



Genomic Insights of “*Candidatus Nitrosocaldaceae*” Based on Nine New Metagenome-Assembled Genomes, Including “*Candidatus Nitrosothermus*” Gen Nov. and Two New Species of “*Candidatus Nitrosocaldus*”

OPEN ACCESS

Edited by:

Olga V. Golyshina,
Bangor University, United Kingdom

Reviewed by:

Anne Daebeler,
University of Vienna, Austria
Sophie S. Abby,
UMR5525 Techniques de l'Ingénierie
Médicale et de la Complexité
Informatique, Mathématiques et
Applications, Grenoble (TIMC-IMAG),
France

*Correspondence:

Wen-Sheng Shu
shuwensheng@m.scnu.edu.cn
Wen-Jun Li
liwenjun3@mail.sysu.edu.cn

Specialty section:

This article was submitted to
Biology of Archaea,
a section of the journal
Frontiers in Microbiology

Received: 21 September 2020

Accepted: 09 December 2020

Published: 08 January 2021

Citation:

Luo Z-H, Narsing Rao MP, Chen H, Hua Z-S, Li Q, Hedlund BP, Dong Z-Y, Liu B-B, Guo S-X, Shu W-S and Li W-J (2021) Genomic Insights of “*Candidatus Nitrosocaldaceae*” Based on Nine New Metagenome-Assembled Genomes, Including “*Candidatus Nitrosothermus*” Gen Nov. and Two New Species of “*Candidatus Nitrosocaldus*”. *Front. Microbiol.* 11:608832. doi: 10.3389/fmicb.2020.608832

Zhen-Hao Luo¹, Manik Prabhu Narsing Rao¹, Hao Chen¹, Zheng-Shuang Hua^{1,2}, Qi Li¹, Brian P. Hedlund^{3,4}, Zhou-Yan Dong¹, Bing-Bing Liu⁵, Shu-Xian Guo⁵, Wen-Sheng Shu^{6*} and Wen-Jun Li^{1,5*}

¹ State Key Laboratory of Biocontrol, Guangdong Provincial Key Laboratory of Plant Resources and Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), School of Life Sciences, Sun Yat-sen University, Guangzhou, China, ² Department of Biological Sciences, Dartmouth College, Hanover, NH, United States, ³ School of Life Sciences, University of Nevada, Las Vegas, Las Vegas, NV, United States, ⁴ Nevada Institute of Personalized Medicine, University of Nevada, Las Vegas, Las Vegas, NV, United States, ⁵ Henan Key Laboratory of Industrial Microbial Resources and Fermentation Technology, College of Biological and Chemical Engineering, Nanyang Institute of Technology, Nanyang, China, ⁶ School of Life Sciences, South China Normal University, Guangzhou, China

“*Candidatus Nitrosocaldaceae*” are globally distributed in neutral or slightly alkaline hot springs and geothermally heated soils. Despite their essential role in the nitrogen cycle in high-temperature ecosystems, they remain poorly understood because they have never been isolated in pure culture, and very few genomes are available. In the present study, a metagenomics approach was employed to obtain “*Ca. Nitrosocaldaceae*” metagenomic-assembled genomes (MAGs) from hot spring samples collected from India and China. Phylogenomic analysis placed these MAGs within “*Ca. Nitrosocaldaceae*.” Average nucleotide identity and average amino acid identity analysis suggested the new MAGs represent two novel species of “*Candidatus Nitrosocaldus*” and a novel genus, herein proposed as “*Candidatus Nitrosothermus*.” Key genes responsible for chemolithotrophic ammonia oxidation and a thaumarchaeal 3HP/4HB cycle were detected in all MAGs. Furthermore, genes coding for urea degradation were only present in “*Ca. Nitrosocaldus*,” while biosynthesis of the vitamins, biotin, cobalamin, and riboflavin were detected in almost all MAGs. Comparison of “*Ca. Nitrosocaldales/Nitrosocaldaceae*” with other AOA revealed 526 specific orthogroups. This included genes related to thermal adaptation (cyclic 2,3-diphosphoglycerate, and S-adenosylmethionine decarboxylase), indicating their importance for life at high temperature. In addition, these MAGs acquired genes from members from archaea (Crenarchaeota) and bacteria (Firmicutes), mainly involved in metabolism and stress responses, which might play a role to allow this group to adapt to thermal habitats.

Keywords: “*Candidatus Nitrosocaldaceae*”, hot spring, ammonia-oxidizing archaea, genome comparison, horizontal gene transfer

INTRODUCTION

Geothermal springs represent a model system to study microbial ecology because of their simple community structure and simplified biogeochemical processes (Amend and Shock, 2001). This habitat serves as an important niche for diverse archaea with heterogeneous metabolic capabilities (Alves et al., 2018). Among the various modes of chemolithotrophy, ammonia oxidation can be a major energy source in some geothermal systems, due to relatively high concentrations of ammonia (Dodsworth et al., 2012). The global distribution of archaeal ammonia monooxygenase subunit A (*amoA*) in terrestrial geothermal springs (Zhang et al., 2008), and high rates of nitrification in some hot springs (Reigstad et al., 2008; Dodsworth et al., 2011), suggests a vital role for thermophilic ammonia-oxidizing archaea (AOA) in the nitrogen and carbon cycles in these habitats. “*Candidatus Nitrosocaldus*” members are the only archaea capable of oxidizing ammonia to nitrite at temperatures >65°C (de la Torre et al., 2008; Abby et al., 2018; Daebeler et al., 2018). Members of this family are globally distributed in neutral or slightly alkaline hot spring sediments or geothermally heated soils (Marteinsson et al., 2001; Nunoura et al., 2005; de la Torre et al., 2008; Zhang et al., 2008; Dodsworth et al., 2011; Hou et al., 2013; Hamilton et al., 2014; Abby et al., 2018).

“*Candidatus Nitrosocaldaceae*,” which includes “*Ca. Nitrosocaldus*,” was first proposed in 2008 (de la Torre et al., 2008) to include the first known member of the genus, “*Ca. N. yellowstonensis*.” So far, only three enrichment cultures of “*Ca. Nitrosocaldus*” have been reported: “*Ca. N. yellowstonensis*” from a Yellowstone National Park hot spring, “*Ca. N. islandicus*” from an Icelandic hot spring, and “*Ca. N. cavascurensis*” SCU2 from a hot spring in southern Italy (de la Torre et al., 2008; Abby et al., 2018; Daebeler et al., 2018). Like other AOA, all of them are chemolithoautotrophic and couple ammonia oxidation to carbon fixation using the 3-hydroxypropionate/4-hydroxybutyrate pathway (de la Torre et al., 2008; Abby et al., 2018; Daebeler et al., 2018). Besides ammonia oxidation, additional metabolic activities have been suggested from these three genomes, including aromatic amino acid fermentation, utilization of urea, nitrile, and hydrogen cyanide as alternative ammonia donors, and synthesis of vitamins (de la Torre et al., 2008; Abby et al., 2018; Daebeler et al., 2018). Despite these critical features, the family remains poorly understood because of the lack of pure cultures, limited availability of genomes, and small number of studies measuring their activities *in situ*.

In the present study, a metagenomics approach was employed to obtain “*Ca. Nitrosocaldaceae*” metagenomic-assembled genomes (MAGs) from hot spring samples collected from India and China. With the newly recovered MAGs and available AOA genomes, we have performed comparative genomic analyses to provide insight into the phylogeny, functional diversity, and adaptation mechanisms of this family and propose one novel genus and two novel species.

MATERIALS AND METHODS

Sampling, DNA Extraction, Sequencing, and Phylogenetic Analysis

Hot spring sediment samples were collected from India (Gujarat, Tumba: 22°47'58"N 73°27'37"E, Temp: 55°C) and China (Yunnan, Jinze, JZ: 25°26'28"N 98°27'36"E, Temp: 75°C; Qiaquan, QQ: 24°57'0"N 98°26'11"E, Temp: 69°C; Shuirebaozha, SRBZ: 24°57'0"N 98°26'14"E, Temp: 72°C). General descriptions for the Chinese hot springs have been described elsewhere (Hedlund et al., 2012; Hou et al., 2013).

Genomic DNA was extracted using the PowerSoil DNA isolation kit (MoBio). Metagenomic sequencing data were generated using the Illumina HiSeq 4000 instrument. The raw data were processed as described by Hua et al. (2015) and assembled using SPAdes v.3.14.0 (Bankevich et al., 2012). The sequence coverage was determined by BMAP v.36.77¹ and genome binning based on tetra-nucleotide frequencies, and sequencing depth was conducted with MetaBAT v.2.12.1 (Kang et al., 2015). The quality of the MAGs was evaluated using CheckM v.1.0.7 (Parks et al., 2015). Full-length 16S rRNA gene sequences were extracted from the MAGs using rnammer v.1.2 (Lagesen et al., 2007). “*Ca. Nitrosocaldaceae*” MAGs were preliminarily screened using PhyloPhlAn v.0.99 (Segata et al., 2013). Phylogenetic tree was constructed using a concatenated set of 122 universal marker proteins by GTDB-Tk (Parks et al., 2017). MUSCLE v.3.8.31 (Edgar, 2004) with default parameters used to align these sequences. Poorly aligned regions were removed using Trimal v.1.4 (Capella-Gutiérrez et al., 2009). The phylogeny was inferred with IQtree v.1.6.0 (Nguyen et al., 2014) by ultra-rapid bootstraps (1,000) with the parameters `-alrt 1000 -bb 1000 -nt AUTO`. Besides, nucleotide sequences of *amoA* gene were retrieved (Alves et al., 2018) and aligned with MAFFT v.7.407 (Katoh and Standley, 2013). Protein sequences coding for nitrite reductase (NirK), multicopper oxidases (Decleire et al., 2016; Kerou et al., 2016) and indolepyruvate ferredoxin oxidoreductase (Daebeler et al., 2018) were retrieved and aligned with MAFFT. The alignments were filtered using Trimal v.1.4 (Capella-Gutiérrez et al., 2009) to eliminate columns with ≥ 95% gaps, and all gene trees were inferred with IQtree v.1.6.0 (Nguyen et al., 2014). All trees were uploaded to iTOL v.4 (Letunic and Bork, 2019) for visualization.

Genome Annotation, Horizontal Gene Transfer Predictions, and Comparative Genomics

In order to gain a comprehensive insight into “*Ca. Nitrosocaldaceae*,” published “*Ca. Nitrosocaldaceae*” MAGs were obtained from NCBI² and JGI³. Protein-coding regions were identified using Prodigal v.2.6.3 (Hyatt et al., 2010) with the “-p single” option and annotated against the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al., 2016) and

¹<http://sourceforge.net/projects/bbmap/>

²<https://www.ncbi.nlm.nih.gov/>

³<https://img.jgi.doe.gov/cgi-bin/m/main.cgi>

Archaeal Clusters of Orthologous Genes (arCOG) (Makarova et al., 2015) database (e-value threshold = 10^{-5}) with DIAMOND (Buchfink et al., 2015). Assignments of key metabolic pathways and specific functions were manually verified based on the KEGG result and the online KEGG mapping tools⁴. Horizontal gene transfers (HGTs) were inferred using HGTector2 (Zhu et al., 2014) and visualized using SankeyMATIC⁵. For confirmation, we use candidate HGT protein sequences selected in our MAGs as queries against the protein sequences of NCBI genomes database available (downloaded in May 2020). Retrieved protein sequences together with protein sequences in “*Ca. Nitrosocaldaceae*” were then aligned using MAFFT, the alignments were filtered with TrimalAL and trees were constructed with IQtree and visualized in iTOL. All parameter sets were the same as mentioned above. For simplicity, the initial phylogenetic trees were used to choose a small set of sequences in the final trees. The rRNAs and tRNAs were predicted using RNAmmer v.1.2 (Lagesen et al., 2007) and tRNAscan-SE v.2.0.2, respectively (Lowe and Eddy, 1997). The average nucleotide identity (ANIm) values of MAGs were determined using pyani (Pritchard et al., 2016), and the average amino acid identity (AAI) values were determined using CompareM⁶. For further comparative analysis, MAGs with completeness $\geq 80\%$ were selected. OrthoFinder v.2.4.0 (parameters “-s blast_gz -t 20 -f protein_files”) with BLASTP (-e 0.001) was used to assess orthology among the coding sequences (CDS) of all selected genomes (Emms and Kelly, 2015, 2019). The term “shared set” was defined for orthogroups that were present in a given branch and other branches, whereas the “specific set” refers to orthogroups that were only present in a given branch. Then, further sub-defining the “specific set,” the term “lineage-core set” was used for orthogroups identified in $\geq 90\%$ of the analyzed genomes, and the term “lineage-accessory set” was used for orthogroups that were present in $< 90\%$ of the analyzed genomes. Detailed information on published AOA used in this study is provided in **Supplementary Table S1**.

