Golosova, O., Fursov, M., and Team, U. (2012). Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics* 28, 1166–1167. doi: 10.1093/bioinformatics/bts091
- Oshima, T. (2007). Unique polyamines produced by an extreme thermophile, *Thermus thermophilus*. *Amino Acids* 33, 367–372. doi: 10.1007/s00726-007-0526-z
- Otte, J., Mall, A., Schubert, D. M., Konneke, M., and Berg, I. A. (2015). Malonic semialdehyde reductase from the archaeon *Nitrosopumilus maritimus* is involved in the autotrophic 3-hydroxypropionate/4-hydroxybutyrate cycle. *Appl. Environ. Microbiol.* 81, 1700–1707. doi: 10.1128/AEM.03390-14
- Ovreas, L., Forney, L., Daae, F. L., and Torsvik, V. (1997). Distribution of bacterioplankton in meromictic Lake Saelenvannet, as determined by denaturing gradient gel electrophoresis of PCR-amplified gene fragments coding for 16S rRNA. *Appl. Environ. Microbiol.* 63, 3367–3373.
- Park, S. J., Kim, J. G., Jung, M. Y., Kim, S. J., Cha, I. T., Ghai, R., et al. (2012). Draft genome sequence of an ammonia-oxidizing archaeon, "*Candidatus Nitrosopumilus sediminis*" AR2, from Svalbard in the Arctic Circle. *J. Bacteriol.* 194, 6948–6949. doi: 10.1128/JB.01869-12
- Pelvé, E. A., Lindas, A. C., Martens-Habbena, W., De La Torre, J. R., Stahl, D. A., and Bernander, R. (2011). Cdv-based cell division and cell cycle organization

- in the thaumarchaeon *Nitrosopumilus maritimus*. *Mol. Microbiol.* 82, 555–566. doi: 10.1111/j.1365-2958.2011.07834.x
- Pester, M., Schleper, C., and Wagner, M. (2011). The Thaumarchaeota: an emerging view of their phylogeny and ecophysiology. *Curr. Opin. Microbiol.* 14, 300–306. doi: 10.1016/j.mib.2011.04.007
- Peters, J. W., Schut, G. J., Boyd, E. S., Mulder, D. W., Shepard, E. M., Broderick, J. B., et al. (2015). [FeFe]- and [NiFe]-hydrogenase diversity, mechanism, and maturation. *Biochim. Biophys. Acta* 1853, 1350–1369. doi: 10.1016/j.bbamcr.2014.11.021
- Petitjean, C., Deschamps, P., Lopez-Garcia, P., and Moreira, D. (2014). Rooting the domain archaea by phylogenomic analysis supports the foundation of the new kingdom Proteoarchaeota. *Genome Biol. Evol.* 7, 191–204. doi: 10.1093/gbe/evu274
- Poli, A., Di Donato, P., Abbamondi, G. R., and Nicolaus, B. (2011). Synthesis, production, and biotechnological applications of exopolysaccharides and polyhydroxyalkanoates by archaea. *Archaea* 2011:693253. doi: 10.1155/2011/693253
- Prangishvili, D., Bamford, D. H., Forterre, P., Iranzo, J., Koonin, E. V., and Krupovic, M. (2017). The enigmatic archaeal virosphere. *Nat. Rev. Microbiol.* 15, 724–739. doi: 10.1038/nrmicro.2017.125
- Prosser, J. I., and Nicol, G. W. (2008). Relative contributions of archaea and bacteria to aerobic ammonia oxidation in the environment. *Environ. Microbiol.* 10, 2931–2941. doi: 10.1111/j.1462-2920.2008.01775.x
- Qin, W., Amin, S. A., Martens-Habben, W., Walker, C. B., Urakawa, H., Devol, A. H., et al. (2014). Marine ammonia-oxidizing archaeal isolates display obligate mixotrophy and wide ecotypic variation. *Proc. Natl. Acad. Sci. U.S.A.* 111, 12504–12509. doi: 10.1073/pnas.1324115111
- Ramos-Vera, W. H., Weiss, M., Strittmatter, E., Kockelkorn, D., and Fuchs, G. (2011). Identification of missing genes and enzymes for autotrophic carbon fixation in crenarchaeota. *J. Bacteriol.* 193, 1201–1211. doi: 10.1128/JB.01156-10
- Raymann, K., Brochier-Armanet, C., and Gribaldo, S. (2015). The two-domain tree of life is linked to a new root for the Archaea. *Proc. Natl. Acad. Sci. U.S.A.* 112, 6670–6675. doi: 10.1073/pnas.1420858112
