bioRxiv preprint doi: https://doi.org/10.1101/2025.06.25.661132; this version posted June 25, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

Title

Cultivation of Methanonezhaarchaeia, the third class of methanogens within the phylum Thermoproteota

Short title

Cultivation of Methanonezhaarchaeia

Authors

Anthony J. Kohtz[1,†,*], Sylvia Nupp[1], Roland Hatzenpichler[1,2,*]

Affiliations

[1], Department of Chemistry and Biochemistry, Center for Biofilm Engineering, and Thermal Biology Institute, Montana State University, Bozeman, MT-59717, USA

[2], Department of Microbiology and Cell Biology, Montana State University, Bozeman, MT-59717, USA

†, Present address: Department of Microbiology, Biomedicine Discovery Institute, Monash University, Clayton, VIC 3800, Australia

*, Correspondence: anthony.kohtz@monash.edu, rolandhatzenpichler@gmail.com

Abstract

Methane is a potent greenhouse gas that is largely produced through the activity of methanogenic archaea, which contribute to Earth's dynamic climate and biogeochemical cycles. In the past decade, metagenomics revealed that lineages outside of the traditional Euryarchaeota superphylum encode genes for methanogenesis. This was recently confirmed through the cultivation of two classes of methanogenic Thermoproteota. Thus far, all methanogens within the Thermoproteota are predicted or were shown to be methylotrophic. The only exception to this is the Nezhaarchaea, which, based on metagenomics, are predicted to be CO_2-reducing methanogens. Here, we demonstrate methanogenic activity in a third class of Thermoproteota, the Methanonezhaarchaeia. We expand the metabolic diversity of this class by cultivating a methylotrophic species, *Candidatus* Methanonezhaarchaeum fastidiosum YNP3N. We describe novel genes involved in methanogenesis that are not found in other methanogens. We investigate the metabolic diversity of Methanonezhaarchaeia, including metabolic modifications accompanying frequent loss of methanogenesis in this class. This highlights gaps in our understanding of the biochemistry, diversity, and evolution of non-traditional methanogens and their contributions to carbon cycling.

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.25.661132; this version posted June 25, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

Introduction

Methanogenesis, the production of methane coupled to energy conservation, is a uniquely archaeal metabolism (*1, 2*). Through their activity, methanogens are responsible for mediating global methane emissions, which significantly affects Earth's climate (*3*). While all methanogens share a unique methane-forming enzyme, methyl coenzyme-M reductase (MCR), different precursor molecules can be converted to methane via distinct pathways : (i) CO_2-reducing, (ii) acetoclastic, (iii) methyl-dismutating, (iv) methyl-reducing, (v) methoxyl-dismutating, and (vi) alkanotrophic methanogenesis(*1*).

In addition to discoveries regarding their metabolism (*4-7*), our understanding of which archaea engage in methanogenesis has radically changed over the past decade. Traditionally, all methanogens were thought to belong to the superphylum Euryarchaeota. However, this view rapidly changed after several independent metagenomic studies discovered genes for anaerobic alkane-cycling in numerous uncultured archaeal lineages within the phylum Thermoproteota (*a.k.a.* TACK superphylum), including Korarchaeia (*5, 7*), Methanomethylicia (previously called Verstraetearchaeota) (*8*), Nezhaarchaea (*9*), Nitrososphaeria (*10, 11*), Aukarchaeles (*12*), and Bathyarchaeia (*13, 14*). In addition, even in traditional Euryarchaeota, multiple deep branching methanogenic and alkane degrading lineages have been discovered in the past decade, showing our understanding of the mediators of these metabolisms are not yet complete (*5, 6, 9, 15-18*). Only recently, the first three cultures of Korarchaeia and Methanomethylicia were recovered, enabling an experimental validation of their methanogenic physiology (*12, 19, 20*). The vast expansion in experimentally confirmed as well as potential non-traditional methanogens has wide-ranging implications for our understanding of global methane cycling, the evolution and biology of archaea, and the biotechnological potential of these novel methanogens (*1, 11, 21, 22*).