RESULTS AND DISCUSSION

General Features of “*Ca. Nitrosocaldaceae*” MAGs

In total, nine “*Ca. Nitrosocaldaceae*” MAGs were obtained from Chinese and Indian hot spring sediments. Among them, four MAGs were obtained from Jinze (JZ-1.bins.77, JZ-2.bins.172, JZ-2.bins.249, and JZ-3.bins.102), three from Qiaquan (QQ.bins.88, QQ.bins.97, and QQ.bins.115137), one from Shuirebaozha (SRBZ.bins.174), and one from Tumba (Tumba.bins.72).

In addition to the two cultivated members of the genus “*Ca. Nitrosocaldus*” (with genomic data), “*Ca. N. cavascurensis* SCU2” (Abby et al., 2018) and “*Ca. N.*

islandicus” (Daebeler et al., 2018), a MAG (Thaumarchaeota archaeon J079) obtained from a Japanese hot spring (Ward et al., 2019) was also assigned to this family based on phylogenetic analysis. In phylogenetic trees based on alignments of 122 archaeal marker genes (**Figure 1A**), a monophyletic group containing the two cultivated members of “*Ca. Nitrosocaldus*” and the nine additional MAGs were recovered.

Phylogenetic analysis (**Figure 1A**) showed that the twelve MAGs formed four well-supported clades within the family “*Ca. Nitrosocaldaceae*.” Clade A consisted of JZ-2.bins.172, QQ.bins.88, and Thaumarchaeota archaeon J079. Clade B corresponded to QQ.bins.115137 and SRBZ.bins.174. Clade C included Tumba.bins.72 and two recently proposed “*Ca. Nitrosocaldus*” species, “*Ca. N. islandicus*” and “*Ca. N. cavascurensis* SCU2.” Clade D comprised of JZ-2.bins.249, JZ-3.bins.102, JZ-1.bins.77, and QQ.bins.97.

AAI values suggested that clades B, C, and D can be ascribed to the genus “*Ca. Nitrosocaldus*,” as their AAI values (**Supplementary Table S2** and **Supplementary Figure S1A**) were within the proposed limits (65–95%) (Konstantinidis et al., 2017). In contrast, AAI values comparing Clade A genomes to “*Ca. Nitrosocaldus*” genomes were well below those values (56.6–57.8%). In addition, other genomic characteristics consistently separated Clade A from “*Ca. Nitrosocaldus*,” reflected by GC contents (Clade A, 31.9–33.1%, average 32.4%, $n = 3$; “*Ca. Nitrosocaldus*,” 41.5–42.8%, average 42.0%, $n = 9$, **Figure 1B**) and overall KO profiles (**Figure 1C**, ADONIS $R^2 = 0.48$, $P < 0.01$).

Therefore, we propose the name “*Candidatus Nitrosothermus koennekii*” gen. nov., sp. nov. to circumscribe the organisms represented by clade A.

Etymology: Ni.tro.so.ther’mus. L. masc. adj. *nitrosus*, full of natron, here intended to mean nitrous; Gr. masc. adj. *thermos*, hot; N.L. masc. n. *Nitrosothermus*, a nitrite-forming organism from a hot environment; koen.ne’ke.i. N.L. gen. n. *koennekei*, named in honor of Martin Köneke. The type genome is JZ-2.bins.172^T.

The ANI values (**Supplementary Table S3** and **Supplementary Figure S1B**) separating clades B, C, and D were below the cut-off (95–96%) for species identification (Jain et al., 2018), and we propose the names “*Candidatus Nitrosocaldus tengchongensis*” sp. nov. and “*Candidatus Nitrosocaldus schleperae*” sp. nov. for clade B and clade D, respectively. The type genomes are QQ.bins.115137^T for “*Candidatus Nitrosocaldus tengchongensis*” and JZ-1.bins.77^T for “*Candidatus Nitrosocaldus schleperae*.”

Etymology: teng.chong.en’sis. N.L. masc. adj. *tengchongensis* pertaining to Tengchong, Yunnan Province, south-west China, where the type strain was isolated; schle’pe.rae. N.L. gen. n. *schleperae*, named in honor of Christa Schleper.

Despite the wide geographic range of clade C (Italy, Iceland, and India) and the previous descriptions of two “*Ca. Nitrosocaldus*” species, “*Ca. N. cavascurensis*” (Abby et al., 2018) and “*Ca. N. islandicus*” (Daebeler et al., 2018), the high ANI values between members of clade C (98.41–99.80%) (**Supplementary Table S3**) suggests clade C as a single species group.

⁴<https://www.genome.jp/kegg/kegg1b.html>

⁵<http://sankeymatic.com/>

⁶<https://github.com/dpark1134/CompareM>

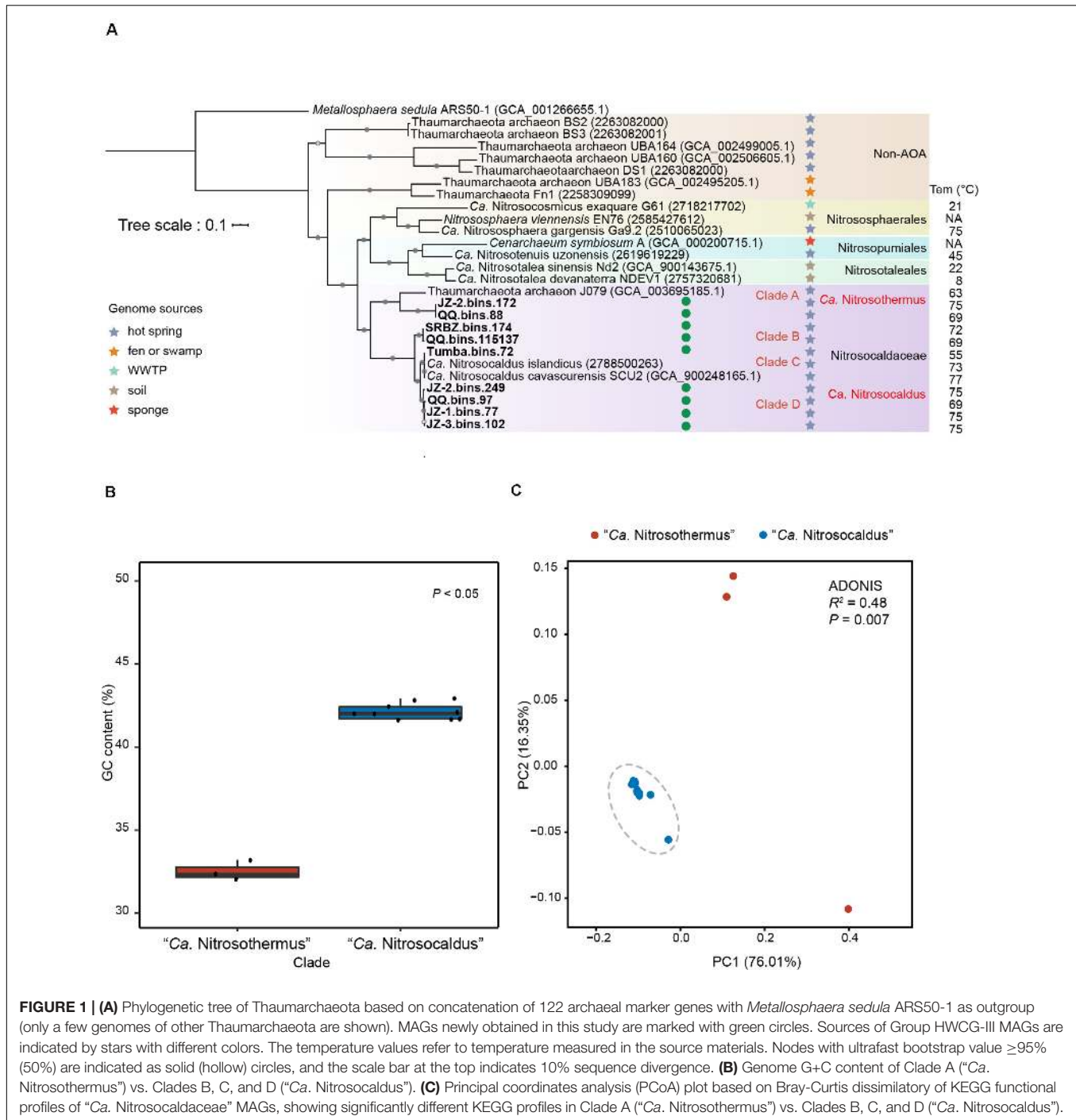


FIGURE 1 | (A) Phylogenetic tree of Thaumarchaeota based on concatenation of 122 archaeal marker genes with *Metallosphaera sedula* ARS50-1 as outgroup (only a few genomes of other Thaumarchaeota are shown). MAGs newly obtained in this study are marked with green circles. Sources of Group HWCG-III MAGs are indicated by stars with different colors. The temperature values refer to temperature measured in the source materials. Nodes with ultrafast bootstrap value $\geq 95\%$ (50%) are indicated as solid (hollow) circles, and the scale bar at the top indicates 10% sequence divergence. **(B)** Genome G+C content of Clade A ("Ca. Nitrosothermus") vs. Clades B, C, and D ("Ca. Nitrosocaldus"). **(C)** Principal coordinates analysis (PCoA) plot based on Bray-Curtis dissimilarity of KEGG functional profiles of "Ca. Nitrosocaldaceae" MAGs, showing significantly different KEGG profiles in Clade A ("Ca. Nitrosothermus") vs. Clades B, C, and D ("Ca. Nitrosocaldus").

The estimated genome sizes were small (1.26–1.56 Mb), with estimated completeness of 58.8–100% and estimated contamination of 0–3.4%. Most MAGs have detectable 16S rRNA genes (except Thaumarchaeota archaeon J079) and more than 18 tRNAs (except JZ-2.bin.249), indicating that they were well curated and appropriate for further analysis. The number of coding sequences ranged from 857 to 1,767, with about half of them being assignable by KO profiles. The detailed features of these MAGs are given in **Table 1**.

Metabolic Potential of "Ca. Nitrosocaldaceae" Nitrogen Metabolism

Similar to other AOA (Walker et al., 2010; Spang et al., 2012), the key genes responsible for ammonia oxidation encoded by *amoABC* were present in all "Ca. Nitrosocaldaceae" MAGs (**Figure 2**, **Supplementary Figure S2**, and **Supplementary Table S4**).