- Reigstad, L. J., Richter, A., Daims, H., Urich, T., Schwark, L., and Schleper, C. (2008). Nitrification in terrestrial hot springs of Iceland and Kamchatka. *FEMS Microbiol. Ecol.* 64, 167–174. doi: 10.1111/j.1574-6941.2008.00466.x
- Rouillon, C., and White, M. F. (2011). The evolution and mechanisms of nucleotide excision repair proteins. *Res. Microbiol.* 162, 19–26. doi: 10.1016/j.resmic.2010.09.003
- Saier, M. H. Jr., Reddy, V. S., Tamang, D. G., and Vastermark, A. (2014). The transporter classification database. *Nucleic Acids Res.* 42, D251–D258. doi: 10.1093/nar/gkt1097
- Samson, R. Y., Dobro, M. J., Jensen, G. J., and Bell, S. D. (2017). The structure, function and roles of the archaeal ESCRT apparatus. *Subcell. Biochem.* 84, 357–377. doi: 10.1007/978-3-319-53047-5_12
- Samson, R. Y., Obita, T., Hodgson, B., Shaw, M. K., Chong, P. L., Williams, R. L., et al. (2011). Molecular and structural basis of ESCRT-III recruitment to membranes during archaeal cell division. *Mol. Cell* 41, 186–196. doi: 10.1016/j.molcel.2010.12.018
- Schleper, C., Jurgens, G., and Jonuscheit, M. (2005). Genomic studies of uncultivated archaea. *Nat. Rev. Microbiol.* 3, 479–488. doi: 10.1038/nrmicro1159
- Schleper, C., and Nicol, G. W. (2010). Ammonia-oxidising archaea—physiology, ecology and evolution. *Adv. Microb. Physiol.* 57, 1–41. doi: 10.1016/B978-0-12-381045-8.00001-1
- Shen, T., Stieglmeier, M., Dai, J., Urich, T., and Schleper, C. (2013). Responses of the terrestrial ammonia-oxidizing archaeon *Ca. Nitrososphaera viennensis* and the ammonia-oxidizing bacterium *Nitrososphaera multififormis* to nitrification inhibitors. *FEMS Microbiol. Lett.* 344, 121–129. doi: 10.1111/1574-6968.12164
- Shi, Y., Tyson, G. W., Eppley, J. M., and Delong, E. F. (2011). Integrated metatranscriptomic and metagenomic analyses of stratified microbial assemblages in the open ocean. *ISME J.* 5, 999–1013. doi: 10.1038/ismej.2010.189
- Shima, S., Herault, D. A., Berkessel, A., and Thauer, R. K. (1998). Activation and thermostabilization effects of cyclic 2, 3-diphosphoglycerate on enzymes from the hyperthermophilic *Methanopyrus kandleri*. *Arch. Microbiol.* 170, 469–472. doi: 10.1007/s002030050669
- Siguier, P., Perochon, J., Lestrade, L., Mahillon, J., and Chandler, M. (2006). ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res.* 34, D32–D36. doi: 10.1093/nar/gkj014
- Simon, J., and Klotz, M. G. (2013). Diversity and evolution of bioenergetic systems involved in microbial nitrogen compound transformations. *Biochim. Biophys. Acta* 1827, 114–135. doi: 10.1016/j.bbabi.2012.07.005
- Soding, J. (2005). Protein homology detection by HMM-HMM comparison. *Bioinformatics* 21, 951–960. doi: 10.1093/bioinformatics/bti125
- Sondergaard, D., Pedersen, C. N., and Greening, C. (2016). HydDB: a web tool for hydrogenase classification and analysis. *Sci. Rep.* 6:34212. doi: 10.1038/srep34212
- Spang, A., Hatzenpichler, R., Brochier-Armanet, C., Rattei, T., Tischler, P., Spieck, E., et al. (2010). Distinct gene set in two different lineages of ammonia-oxidizing archaea supports the phylum Thaumarchaeota. *Trends Microbiol.* 18, 331–340. doi: 10.1016/j.tim.2010.06.003
- Spang, A., Poehlein, A., Offre, P., Zumbragel, S., Haider, S., Rychlik, N., et al. (2012). The genome of the ammonia-oxidizing *Candidatus Nitrososphaera gargensis*: insights into metabolic versatility and environmental adaptations. *Environ. Microbiol.* 14, 3122–3145. doi: 10.1111/j.1462-2920.2012.02893.x
- Stahl, D. A., and Amann, R. (1991). “Development and application of nucleic acid probes,” in *Nucleic Acid Techniques in Bacterial Systematics*, eds E. Stackebrandt and M. Goodfellow (New York, NY: John Wiley & Sons), 205–248.