Almost all lineages of (potential) methanogens within the Thermoproteota are predicted (*5, 7-14*) or have been demonstrated (*12, 19, 20*) to be methyl-reducing methanogens. This idea is supported by the presence of methyltransferase genes and the lack of both the Wood-Ljungdahl Pathway (WLP) and the tetrahydromethanopterin S-methyltransferase (MTR) that can connect methanogenesis to the WLP. So far, the only lineage within the Thermoproteota that encodes both MCR and MTR are the 'Nezhaarchaea', which, based on metagenomics, are predicted to be CO_2-reducing methanogens (*9, 11*). Nezhaarchaea were first discovered in hot spring samples (*9, 11*) but are also present in deep-sea hydrothermal systems (*23*), suggesting that they are important for methane cycling in high temperature environments. So far, the only experimental evidence supportive of a methanogenic lifestyle was a study demonstrating that Nezhaarchaea are the most abundant MCR-encoding archaea in some hot springs and transcribe MCR in methanogenic microcosms seeded with biomass from those hot springs (*24*). While this was an important step forward, most of their metabolism remains unresolved due to low recruitment of metatranscriptomic reads to genes beyond MCR and Nezhaarchaea remain enigmatic without cultures.

Here we demonstrate methylotrophic methanogenesis in the Nezhaarchaea through selective cultivation, fluorescence microscopy, and metagenomic and metatranscriptomic sequencing. We further assess the wider metabolic diversity in this lineage, suggesting loss of methanogenesis across this lineage is associated with rewiring of membrane-associated redox complexes.

Results

Cultivation of a new methylotrophic methanogen

We recently demonstrated the Lower Culex Basin (LCB) in Yellowstone National Park to be an area hosting a taxonomically diverse range of novel MCR-encoding archaea affiliating with the superphylum Thermoproteota (*25*). Notably, methanogenic cultures of the first two classes of Thermoproteotal methanogens, affiliating to the Korarchaeia and Methanomethylicia, were obtained from hot springs located in the LCB (*19, 20*). Furthermore, mesocosm experiments using sediments from hot spring 'LCB003'

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.25.661132; this version posted June 25, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

revealed that members of a third class of Thermoproteotal MCR-encoding archaea, Nezhaarchaea, became slightly enriched (up to 8% relative abundance) under methanogenic conditions when incubated with methanol under a N_2/CO_2 headspace (*25*).

To further explore the metabolism of Nezhaarchaea and attempt their cultivation, we incubated sediments from hot spring LCB003 in anoxic media under a variety of conditions designed to stimulate methylotrophic methanogenesis at the *in situ* temperature of 77 °C and pH of 6.5. We found that incubations with monomethylamine (MMA), dimethylamine, or trimethylamine, with or without isopropanol), and a N_2/CO_2 headspace repeatedly led to the enrichment of Nezhaarchaea along with pronounced methane production. Further incubations revealed that isopropanol is not required for enrichment and has little effect on total methane production (Fig. 1A). MMA exhibited the most consistent response in methane production, and all further work was performed using MMA as a substrate. The cultures exhibited long lag times of around 50-60 days when first inoculated into anoxic media, and this response was repeatable over multiple years of sampling from the same hot spring. Cultures also frequently stopped growing after 1-2 transfers into fresh media; however, we were able to eventually obtain a stable methanogenic culture for detailed study.