TABLE 1 | Genomic features of "Candidatus Nitrosocaldaceae".

Genome feature	1	2	3	4	5	6	7	8	9	10	11	12
No. of scaffolds	434	38	34	1	1	12	56	196	111	88	16	26
Genome size (Mbp)	1.06	1.51	1.39	1.50	1.54	1.50	1.18	1.04	0.63	1.39	1.39	1.39
Estimated Genome Size (Mbp)	1.26	1.51	1.43	1.52	1.56	1.52	1.29	1.40	1.08	1.41	1.40	1.41
Genomic G+C content (%)	33.1	32.2	31.9	41.6	41.5	41.5	42.8	42.7	42.3	41.9	42	41.9
N_{50} value (bp)	3,037	1,56,292	1,85,164	15,77,284	16,17,394	2,23,702	29,463	6,454	6,672	29,350	1,23,923	1,02,733
No. of protein coding genes	1,569	1,767	1,648	1,718	1,753	1,729	1,431	1,457	857	1,700	1,665	1,646
No. of rRNAs	1	2	2	2	2	2	2	2	1	2	2	2
Presence of 16S rRNA gene	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. of tRNAs	32	40	37	36	36	41	31	30	14	36	38	39
No. of genes annotated by KO	811 (51.7%)	877 (49.6%)	819 (49.7%)	832 (48.4%)	836 (47.6%)	840 (48.5%)	746 (52.1%)	751 (51.5%)	488 (56.9%)	849 (49.9%)	818 (49.1%)	812 (49.3%)
Completeness ^a (%)	84.43	100	97.5	99	99	99	91.9	74.7	58.8	98	99	98
Contamination ^a (%)	2.5	0.9	0	0	0	0.9	0	0	3.4	0	0	0
Accession number	GCA_003695185.1 ^b	GCA_013538795.1 ^b	GCA_013538715.1 ^b	GCA_900248165.1 ^b	2788500263 ^c	GCA_011058825.1 ^b	GCA_013538675.1 ^b	GCA_013538755.1 ^b	GCA_013538775.1 ^b	GCA_013538695.1 ^b	GCA_013538805.1 ^b	GCA_013538705.1 ^b

1, *Thaumarchaeota* archaeon J079; 2, JZ-2.bins.172; 3, QQ.bins.88; 4, "Ca. Nitrosocaldus cavascurensis"; 5, "Ca. Nitrosocaldus islandicus"; 6, Tumba.bins.72; 7, QQ.bins.115137; 8, SRBZ.bins.174; 9, JZ-2.bins.249, 10, JZ-3.bins.102; 11, JZ-1.bins.77; 12, QQ.bins.97. ^aThe evaluations of genome completeness and contamination were conducted with CheckM. Although the genomes of "Ca. Nitrosocaldus islandicus" and "Ca. Nitrosocaldus cavascurensis" are only 99% complete in the table, these are, in fact, closed genomes that are missing PF000395. ^bThe GenBank accession number. ^cThe JGI accession number.

Ammonia oxidation not only yields free energy but also reducing force for anabolic reactions. Previously, an additional copy of *amoC* that may participate in an ammonia starvation and stress response was observed in some AOA MAGs (Berube et al., 2007; Berube and Stahl, 2012; Spang et al., 2012). In our study, an extra *amoC* could be found in two “Ca. Nitrosothermus”

MAGs (**Supplementary Table S4**), indicating that they may grow at a lower concentration of ammonia than other “Ca. Nitrosocaldaceae” members. Interestingly, a full-length extra *amoA* gene (NODE_52239_length_4113_cov_3.91646_6) was detected in MAG JZ-2.bins.249. The sequence assembly of the extra *amoA*-containing scaffold

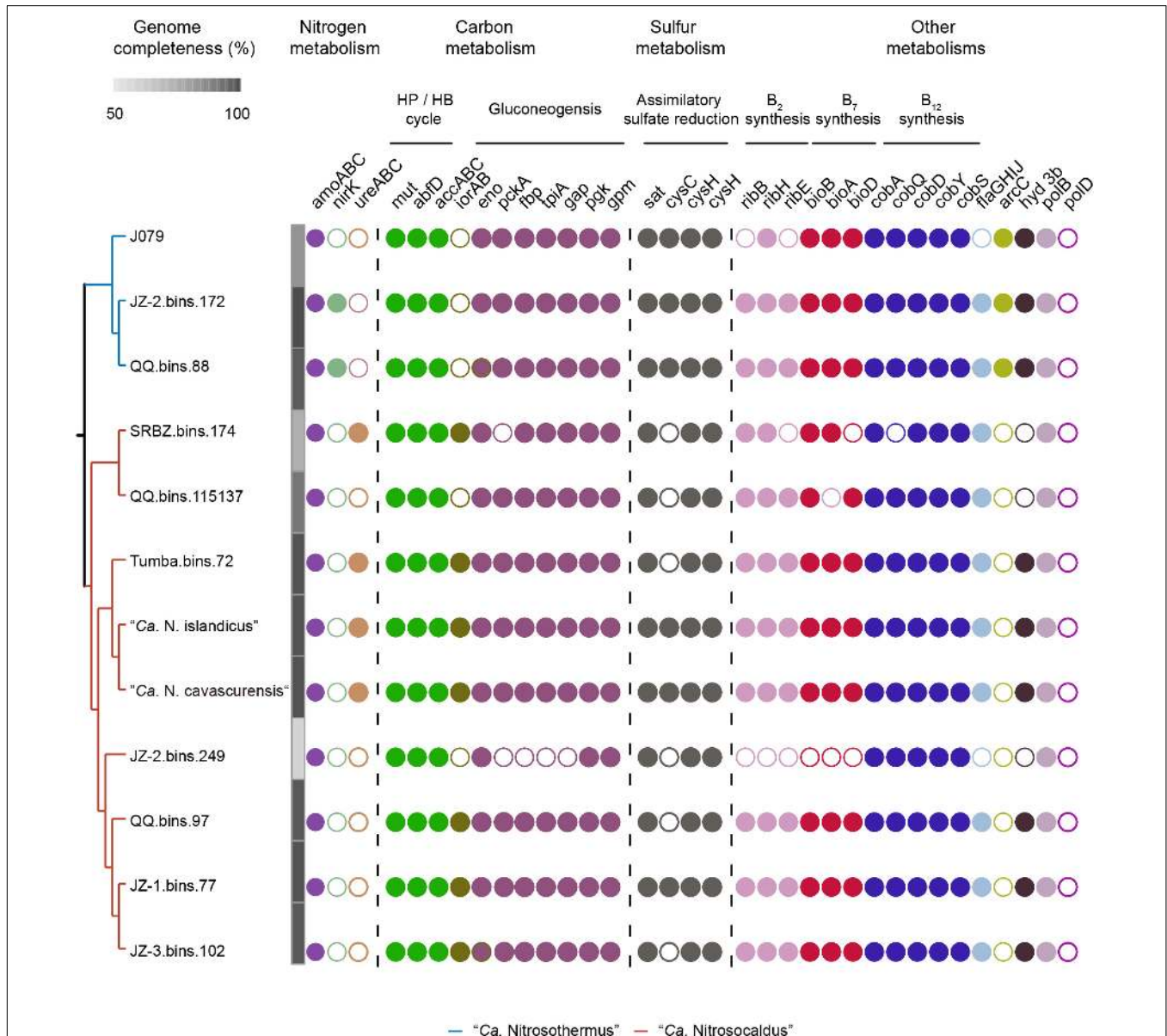


FIGURE 2 | The overall distributions of the genes of interest in “Ca. Nitrosocaldaceae.” The phylogenetic tree in the left was pruned according to **Figure 1**. For visualization, the branch length is ignored. The two genera were colored according to **Figure 1**. MAGs of “Ca. Nitrosothermus” are marked with blue while MAGs of “Ca. Nitrosocaldus” with red. The solid and hollow circles represent the presence and absence of the genes. J079 Thaumarchaeota archaeon J079, HP/HB cycle hydroxypropionate/hydroxybutyrate cycle, *amoABC* ammonia monooxygenase subunit A, B, and C, *nirK* nitrite reductase, *ureABC* urease subunit gamma, beta and alpha, *mut* cobalamin-dependent methylmalonyl-CoA mutase, *abfD* 4-hydroxybutyryl-CoA dehydratase, *accABC* biotin-dependent acetyl-CoA/propionyl-CoA carboxylase, *iorAB* indolepyruvate ferredoxin oxidoreductase subunit alpha and beta, *pckA* phosphoenolpyruvate carboxykinase, *fbp* fructose 1,6-bisphosphate aldolase/phosphatase, *tpiA* triosephosphate isomerase, *gap*, glyceraldehyde-3-phosphate dehydrogenase, *pgk* phosphoglycerate kinase, *gpm* phosphoglycerate mutase, *eno* enolase, *sat* sulfate adenyllyltransferase, *cysC* adenyllylsulfate kinase, *cysH* phosphoadenosine phosphosulfate reductase, *sir* sulfite reductase (ferredoxin), *flaGHIJ* archaeal flagellar protein FlaG, FlaH, FlaI, and FlaH, *arcC* carbamate kinase, hyd 3b group 3b hydrogenase, *polB* archaea type DNA polymerase, *polD* D-family polymerase. The full name of genes involved in B₂ (riboflavin), B₇ (biotin), and B₁₂ (cobalamin) syntheses are listed in **Supplementary Table S4**.

(NODE_52239_length_4113_cov_3.91646) seems to be reliable, as the GC content and read depth of this scaffold was similar to other scaffolds in this bin (**Supplementary Figure S3**). All the closest protein sequences of this scaffold showed high similarities to “*Ca. Nitrosocaldus cavascurensis*” (**Supplementary Table S5**). However, the additional *amoA* gene was near the end of the scaffold and the gene content and orientation between the two scaffolds were highly identical. Besides, the extra *amoA* genes was found in the most contaminated MAG (contamination: 3.4%) but these two copies of *amoA* genes were very closely related (100% in query coverage and 99% in amino acid identity, **Figure 3**). We conclude that the placement of the second *amoA*

gene copy into the MAG might be due to strain diversity, which lead to incorrect assignment of two very similar contigs into the same MAG. Therefore, we assume that the species represented by the MAG were likely to harbor only one *amoA* gene. In addition, all *amoA* genes in “*Ca. Nitrosocaldaceae*” were placed in the NC-clade (**Figure 3**), which only occupies a small fraction of all known *amoA* sequences (Alves et al., 2018), stressing the limited research on this family.