- Stahl, D. A., and de la Torre, J. R. (2012). Physiology and diversity of ammonia-oxidizing archaea. *Annu. Rev. Microbiol.* 66, 83–101. doi: 10.1146/annurev-micro-092611-150128
- Stieglmeier, M., Klingl, A., Alves, R. J., Rittmann, S. K., Melcher, M., Leisch, N., 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. doi: 10.1099/ijs.0.063172-0
- Szabo, Z., Sani, M., Groeneveld, M., Zolghadr, B., Scheler, J., Albers, S. V., et al. (2007). Flagellar motility and structure in the hyperthermoacidophilic archaeon *Sulfolobus solfataricus*. *J. Bacteriol.* 189, 4305–4309. doi: 10.1128/JB.00042-07
- Tourna, M., Stieglmeier, M., Spang, A., Konneke, M., Schintlmeister, A., Urich, T., et al. (2011). *Nitrososphaera viennensis*, an ammonia oxidizing archaeon from soil. *Proc. Natl. Acad. Sci. U.S.A.* 108, 8420–8425. doi: 10.1073/pnas.1013488108
- Treusch, A. H., Leininger, S., Kletzin, A., Schuster, S. C., Klenk, H. P., and Schleper, C. (2005). Novel genes for nitrite reductase and Amo-related proteins indicate a role of uncultivated mesophilic crenarchaeota in nitrogen cycling. *Environ. Microbiol.* 7, 1985–1995. doi: 10.1111/j.1462-2920.2005.00906.x
- Tripepi, M., Imam, S., and Pohlschroder, M. (2010). *Haloflexax volcanii* flagella are required for motility but are not involved in PibD-dependent surface adhesion. *J. Bacteriol.* 192, 3093–3102. doi: 10.1128/JB.00133-10
- Vajjala, N., Martens-Habben, W., Sayavedra-Soto, L. A., Schauer, A., Bottomley, P. J., Stahl, D. A., et al. (2013). Hydroxylamine as an intermediate in ammonia oxidation by globally abundant marine archaea. *Proc. Natl. Acad. Sci. U.S.A.* 110, 1006–1011. doi: 10.1073/pnas.1214272110
- Vallenet, D., Calteau, A., Cruveiller, S., Gachet, M., Lajus, A., Josso, A., et al. (2017). MicroScope in 2017: an expanding and evolving integrated resource for community expertise of microbial genomes. *Nucleic Acids Res.* 45, D517–D528. doi: 10.1093/nar/gkw1101
- Vallenet, D., Engelen, S., Mornico, D., Cruveiller, S., Fleury, L., Lajus, A., et al. (2009). MicroScope: a platform for microbial genome annotation and comparative genomics. *Database* 2009:bap021. doi: 10.1093/database/bap021
- van Wolferen, M., Ajon, M., Driessen, A. J., and Albers, S. V. (2013). How hyperthermophiles adapt to change their lives: DNA exchange in extreme conditions. *Extremophiles* 17, 545–563. doi: 10.1007/s00792-013-0552-6
- Williams, T. A., Szollosi, G. J., Spang, A., Foster, P. G., Heaps, S. E., Boussau, B., et al. (2017). Integrative modeling of gene and genome evolution roots the archaeal tree of life. *Proc. Natl. Acad. Sci. U.S.A.* 114, E4602–E4611. doi: 10.1073/pnas.1618463114
- Yan, J., Beattie, T. R., Rojas, A. L., Schermerhorn, K., Gristwood, T., Trinidad, J. C., et al. (2017). Identification and characterization of a heterotrimeric archaeal DNA polymerase holoenzyme. *Nat. Commun.* 8:15075. doi: 10.1038/ncomms15075
- Yan, J., Haaijer, S. C., Op den Camp, H. J., Van Niftrik, L., Stahl, D. A., Konneke, M., et al. (2012). Mimicking the oxygen minimum zones: stimulating interaction of aerobic archaeal and anaerobic bacterial ammonia oxidizers in a

- laboratory-scale model system. *Environ. Microbiol.* 14, 3146–3158. doi: 10.1111/j.1462-2920.2012.02894.x
- Zhang, C. L., Ye, Q., Huang, Z., Li, W., Chen, J., Song, Z., et al. (2008). Global occurrence of archaeal *amoA* genes in terrestrial hot springs. *Appl. Environ. Microbiol.* 74, 6417–6426. doi: 10.1128/AEM.00843-08
- Zhou, J., Bruns, M. A., and Tiedje, J. M. (1996). DNA recovery from soils of diverse composition. *Appl. Environ. Microbiol.* 62, 316–322.
- Zolghadr, B., Weber, S., Szabo, Z., Driessen, A. J., and Albers, S. V. (2007). Identification of a system required for the functional surface localization of sugar binding proteins with class III signal peptides in *Sulfolobus solfataricus*. *Mol. Microbiol.* 64, 795–806. doi: 10.1111/j.1365-2958.2007.05697.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Abby, Melcher, Kerou, Krupovic, Stieglmeier, Rossel, Pfeifer and Schleper. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.