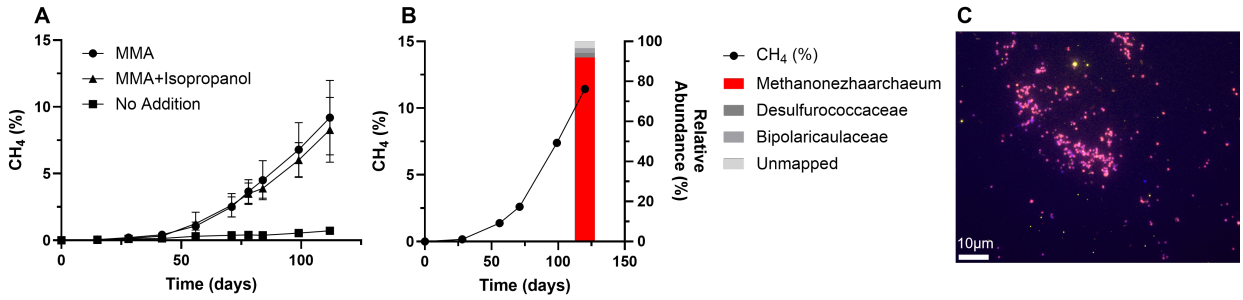


Figure 1. Enrichment cultivation and visualization of *Ca.* Methanonezhaarchaeum fastidiosum YNP3N. (A) Methane production by the culture under varying amendments: MMA only (n = 9), MMA and isopropanol (n = 9), and no addition (n = 3). Symbols are the mean from biological replicates and error bars are one standard deviation, where not visible the error bars are smaller than the symbol. **(B)** Long term cultivation on MMA and multiple transfers produced a sediment-free culture >90% enriched in strain YNP3N according to metagenomics (n = 1). **(C)** DOPE-FISH confirmed the high enrichment of YNP3. Abundant coccoid cells double hybridized (appearing in pink) with a general archaeal probe (Arch915, red) and a probe specific to Methanonezhaarchaea (Mnez277, yellow). Cells were counterstained with DAPI (blue). Bright yellow signals stem from auto-fluorescent sediment particles.

We found that incubations initiated with 50% H_2 in the headspace did not develop any significant enrichment of Nezhaarchaea, suggesting that they may be adapted to low H_2 concentrations or are outcompeted under those conditions. In contrast, the methanogenic korarchaeon *Ca.* Methanodesulfokora washburnensis, which we previously enriched from the same hot spring, rapidly increases in abundance under a 50% H_2 headspace and methanol amendment (*19*). This indicates that these thermoproteotal methanogens inhabit different niches within the same hot spring sediment with regards to both substrate (methylamines vs. methanol) and hydrogen (no/low H_2 vs. high H_2) preferences.

Long read Nanopore and short read Illumina sequencing of the culture recovered a complete circular genome of the nezhaarchaeon, with a size of 1.75 Mbp and a GC content of 53% (Supplementary Table 1). Read mapping of the short read metagenome to the genome revealed that this archaeon dominated the culture at a relative abundance of 90.13% (Figure 1B, Supplementary Table 1). This strong enrichment was independently confirmed by DOPE-FISH (Figure 1C). In contrast, by mapping the raw reads of the LCB003 hot spring metagenome to the genome of the archaeon in culture, we estimate the relative abundance of the organism in its native habitat at 0.46%. Phylogenomic, average nucleotide identity (ANI), and GTDB-tk analyses showed that the cultured archaeon clusters within the Nezhaarchaea lineage and that it represents

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.25.661132; this version posted June 25, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

a new species, for which we propose the name *Candidatus* Methanonezhaarchaeum fastidiosum strain YNP3N (Figure 2A, Figure S1, Supplementary Table 1). We chose this name to emphasize the 'methano-' prefix typical in methanogen taxonomy, to preserve previously proposed names as much as possible (*i.e.*, Nezhaarchaeota, Nezhaarchaea (*9, 11, 24*)), and to emphasize the fastidious and slow-growing nature of this organism. We note that this name change would necessitate slightly renaming this group using the type genus Methanonezhaarchaeum propagated up to class Methanonezhaarchaeia (see Etymology section at the end of the manuscript the and the Protologue in the Supplementary Online Information).

Strain YNP3N encoded the only *mcrA* gene recovered in the metagenome assembly, and phylogenetic analysis of this gene placed it within the Methanonezhaarchaeia, with recently described Bathyarchaeia sequences as a sister group (Figure 2B) (*13*). We also found fragmented *mcrBCD* genes related to *Ca*. Methanoglobus hypatiae (*16*), known to metabolize MMA, in the metagenome assembly. However, no Methanoglobus *mcrAG* was recovered, the *mcrBCD* genes had low coverage, and mapping the metagenomic reads to the *Ca*. M. hypatiae genome revealed a relative abundance of 0.037%. Together, this suggests minimal impacts on the observed methane production by *Ca*. M. hypatiae. We also observed a second species of Methanonezhaarchaeia in the culture, with a low relative abundance of 1.6%. However, the MAG was non-circular, less complete (85.29%) and did not encode *mcr* genes (Supplementary Table 1).