It is well-known that ammonia can be imported into the cell via passive diffusion (Winkler, 2006). The detection of an AMT family ammonia transporter in most MAGs makes it possible to actively transport ammonia into cells when the permeability of

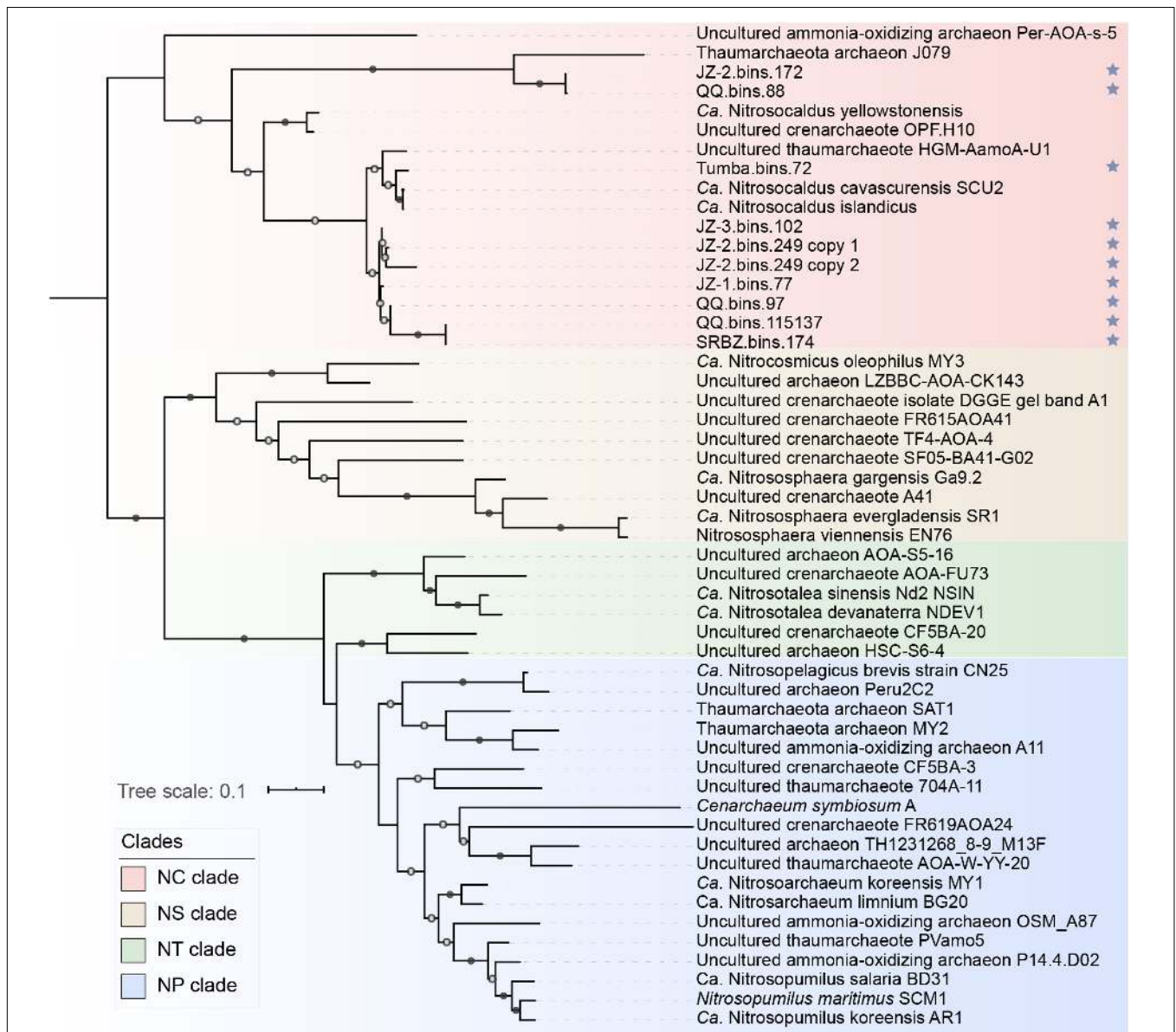


FIGURE 3 | Phylogenetic analysis of *amoA* gene sequences. Sequences with blue stars were newly obtained in this study. Classifications of clades were based on Alves et al. (2018). Only nodes with ultrafast bootstrap values $\geq 95\%$ (50%) were marked with solid (hollow) circles, and the scale bar at the bottom indicates 10% sequence divergence.

the cytoplasmic membrane becomes low, or the concentrations of ammonia were relatively low, as passive diffusion would be limited in these conditions (Winkler, 2006). It has been reported that multiple copper oxidases may play some roles in hydroxylamine oxidation, the second step of ammonia oxidation (Walker et al., 2010). In our study, genes encoding multicopper oxidases (MCO) were identified in some MAGs (**Supplementary Table S4**), which indicates that alternative mechanisms may exist for hydroxylamine oxidation. It has been suggested that NirK may provide NO during the NO-dependent dehydrogenation of hydroxylamine to nitrite (Kozłowski et al., 2016).

Homologs for NirK were found in two MAGs (JZ-2.bins.172 and QQ.bins.88). Interestingly, these MAGs harbor two copies of this protein (**Supplementary Table S4**). The phylogenetic analysis revealed that only one NirK (NODE_3_length_185164_cov_12.5162_189 in QQ.bins.88 and NODE_3_length_165040_cov_5.28639 in JZ-2.bins.172) in each MAG was placed adjacent to the previously defined archaeal NirK branch (**Figure 4**; Kerou et al., 2016). An alignment with these archaeal NirK proteins showed that type-1 and type-2 copper centers were conserved in these two newly recovered proteins (**Supplementary Figure S4**). Consistent with the previous study, they encode a potential transmembrane domain with a signal peptide at N-terminus (Bartossek et al., 2010). After careful phylogenetic analysis, we claim that these two proteins were genuine copper-dependent nitrite reductases with functions similar to characterized NirK instead of homologous MCO. The other homologous proteins clustered with multicopper oxidase sequences (**Figure 4A**), indicating they likely encode multicopper oxidases instead of NirK. Six MCOs belong to lineage 1 (MCO1, **Figure 4A** and **Supplementary Table S6**), which were from four “*Ca. Nitrosocaldus*” MAGs and two “*Ca. Nitrosothermus*” MAGs, while twenty-six MCOs belong to lineage 4 (MCO4, **Figure 4A** and **Supplementary Table S6**) were sourced from all MAGs. Notably, in one MAG (QQ.bins.115137), one ZIP family permease was detected next to MCO1. This pair of genes plays a role in copper sequestration (Kerou et al., 2016). It was assumed that MCO1 and MCO4 were involved in Cu uptake as they could carry out the oxidation of Cu^+ to Cu^{2+} , which could be transported into the cells via the ZIP family permease or a divalent transporter (Reyes et al., 2020). MCO could aid “*Ca. Nitrosocaldaceae*” not only in ammonia oxidation but also copper homeostasis.

Urea degradation provides AOA with ammonia to fuel ammonia oxidation and carbon dioxide for carbon fixation (Walker et al., 2010; Spang et al., 2012). In our study, only four MAGs belonging to “*Ca. Nitrosocaldus*” possess genes coding for urease and urea transporters, which suggests that among “*Ca. Nitrosocaldaceae*” urea utilization was unique in this genus. Consistent with a previous report (Daebeler et al., 2018), all members in “*Ca. Nitrosocaldaceae*” can potentially utilize hydrogen cyanide and nitrile, which were available in geothermal systems (Miller and Urey, 1959; Schulte and Shock, 1995). Nitrilases catalyze the degradation of nitrile, and cyanide hydratases convert HCN to formamide, and both can produce ammonia (Lenza and Vasconcelos, 2001; Pace and Brenner, 2001) to support ammonia oxidation. The

capability to utilize additional nitrogen-containing substrates could provide an advantage when ammonia was limited. The presence of *glnA* and *gdnA* genes coding for glutamine synthetase and glutamate dehydrogenase suggests that “*Ca. Nitrosocaldaceae*” could generate glutamate via ammonia assimilation.

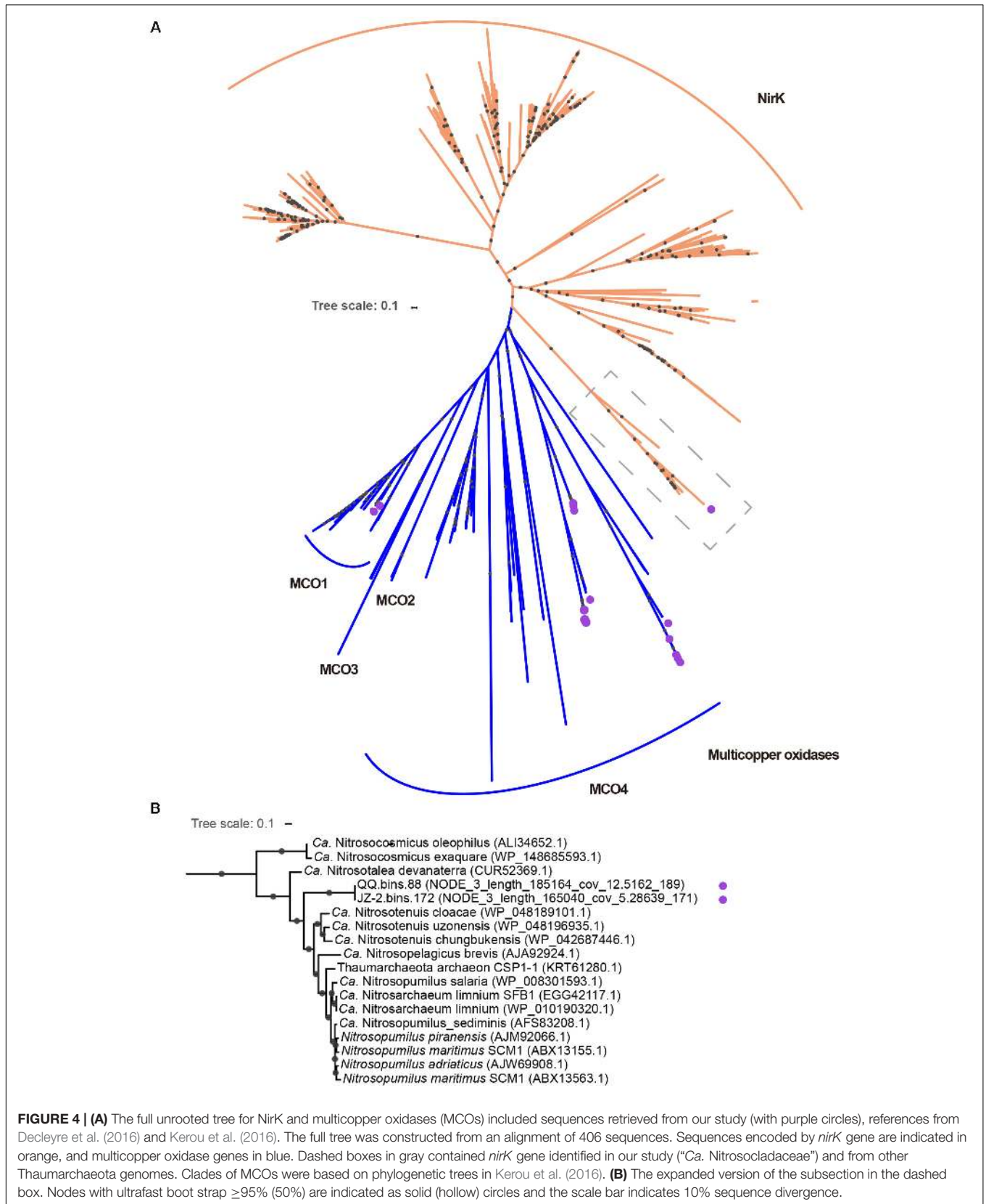
Carbon Metabolism

The key enzymes of the thaumarchaeal 3HP/4HB cycle were detected in all MAGs, indicating the potential for carbon fixation and the most energy-efficient pathway that could generate acetyl-CoA for biosynthesis (Könneke et al., 2014). The detection of an incomplete Embden-Meyerhof-Parnas pathway (EMP), genes coding for phosphoenolpyruvate carboxykinase, and fructose 1,6-bisphosphate aldolase in most MAGs support gluconeogenesis and sugar production (**Figure 2**, **Supplementary Figure S2**, and **Supplementary Table S4**). Except for some low-completeness MAGs, members of this family harbor an oxidative TCA cycle, which could support free-energy yielding reactions with organic substrates and/or provide some important intermediates for biosynthesis (e.g., oxaloacetic acid and α -ketoglutaric acid). Additionally, genes coding for 4-oxalocrotonate tautomerase, essential for the metabolism of aromatic compounds, was detected in all MAGs, indicating that they could generate intermediates for the TCA cycle via conversion of aromatic compounds. The detection of the pentose phosphate pathway and 5-phospho- α -D-ribose 1-diphosphate PRPP biosynthesis in this group suggests they can generate some nucleotide and amino acid precursors.

Furthermore, “*Ca. Nitrosocaldus*” might have the potential for aromatic amino acid fermentation due to the presence of *iorAB*, encoding for indolepyruvate ferredoxin oxidoreductase, which was not common among AOA (Daebeler et al., 2018).

Inferred from the phylogenetic tree (IorB protein sequences), they form a monophyletic clade, which was placed between bacterial IorB proteins, and a Crenarchaeota-Euryarchaeota cluster (**Supplementary Figure S5**). Besides, the presence of hydrogenases could regenerate oxidized ferredoxin that reduced during this process using hydrogen as an energy source, as proposed before (Daebeler et al., 2018). Furthermore, several enzymes, including aspartate aminotransferase argininosuccinate synthase and argininosuccinate lyase, were detected in “*Ca. Nitrosocaldaceae*” MAGs, suggesting a potential anaplerotic contribution of amino acids to the carbon and energy flow. Moreover, genes *pepP* and *pepN*, coding for aminopeptidases, were found in most of the MAGs together with the ABC-type peptidase/nickel transporter system and *iorAB*, suggesting that members of “*Ca. Nitrosocaldaceae*” could assimilate amino acids for energy or intermediates for biosynthesis.

Another enzyme detected in Thaumarchaeota archaeon J079 was carbonic anhydrase, catalyzing the conversion between CO_2 and bicarbonate, indicating that it could provide this MAG with bicarbonate, the substrate for the HP/HB cycle (Kerou et al., 2016). The capability of carbon fixation combined with the detection of various pathways involved in organic carbon metabolism suggests metabolic versatility of “*Ca. Nitrosocaldaceae*” as reported for other Thaumarchaeota



(Tourna et al., 2011; Lehtovirta-Morley et al., 2014), providing a hint for mixotrophy in “Ca. Nitrosocaldaceae.” However, these metabolisms require verification.

Oxidative Phosphorylation

The aerobic lifestyle of “Ca. Nitrosocaldus” has previously been reported in enrichment cultures (de la Torre et al., 2008; Abby et al., 2018; Daebeler et al., 2018). Except for one “Ca. Nitrosothermus” MAG (Thaumarchaeota archaeon J079), the universal presence of cytochrome c-type terminal oxidase (Complex III) in “Ca. Nitrosocaldaceae” MAGs confirmed that they could use oxygen as a terminal electron acceptor. Additionally, Complex I (NADH: ubiquinone oxidoreductase), II (succinate: quinone oxidoreductase), IV (cytochrome c oxidase), and V (ATPase) of the respiratory chain were detected in all MAGs.

Other Metabolisms

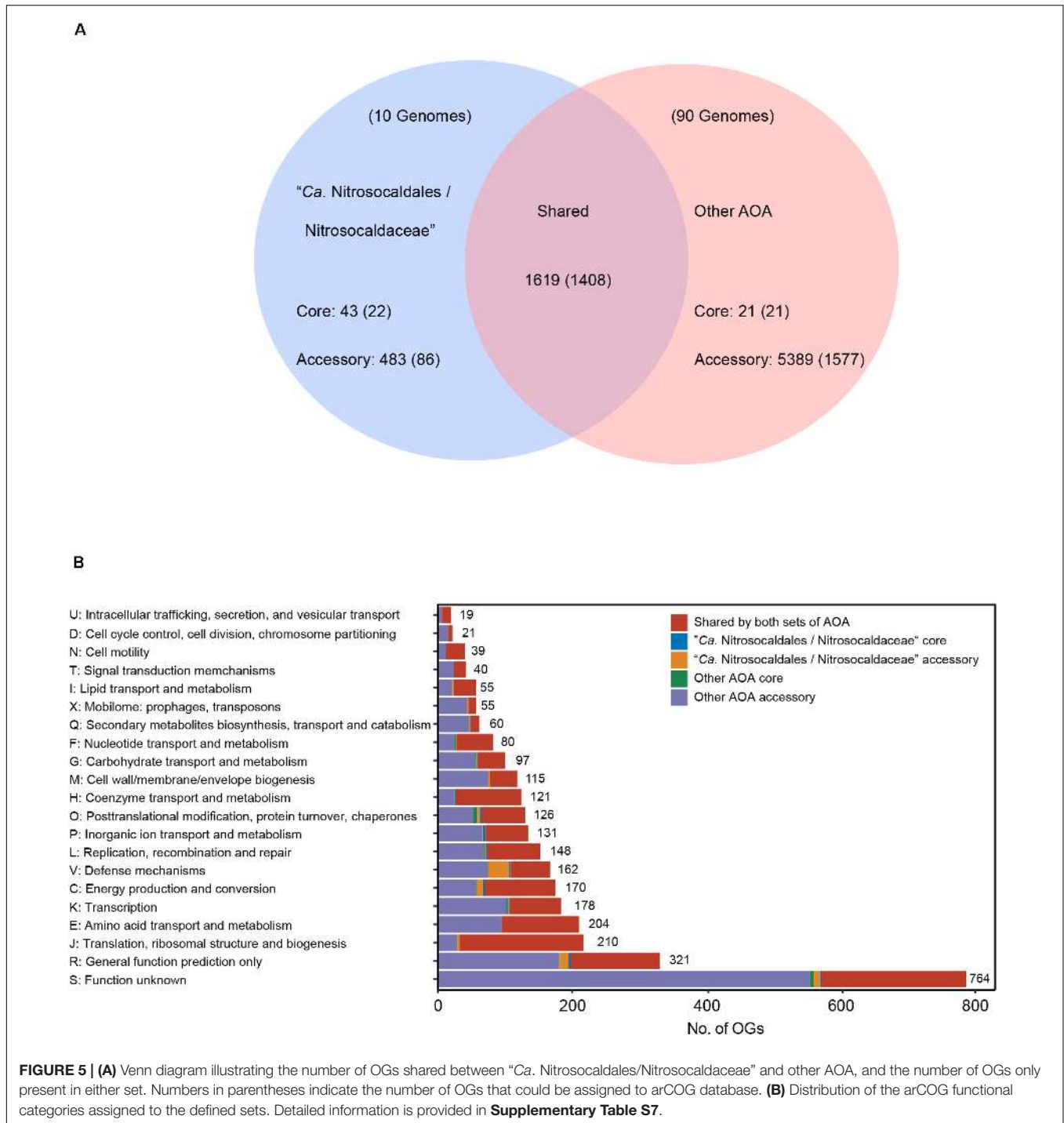
Several vitamins, such as biotin, cobalamin, and riboflavin were essential cofactors that were required for different enzymes (Berg et al., 2007; Mansoorabadi et al., 2007; Kim and Winge, 2013; Chow et al., 2018), including biotin-dependent acetyl-CoA/propionyl-CoA carboxylase and cobalamin-dependent methylmalonyl-CoA mutase, which were part of the HP/HB cycle (Ishii et al., 1996). Key enzymes involved in the biosynthesis of these cofactors were identified in most of the “Ca. Nitrosocaldaceae” MAGs (Figure 2, Supplementary Figure S2, and Supplementary Table S4). Homologs of these genes were also reported in some other Thaumarchaeota lineages (Doxey et al., 2015; Santoro et al., 2015; Heal et al., 2018). The only exception was JZ-2_bins.249, whose completeness was only 58.8%, so we attributed the lack of these genes in this MAG due to genome incompleteness. The genetic capacity for *de novo* synthesis of these cofactors was conserved in “Ca. Nitrosocaldaceae,” indicating that cofactor synthesis could be an important service for terrestrial microbial communities, as these cofactors could be synthesized by some but not all prokaryotes (Zhang et al., 2009; Swithers et al., 2012). Other genes (*sat*, *cysH*, and *sir*) involved in assimilatory sulfate reduction were found in these MAGs, indicating they can conduct this pathway. The products could be used for amino acid synthesis. In contrast, other “Ca. Nitrosocaldaceae” MAGs cannot reduce sulfate in this way as they lack CysC, which was responsible for converting adenylyl sulfate to 3'-phosphoadenylyl sulfate. Motility has been reported in some AOA (Lehtovirta-Morley et al., 2016). In our study, most MAGs encode genes responsible for archaeal flagellar proteins (*flaG*, *flaH*, *flaI*, and *flaJ*). The only exceptions were Thaumarchaeota archaeon J079 and JZ-2_bins.249, which has low completeness (84.4 and 58.8%). Therefore, motility may be a common feature of “Ca. Nitrosocaldaceae,” which could be advantageous for responding to environmental conditions (Blainey et al., 2011). Although phosphate transporter-related genes were commonly found, alkaline phosphatase was only found in “Ca. Nitrosothermus” MAGs, which might confer an advantage in phosphorus-limiting environments (Shen et al., 2016). Moreover, *arcC*, coding for carbamate kinase

and *carB*, coding for carbamoyl-phosphate synthase were only detected in “Ca. Nitrosothermus” (Supplementary Table S4). These two enzymes could potentially provide ammonia under aerobic conditions (Abdelal et al., 1982), which could be valuable given the low concentration of ammonia in these springs (Hedlund et al., 2012; Hou et al., 2013; Ward et al., 2019) and many other alkaline geothermal springs (Holloway et al., 2011). However, the potential function of these genes in ammonia production in “Ca. Nitrosothermus” deserves more attention.

Comparison Between “Ca. Nitrosocaldales/Nitrosocaldaceae” and Other AOA

To gain insight into potential mechanisms enabling thermophily in “Ca. Nitrosocaldales/Nitrosocaldaceae,” comparative genomics of all selected AOA was conducted.

The term “shared set” was defined for orthogroups that were present in a given branch and other branches, whereas the “specific set” refers to orthogroups that were only present in a given branch. Then, further sub-defining the “specific set,” the term “lineage-core set” was used for orthogroups identified in $\geq 90\%$ of the analyzed genomes, and the term “lineage-accessory set” was used for orthogroups that were present in $< 90\%$ of the analyzed genomes. A total of 179,946 coding sequences (CDS) of all selected AOA MAGs were clustered into 15,802 OGs, with 8,247 classified as singletons. Comparisons of OGs between “Ca. Nitrosocaldaceae” and other AOA enabled us to identify a shared set of 1,619 protein families, and a large number of these OGs could be assigned to information-processing genes (replication, transcription, translation) (Figure 5 and Supplementary Table S7). Among them, all genes of the highly conserved central information-processing machinery in Thaumarchaeota were detected (Spang et al., 2010), including ribosomal proteins (S25, S26, S30, L13e, L29), other proteins involved in translation (RNA polymerase subunit B, transcription factor MBF1), topoisomerases IB, proteins involved in cell division (Cell division ATPase of the AAA⁺ class ESCRT system component C, cell division GTPase FtsZ, chromatin segregation and condensation protein ScpA and ScpB), histones H3 and H4, and proteins involved in the repair of macromolecules (ERCC4-like helicase, ERCC4-type nuclease, chaperone DnaK, molecular chaperon GrpE). As discussed above, we have detected genes participating in the thaumarchaeal 3HP/4HB pathway, TCA cycle, gluconeogenesis, non-oxidative PPP, ammonia oxidation, and biosynthesis of several cofactors (riboflavin, biotin, and cobalamin) in the shared OGs (Supplementary Table S7). The presence of these genes in “Ca. Nitrosocaldales/Nitrosocaldaceae”-core and other AOA-core sets confirms the conserved metabolic feature of AOA, which has been noted in previous study (Kerou et al., 2016). It was reported that the genomes of “Ca. Nitrosocaldus islandicus” and “Ca. Nitrosocaldus cavascurensis” encode a family B polymerase, and not family D polymerases (Abby et al., 2018; Daebeler et al., 2018). In our study, we found genes coding for PolB (OG0000200) in the shared set, together with a family Y DNA



polymerase (OG0000329). Consistently, both subunits of PolD (OG0000966 and OG0001021) were detected in the “other AOA-accessory” set, but not in “*Ca. Nitrosocaldales/Nitrosocaldaceae*” (**Supplementary Table S7** and **Figure 2**). PolB was therefore assumed to be the main replicative polymerase in “*Ca. Nitrosocaldales/Nitrosocaldaceae*” as in most thermophilic Thermoproteales, also lack PolD (Barry and Bell, 2006). The absence of PolD in “*Ca. Nitrosocaldales/Nitrosocaldaceae*”

indicates distinct mechanisms of DNA replication exist in “*Ca. Nitrosocaldales/Nitrosocaldaceae*” and other AOA. It might also be possible that DNA primases play a vital role, for example in lagging-strand synthesis, as several genes coding for DNA primases (OG0000234, OG0000417, and OG0000456) were detected in the shared set (**Supplementary Table S7**), and some archaeal primases were known to have polymerase activity (Lao-Sirieix and Bell, 2004).

The differences between the functional profiles of “*Ca. Nitrosocaldales/Nitrosocaldaceae*” and those of other AOA were relatively small but significant (**Supplementary Figure S6**, ADONIS $R^2 = 0.11$, $P = 0.001$). In total, 526 OGs were only present in “*Ca. Nitrosocaldales/Nitrosocaldaceae*,” which we therefore defined as “*Ca. Nitrosocaldales/Nitrosocaldaceae*”-specific OGs, and 5,410 OGs were only present in other AOA (other AOA-specific OGs). Among the 526 OGs, 43 of them could be classified as the core set of “*Ca. Nitrosocaldales/Nitrosocaldaceae*” (**Supplementary Table S7**). Notably, the gene coding for cyclic 2,3-diphosphoglycerate (cDPG) synthetase was found in this set, indicating that cDPG utilization might be a unique strategy used by “*Ca. Nitrosocaldales/Nitrosocaldaceae*” in hot springs (Abby et al., 2018) as it was reported that cDPG might be involved in the thermo-stabilization of proteins (Hensel and König, 1988; Shima et al., 1998). Additionally, cDPG could act as a storage reservoir for energy, phosphorus, or carbon. For example, cDPG could potentially help “*Ca. Nitrosocaldales/Nitrosocaldaceae*” survive during phosphate limitation, as reported in other archaea (Roberts, 2005). The detection of S-adenosylmethionine decarboxylase in this set suggests polyamine production, which was vital in cellular tolerance to high temperature (Tabor and Tabor, 1984; Morgan et al., 1987). Additionally, we predict that “*Ca. Nitrosocaldales/Nitrosocaldaceae*” can specifically export L-alanine to avoid the accumulation of amino acids to toxic levels (Kim et al., 2015), indicated by the presence of gene coding for L-alanine export (AlaE) in all “*Ca. Nitrosocaldales/Nitrosocaldaceae*” MAGs (**Supplementary Tables S4, S7**). The L-alanine could be generated in “*Ca. Nitrosocaldales/Nitrosocaldaceae*” MAGs (**Supplementary Table S4**) via cysteine desulfurase/selenocysteine lyase from selenocysteine or coupling alanine aminotransferase with glutamate dehydrogenase from pyruvate, as proposed in *Pyrococcus furiosus* (Kengen and Stams, 1994). However, alternative mechanisms of L-alanine accumulation might exist as the distribution of alanine aminotransferase was limited in “*Ca. Nitrosocaldales/Nitrosocaldaceae*” (totally absent in “*Ca. Nitrosothermus*”), which requires further study.

In terms of the “*Ca. Nitrosocaldales/Nitrosocaldaceae*”-accessory set, the largest proportion of them (34.9%) was related to defense mechanisms (**Figure 5**). Among them, genes of the CRISPR-Cas system, participating in resistance to viruses (Barrangou et al., 2007), were detected. Some of these genes have been associated with hyperthermophiles like Cas10, Cmr4g7, Cmr6g7, and Cmr1g7 (Makarova et al., 2003), indicating that these components of the CRISPR-Cas system might be necessary for managing viral infections in high-temperature ecosystems, where many archaeal viruses exist (Prangishvili, 2013; Gudbergsdóttir et al., 2016). Furthermore, a gene coding for indolepyruvate ferredoxin oxidoreductase was also included in the “*Ca. Nitrosocaldales/Nitrosocaldaceae*” accessory set, indicating the ability to ferment aromatic amino acids might be unique to “*Ca. Nitrosocaldales/Nitrosocaldaceae*.” Notably, the gene coding for cobalamin-dependent methionine

synthase (MetH), a key enzyme of one-carbon metabolism, clustered with accessory set of other AOA whereas the cobalamin-independent methionine synthase (MetE) was present in “*Ca. Nitrosocaldales/Nitrosocaldaceae*.” It has been reported that MetH was more efficient than MetE (Helliwell et al., 2016), stressing the differences between “*Ca. Nitrosocaldales/Nitrosocaldaceae*” and other AOA.

In summary, the comparisons of function profiles between “*Ca. Nitrosocaldales/Nitrosocaldaceae*” and other AOA could identify potential mechanisms of thermophilic adaptation in “*Ca. Nitrosocaldales/Nitrosocaldaceae*.”

The Role of HGT in “*Ca. Nitrosocaldaceae*”

Previous studies have stressed the role of HGT in the adaptation of Thaumarchaeota (López-García et al., 2015; Herbold et al., 2017; Ren et al., 2019). However, little is known how HGT events have affected “*Ca. Nitrosocaldales/Nitrosocaldaceae*.” To probe the effect of HGT in adaptations of “*Ca. Nitrosocaldales/Nitrosocaldaceae*,” initial analysis was conducted assisted by an automated pipeline HGTector2 (Zhu et al., 2014) to identify candidate HGT genes and some were subsequently confirmed by phylogenetic analyses. In our study, more than 8% of “*Ca. Nitrosocaldales/Nitrosocaldaceae*” genes might have been horizontally transferred, and as typical, few candidate HGTs were involved in information processing (**Supplementary Figure S7A**). Among all candidate HGTs, genes related to carbohydrate metabolism (~10.55% of the total candidate HGTs), amino acid metabolism (~10.47%), energy metabolism (~8.53%), membrane transport (~6.51%), and metabolism of cofactors and vitamins (6.33%), were the top five most abundant functional classifications (**Supplementary Figure S7B**). Among the potential donors, more genes were acquired from archaea (505), with Crenarchaeota transferring the largest number (146). Nevertheless, the contribution of bacteria (422) was significant, including many genes (47) that were acquired from Firmicutes (**Supplementary Figure S8**).

In the present study, it was noticed that all copies of the SOD genes, encoding for superoxide dismutase required for detoxification of reactive oxygen species (ROS), were potentially horizontally transferred. ROS would be produced during ammonia oxidation (Kim et al., 2016) and aerobic respiration. Most of these SOD might have been transferred from Crenarchaeota or Euryarchaeota (**Supplementary Figure S9**). Other essential HGT genes necessary for protection from ROS include *speE*, and *trxB*. Among them, *trxB*, encoding for thioredoxin reductase, could help repair oxidatively damaged cytoplasmic proteins and protect cells from oxidative stress (Cheng et al., 2017). Spermidine synthase, encoded by gene *speE*, could provide DNA protection and stability in thermal habitats (Cheng et al., 2009). These genes would enable “*Ca. Nitrosocaldales/Nitrosocaldaceae*” to broaden their niche into low-oxygen or anoxic conditions. Gene *pspA* was included among the HGT candidate, which encodes phage shock protein A involved in response

to several stresses (heat, ethanol, and osmotic shock) (Brissette et al., 1990; Kleerebezem et al., 1996). Phylogenetic analysis showed that they might have been transferred from Euryarchaeota (**Supplementary Figure S10**). Genes encoding for ATP-dependent helicase Lhr and Lhr-like helicase might have transferred from other archaea. These genes have been reported to play an important role in DNA repair (Rand et al., 2003). Other HGTs from archaea that might be important in coping with stress included genes coding for multiple antibiotic resistance protein imported from Bathyarchaeota and genes coding for Type I restriction enzyme proteins that protect cells from viral infection (Ahlgren et al., 2017) imported from Euryarchaeota. All the above results revealed that archaeal HGT genes contributed largely to stress responses in “*Ca. Nitrosocaldaceae*.” Gene *arsC* encoding for arsenate reductase, involved in detoxification of arsenate (Rosen and Liu, 2009), was imported from bacteria (**Supplementary Figure S11**).

Some genes involved in the CRISPR-Cas system might belong to HGT genes, which assist “*Ca. Nitrosocaldaceae*” to defend against virus infection. In addition, “*Ca. Nitrosocaldaceae*” acquired several genes coding for proteins involved in phosphate transport from bacteria, including permease protein, ATP-binding protein, and substrate-binding protein, which likely enable “*Ca. Nitrosocaldaceae*” to take up phosphate in low and/or fluctuating phosphate concentrations. Detailed information on HGT genes is shown in **Supplementary Table S8**.

Frequent HGTs have previously been identified in thermophiles, facilitating their adaptation in high-temperature habitats (Aravind et al., 1998; Rhodes et al., 2011). Our study showed that “*Ca. Nitrosocaldaceae*” might have acquired genes from both domains, which may be an essential driver to allow this family to adapt to phosphate-limited and thermal habitats but the confirmation necessitates further studies.

CONCLUSION

In the present study, a total of nine MAGs belonging to “*Ca. Nitrosocaldaceae*” from India and China were recovered from hot spring sediments, which enabled us to obtain a better picture of phylogenetic diversity and metabolic potential of this family. In all, we showed that the “*Ca. Nitrosocaldaceae*” belong to four clades and propose two new species of “*Ca. Nitrosocaldus*” and one new genus, “*Ca. Nitrosothermus*,” to accommodate the new genomes. Similar to other AOA, the potential for ammonia oxidation and carbon fixation via thaumarchaeal 3HP/4HB pathway was conserved in all “*Ca. Nitrosocaldaceae*” MAGs but urea utilization and biosynthesis of vitamins, biotin, cobalamin, and riboflavin were detected in some MAGs. The potential adaptive features of “*Ca. Nitrosocaldaceae*” were explored in this study. AOA shared many conserved genes,

and their central metabolism was highly conserved (ammonia oxidation and carbon fixation). However, specific features exist in “*Ca. Nitrosocaldaceae*” (e.g., cDPG synthesis), stressing their adaptation mechanisms. It was found that >8% (from bacteria and archaea) of the genes in “*Ca. Nitrosocaldaceae*” might be horizontally transferred, and the majority of them were acquired from the same domain, suggesting that HGT played an important role in adapting to thermal habitats. In summary, our study gives an insight into the metabolic potentials and possible adaptations of “*Ca. Nitrosocaldaceae*” in thermal habitats, which shed light on the elucidation of the adaptation mechanisms enabling the ecological success of AOA in different environments.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

Z-HL, MPNR, and BH wrote the manuscript. HC, Z-SH, QL, Z-YD, B-BL, and S-XG performed the metagenomic analysis, genome binning, functional annotation, and evolutionary analysis. W-SS and W-JL guided the project. All authors discussed the results and commented on the manuscript.

FUNDING

This work was financially supported by the National Natural Science Foundation of China (No. 91951205) and the Natural Science Foundation of Guangdong Province, China (No. 2016A030312003). BH was funded by the U.S. National Science Foundation through DEB 1557042. The International Workshop on Geo-omics of Archaea was funded in part by the U.S. National Science Foundation (No. DEB 1928924).

ACKNOWLEDGMENTS

We thank Aharon Oren for nomenclature help and Alan J. M. Baker for revising the manuscript. **Supplementary Figure S2** was created with BioRender.com.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2020.608832/full#supplementary-material>

REFERENCES

- Abby, S. S., Melcher, M., Kerou, M., Krupovic, M., Stieglmeier, M., Rossel, C., et al. (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
- Abdelal, A. T., Bibb, W. F., and Nainan, O. (1982). Carbamate kinase from *Pseudomonas aeruginosa*: purification, characterization, physiological role, and regulation. *J. Bacteriol.* 151, 1411–1419. doi: 10.1128/jb.151.3.1411-1419.1982
- Ahlgren, N. A., Chen, Y., Needham, D. M., Parada, A. E., Sachdeva, R., Trinh, V., et al. (2017). Genome and epigenome of a novel marine Thaumarchaeota strain suggest viral infection, phosphorothioation DNA modification and multiple restriction systems. *Environ. Microbiol.* 19, 2434–2452. doi: 10.1111/1462-2920.13768
- Alves, R. J. E., Minh, B. Q., Urich, T., von Haeseler, A., and Schleper, C. (2018). Unifying the global phylogeny and environmental distribution of ammonia-oxidising archaea based on amoA genes. *Nat. Commun.* 9, 1–17. doi: 10.1038/s41467-018-03861-1
- Amend, J. P., and Shock, E. L. (2001). Energetics of overall metabolic reactions of thermophilic and hyperthermophilic archaea and bacteria. *FEMS Microbiol. Rev.* 25, 175–243. doi: 10.1016/S0168-6445(00)00062-0
- Aravind, L., Tatusov, R. L., Wolf, Y. I., Walker, D. R., and Koonin, E. V. (1998). Evidence for massive gene exchange between archaeal and bacterial hyperthermophiles. *Trends. Genet.* 14, 442–444. doi: 10.1016/S0168-9525(98)01553-4
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021
- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., et al. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315, 1709–1712. doi: 10.1126/science.1138140
- Barry, E. R., and Bell, S. D. (2006). DNA replication in the archaea. *Microbiol. Mol. Biol. Rev.* 70, 876–887. doi: 10.1128/MMBR.00029-06
- 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
- Berg, I. A., Kockelkorn, D., Buckel, W., and Fuchs, G. (2007). A 3-Hydroxypropionate/ 4-Hydroxybutyrate autotrophic carbon dioxide assimilation pathway in Archaea. *Science* 318, 1782–1786. doi: 10.1126/science.1149976
- Berube, P. M., Samudrala, R., and Stahl, D. A. (2007). Transcription of all amoC copies is associated with recovery of *Nitrosomonas europaea* from ammonia starvation. *J. Bacteriol.* 189, 3935–3944. doi: 10.1128/JB.01861-06
- Berube, P. M., and Stahl, D. A. (2012). The divergent AmoC3 subunit of ammonia monooxygenase functions as part of a stress response system in *Nitrosomonas europaea*. *J. Bacteriol.* 194, 3448–3456. doi: 10.1128/JB.00133-12
- Blainey, P. C., Mosier, A. C., Potanina, A., Francis, C. A., and Quake, S. R. (2011). Genome of a low-salinity ammonia-oxidizing archaeon determined by single-cell and metagenomic analysis. *PLoS One* 6:e16626. doi: 10.1371/journal.pone.0016626
- Brissette, J. L., Russel, M., Weiner, L., and Model, P. (1990). Phage shock protein, a stress protein of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 87, 862–866. doi: 10.1073/pnas.87.3.862
- Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59–60. doi: 10.1038/nmeth.3176
- Capella-Gutiérrez, S., Silla-Martínez, J. M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972–1973. doi: 10.1093/bioinformatics/btp348
- Cheng, C., Dong, Z., Han, X., Wang, H., Jiang, L., Sun, J., et al. (2017). Thioredoxin A Is Essential for Motility and Contributes to Host Infection of *Listeria monocytogenes* via Redox Interactions. *Front. Cell Infect. Microbiol.* 7:287. doi: 10.3389/fcimb.2017.00287
- Cheng, L., Zou, Y., Ding, S., Zhang, J., Yu, X., Cao, J., et al. (2009). Polyamine accumulation in transgenic tomato enhances the tolerance to high temperature stress. *J. Integr. Plant Biol.* 51, 489–499. doi: 10.1111/j.1744-7909.2009.00816.x
- Chow, J., Danso, D., Ferrer, M., and Streit, W. R. (2018). The Thaumarchaeon *N. gargensis* carries functional bioABD genes and has a promiscuous *E. coli* Δ bioH-complementing esterase EstN1. *Sci. Rep.* 8, 1–14. doi: 10.1038/s41598-018-32059-0
- Daebeler, A., Herbold, C. W., Vierheilig, J., Sedlacek, C. J., Pjevac, P., Albertsen, M., et al. (2018). Cultivation and genomic analysis of "Candidatus Nitrosocaldus islandicus," an obligately thermophilic, ammonia-oxidizing Thaumarchaeon from a hot spring biofilm in Graendalur Valley, Iceland. *Front. Microbiol.* 9:193. doi: 10.3389/fmicb.2018.00193
- de la Torre, J. R., Walker, C. B., Ingalls, A. E., Könneke, 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
- Declyre, H., Heylen, K., Bjorn, T., and Willems, A. (2016). Highly diverse nirK genes comprise two major clades that harbour ammonium-producing denitrifiers. *BMC Genomics* 17:155. doi: 10.1186/s12864-016-2465-0
- 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
- Dodsworth, J. A., McDonald, A. I., and Hedlund, B. P. (2012). Calculation of total free energy yield as an alternative approach for predicting the importance of potential chemolithotrophic reactions in geothermal springs. *FEMS Microbiol. Ecol.* 81, 446–454. doi: 10.1111/j.1574-6941.2012.01369.x
- Doxey, A. C., Kurtz, D. A., Lynch, M. D., Sauder, L. A., and Neufeld, J. D. (2015). Aquatic metagenomes implicate Thaumarchaeota in global cobalamin production. *ISME J.* 9, 461–471. doi: 10.1038/ismej.2014.142
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucl. Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340
- Emms, D. M., and Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* 16:157. doi: 10.1186/s13059-015-0721-2
- Emms, D. M., and Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* 20:238. doi: 10.1186/s13059-019-1832-y
- Gudbergssdóttir, S. R., Menzel, P., Krogh, A., Young, M., and Peng, X. (2016). Novel viral genomes identified from six metagenomes reveal wide distribution of archaeal viruses and high viral diversity in terrestrial hot springs. *Environ. Microbiol.* 18, 863–874. doi: 10.1111/1462-2920.13079
- Hamilton, T. L., Koonce, E., Howells, A., Havig, J. R., Jewell, T., de la Torre, J. R., et al. (2014). Competition for ammonia influences the structure of chemotrophic communities in geothermal springs. *Appl. Environ. Microbiol.* 80, 653–661. doi: 10.1128/AEM.02577-13
- Heal, K. R., Qin, W., Amin, S. A., Devol, A. H., Moffett, J. W., Armbrust, E. V., et al. (2018). Accumulation of NO₂-cobalamin in nutrient-stressed ammonia oxidizing archaea and in the oxygen deficient zone of the eastern tropical North Pacific. *Environ. Microbiol. Rep.* 10, 453–457. doi: 10.1111/1758-2229.12664
- Hedlund, B. P., Cole, J. K., Williams, A. J., Hou, W., Zhou, E., Li, W., et al. (2012). A review of the microbiology of the Rehai geothermal field in Tengchong, Yunnan Province, China. *Geoscience Front.* 3, 273–288. doi: 10.1016/j.gsf.2011.12.006
- Helliwell, K. E., Lawrence, A. D., Holzer, A., Kudahl, U. J., Sasso, S., Kräutler, B., et al. (2016). Cyanobacteria and eukaryotic algae use different chemical variants of Vitamin B12. *Curr. Biol.* 26, 999–1008. doi: 10.1016/j.cub.2016.02.041
- Hensel, R., and König, H. (1988). Thermoadaptation of methanogenic bacteria by intracellular ion concentration. *FEMS Microbiol. Lett.* 49, 75–79. doi: 10.1111/j.1574-6968.1988.tb02685.x
- Herbold, C. W., Lehtovirta-Morley, L. E., Jung, M. Y., Jehlich, N., Hausmann, B., Han, P., et al. (2017). Ammonia-oxidising archaea living at low pH: insights from comparative genomics. *Environ. Microbiol.* 19, 4939–4952. doi: 10.1111/1462-2920.13971
- Holloway, J. M., Nordstrom, D. K., Böhlke, J. K., McCleskey, R. B., and Ball, J. W. (2011). Ammonium in thermal waters of Yellowstone national park: processes affecting speciation and isotope fractionation. *Geochim. Cosmochim. Acta* 75, 4611–4636. doi: 10.1016/j.gca.2011.05.036
- Hou, W. G., Wang, S., Dong, H. L., Jiang, H. C., Briggs, B. R., Peacock, J. P., et al. (2013). A comprehensive census of microbial diversity in hot springs of Tengchong, Yunnan Province China using 16S rRNA gene pyrosequencing. *PLoS One* 8:e53350. doi: 10.1371/journal.pone.0053350

- Hua, Z. S., Han, Y. J., Chen, L. X., Liu, J., Hu, M., Li, S. J., et al. (2015). Ecological roles of dominant and rare prokaryotes in acid mine drainage revealed by metagenomics and metatranscriptomics. *ISME J.* 9, 1280–1294. doi: 10.1038/ismej.2014.212
- Hyatt, D., Chen, G. L., Locascio, P. F., Land, M. L., Larimer, F. W., and Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinform.* 11:119. doi: 10.1186/1471-2105-11-119
- Ishii, M., Miyake, T., Satoh, T., Sugiyama, H., Oshima, Y., Kodama, T., et al. (1996). Autotrophic carbon dioxide fixation in *Acidianus brierleyi*. *Arch. Microbiol.* 166, 368–371. doi: 10.1007/BF01682981
- Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., and Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* 30:5114. doi: 10.1038/s41467-018-07641-9
- Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2016). KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucl. Acids Res.* 45, D353–D361. doi: 10.1093/nar/gkw1092
- Kang, D. D., Froula, J., Egan, R., and Wang, Z. (2015). MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* 3:e1165. doi: 10.7717/peerj.1165
- Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010
- Kengen, S. W. M., and Stams, A. J. M. (1994). Formation of L-alanine as a reduced end product in carbohydrate fermentation by the hyperthermophilic archaeon *Pyrococcus furiosus*. *Arch. Microbiol.* 161, 168–175. doi: 10.1007/BF00276479
- Kerou, M., Offre, P., Valledor, L., Abby, S. S., Melcher, M., Nagler, M., et al. (2016). 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, 7937–7946. doi: 10.1073/pnas.1601212113
- Kim, H. J., and Winge, D. R. (2013). Emerging concepts in the flavinylation of succinate dehydrogenase. *Biochim. Biophys. Acta Bioenerg.* 1827, 627–636. doi: 10.1016/j.bbabi.2013.01.012
- Kim, J. G., Park, S. J., Sinnighe Damsté, J. S., Schouten, S., Rijpstra, W. I. C., Jung, W. 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
- Kim, S., Ihara, K., Katsube, S., Hori, H., Ando, T., Isogai, E., et al. (2015). Characterization of the L-alanine exporter AlaE of *Escherichia coli* and its potential role in protecting cells from a toxic-level accumulation of L-alanine and its derivatives. *Microbiologyopen* 4, 632–643. doi: 10.1002/mbo3.269
- Kleerebezem, M., Crielaard, W., and Tommassen, J. (1996). Involvement of stress protein PspA (phage shock protein A) of *Escherichia coli* in maintenance of the protonmotive force under stress conditions. *EMBO J.* 15, 162–171. doi: 10.1002/j.1460-2075.1996.tb00344.x
- Könneke, M., Schubert, D. M., Brown, P. C., Hügl, 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
- Konstantinidis, K. T., Rosselló-Móra, R., and Amann, R. (2017). Uncultivated microbes in need of their own taxonomy. *ISME J.* 11, 2399–2406. doi: 10.1038/ismej.2017.113
- Kozlowski, J. A., Stieglmeier, M., Schleper, C., Klotz, M. G., and Stein, L. Y. (2016). Pathways and key intermediates required for obligate aerobic ammoniadependent chemolithotrophy in bacteria and Thaumarchaeota. *ISME J.* 10, 1836–1845. doi: 10.1038/ismej.2016.2
- Lagesen, K., Hallin, P., Rødland, E. A., Staerfeldt, H. H., Rognes, T., and Ussery, D. W. (2007). RNAMmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35, 3100–3108. doi: 10.1093/nar/gkm160
- Lao-Sirieix, S., and Bell, S. D. (2004). The heterodimeric primase of the hyperthermophilic archaeon *Sulfolobus solfataricus* possesses DNA and RNA primase, polymerase and 3'-terminal nucleotidyl transferase activities. *J. Mol. Biol.* 344, 1251–1263. doi: 10.1016/j.jmb.2004.10.018
- Lehtovirta-Morley, L. E., Ge, C., Ross, J., Yao, H., Nicol, G. W., and Prosser, J. I. (2014). Characterisation of terrestrial acidophilic archaeal ammonia oxidizers and their inhibition and stimulation by organic compounds. *FEMS Microbiol. Ecol.* 89, 542–552. doi: 10.1111/1574-6941.12353
- Lehtovirta-Morley, L. E., Sayavedra-Soto, L. A., Gallois, N., Schouten, S., Stein, L. Y., Prosser, J. I., et al. (2016). Identifying potential mechanisms enabling acidophily in the ammonia-oxidizing archaeon “*Candidatus Nitrosotalea devanatterra*”. *Appl. Environ. Microbiol.* 82, 2608–2619. doi: 10.1128/AEM.04031-15
- Lenza, R. F., and Vasconcelos, W. L. (2001). Structural evolution of silica sols modified with formamide. *Mater. Res.* 4, 175–179. doi: 10.1590/S1516-14392001000300006
- Letunic, I., and Bork, P. (2019). Interactive Tree of Life (iTOL) v4: recent updates and new developments. *Nucl. Acids Res.* 47, W256–W259. doi: 10.1093/nar/gkz239
- López-García, P., Zivanovic, Y., Deschamps, P., and Moreira, D. (2015). Bacterial gene import and mesophilic adaptation in archaea. *Nat. Rev. Microbiol.* 13, 447–456. doi: 10.1038/nrmicro3485
- Lowe, T. M., and Eddy, S. R. (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucl. Acids Res.* 25, 955–964. doi: 10.1093/nar/25.5.955
- Makarova, K. S., Wolf, Y. I., and Koonin, E. V. (2003). Potential genomic determinants of hyperthermophily. *Trends. Genet.* 19, 172–176. doi: 10.1016/S0168-9525(03)00047-7
- 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
- Mansoorabadi, S. O., Thibodeaux, C. J., and Liu, H. W. (2007). The diverse roles of flavin coenzymes nature’s most versatile thespians. *J. Org. Chem.* 72, 6329–6342. doi: 10.1002/chin.200748273
- Marteinsson, V. T., Hauksdóttir, S., Hobel, C. F. V., Kristmannsdóttir, H., Hreggvidsson, G. O., and Kristjánsson, J. K. (2001). Phylogenetic diversity analysis of subterranean hot springs in Iceland. *Appl. Environ. Microbiol.* 67, 4242–4248. doi: 10.1128/AEM.67.9.4242-4248.2001
- Miller, S. L., and Urey, H. C. (1959). Organic compound synthesis on the primitive earth. *Science* 130, 245–251. doi: 10.1126/science.130.3370.245
- Morgan, J. E., Blankenship, J. W., and Matthews, H. R. (1987). Polyamines and acetyl polyamines increase the stability and alter the conformation of nucleosome core particles. *Biochemistry* 26, 3643–3649. doi: 10.1021/bi00386a058
- Nguyen, L. T., Schmidt, H. A., von Haeseler, A., and Minh, B. Q. (2014). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300
- 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
- Pace, H. C., and Brenner, C. (2001). The nitrilase superfamily: classification, structure and function. *Genome Biol.* 2, 1–9. doi: 10.1186/gb-2001-2-1-reviews0001
- Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25, 1043–1055. doi: 10.1101/gr.186072.114
- Parks, D. H., Rinke, C., Chuvochina, M., Chaumeil, P. A., Woodcroft, B. J., Evans, P. N., et al. (2017). Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nat. Microbiol.* 2, 1533–1542. doi: 10.1038/s41564-017-0012-7
- Prangishvili, D. (2013). The wonderful world of archaeal viruses. *Annu. Rev. Microbiol.* 67, 565–585. doi: 10.1146/annurev-micro-092412-155633
- Pritchard, L., Glover, R. H., Humphris, S., Elphinstone, J. G., and Toth, I. K. (2016). Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Anal. Methods.* 8, 12–24. doi: 10.1039/C5AY02550H
- Rand, L., Hinds, J., Springer, B., Sander, P., Buxton, R. S., and Davis, E. O. (2003). The majority of inducible DNA repair genes in *Mycobacterium tuberculosis* are induced independently of RecA. *Mol. Microbiol.* 50, 1031–1042. doi: 10.1046/j.1365-2958.2003.03765.x
- 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

- Ren, M., Feng, X., Huang, Y., Wang, H., Hu, Z., Clingepel, S., et al. (2019). Phylogenomics suggests oxygen availability as a driving force in Thaumarchaeota evolution. *ISME J.* 13, 2150–2161. doi: 10.1038/s41396-019-0418-8
- Reyes, C., Hodgskiss, L. H., Kerou, M., Pribasnik, T., Abby, S. S., Bayer, B., et al. (2020). Genome wide transcriptomic analysis of the soil ammonia oxidizing archaeon *Nitrososphaera viennensis* upon exposure to copper limitation. *ISME J.* 14, 2659–2674. doi: 10.1038/s41396-020-0715-2
- Rhodes, M. E., Spear, J. R., Oren, A., and House, C. H. (2011). Differences in lateral gene transfer in hypersaline versus thermal environments. *BMC Evol. Biol.* 11:199. doi: 10.1186/1471-2148-11-199
- Roberts, M. F. (2005). Organic compatible solutes of halotolerant and halophilic microorganisms. *Saline Syst.* 1:5. doi: 10.1186/1746-1448-1-5
- Rosen, B. P., and Liu, Z. (2009). Transport pathways for arsenic and selenium: a minireview. *Environ. Int.* 35, 512–515. doi: 10.1016/j.envint.2008.07.023
- Santoro, A. E., Dupont, C. L., Richter, R. A., Craig, M. T., Carini, P., McIlvin, M. R., et al. (2015). Genomic and proteomic characterization of "Candidatus Nitrosopelagicus brevis": an ammonia-oxidizing archaeon from the open ocean. *Proc. Natl. Acad. Sci. U.S.A.* 112, 1173–1178. doi: 10.1073/pnas.1416223112
- Schulte, M., and Shock, E. (1995). Thermodynamics of strecker synthesis in hydrothermal systems. *Origins. Life. Evol. B.* 25, 161–173. doi: 10.1007/BF01581580
- Segata, N., Börnigen, D., Morgan, X. C., and Huttenhower, C. (2013). PhyloPhlAn is a new method for improved phylogenetic and taxonomic placement of microbes. *Nat. Commun.* 4:2304. doi: 10.1038/ncomms3304
- Shen, C., Li, X., Rasooly, A., Guo, L., Zhang, K., and Yang, M. (2016). A single electrochemical biosensor for detecting the activity and inhibition of both protein kinase and alkaline phosphatase based on phosphate ions induced deposition of redox precipitates. *Biosens. Bioelectron.* 85, 220–225. doi: 10.1016/j.bios.2016.05.025
- Shima, S., Hérault, 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
- 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., Zumbärgel, 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
- Swithers, K. S., Petrus, A. K., Secinaro, M. A., Nesbø, C. L., Gogarten, J. P., Noll, K. M., et al. (2012). Vitamin B12 synthesis and salvage pathways were acquired by horizontal gene transfer to the Thermotogales. *Genome Biol. Evol.* 4, 842–851. doi: 10.1093/gbe/evs057
- Tabor, C. W., and Tabor, H. (1984). Polyamines. *Annu. Rev. Biochem.* 53, 749–790. doi: 10.1146/annurev.bi.53.070184.003533
- Tourna, M., Stieglmeier, M., Spang, A., Könneke, 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
- Walker, C. B., de La Torre, J. R., Klotz, M. G., Urakawa, H., Pinel, N., Arp, D. J., et al. (2010). *Nitrosopumilus maritimus* genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. *Proc. Natl. Acad. Sci. U.S.A.* 107, 8818–8823. doi: 10.2307/25681500
- Ward, L. M., Idei, A., Nakagawa, M., Ueno, Y., Fischer, W. W., and McGlynn, S. E. (2019). Geochemical and metagenomic characterization of Jinata Onsen, a Proterozoic-analog hot spring, reveals novel microbial diversity including iron-tolerant phototrophs and thermophilic lithotrophs. *Microbes. Environ.* 34, 278–292. doi: 10.1264/jjsme2.ME19017
- Winkler, F. K. (2006). Amt/MEP/Rh proteins conduct ammonia. *Pflugers Archiv.* 451, 701–707. doi: 10.1007/s00424-005-1511-6
- 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
- Zhang, Y., Rodionov, D. A., Gelfand, M. S., and Gladyshev, V. N. (2009). Comparative genomic analyses of nickel, cobalt and vitamin B12 utilization. *BMC Genom.* 10:78. doi: 10.1186/1471-2164-10-78
- Zhu, Q., Kosoy, M., and Dittmar, K. (2014). Hgtector: an automated method facilitating genome-wide discovery of putative horizontal gene transfers. *BMC Genom.* 15:717. doi: 10.1186/1471-2164-15-717

Conflict of Interest: 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 © 2021 Luo, Narsing Rao, Chen, Hua, Li, Hedlund, Dong, Liu, Guo, Shu and Li. 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(s) 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.