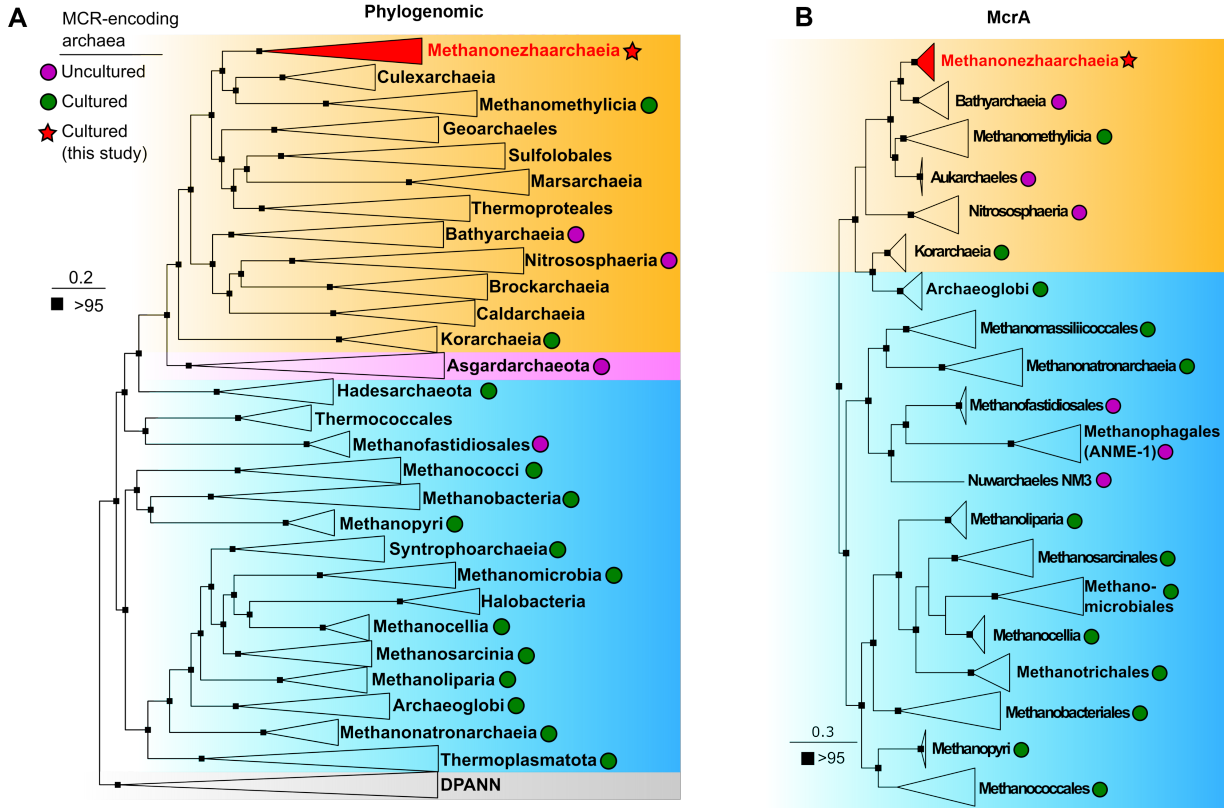


Figure 2. Phylogenetic placement of Methanonezhaarchaeia based on single copy marker genes and McrA. (A) Phylogenomic tree of archaea based on a set of 43 single-copy marker genes. Major groups of archaea are highlighted in orange (Thermoproteota/TACK superphylum), pink (Asgardarchaeota), blue (Euryarchaeota), and gray (DPANN). For inclusion, genomes were required to have at least 37 of the marker genes. Squares at the nodes indicate ultrafast bootstrap support, only values above 95 are shown. **(B)** Phylogenetic tree of McrA sequences demonstrating the affiliation of YNP3N within the Methanonezhaarchaeia (Thermoproteota, orange) compared to traditional Euryarchaeota (blue). Squares at the nodes indicate ultrafast bootstrap support, only values above 95 are shown.

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.25.661132; this version posted June 25, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

A new FISH probe was designed to target Methanonezhaarchaeia. Application of this probe and a general archaeal probe revealed abundant coccoid cells that were labeled with both probes, confirming the high enrichment of strain YNP3N (Fig. 1C). Searches against public databases using the 16S rRNA and *mcrA* gene sequences showed that Methanonezhaarchaeia are globally prevalent in terrestrial hot springs and marine hydrothermal systems (Supplementary Table 2,3). However, it is important to note that Methanonezhaarchaeia likely avoided detection in the past because commonly used PCR primers exhibit multiple mismatches to their *mcrA* genes(*25*).

Metabolism and gene expression

Reconstruction of the metabolic potential revealed that strain YNP3N encodes a more diverse metabolism for methanogenic substrates relative to previously published Methanonezhaarchaeia MAGs, which had been proposed to exclusively use H_2/CO_2 for methanogenesis. As expected, the MCR complex genes were highly expressed during methanogenic growth. During growth on monomethylamine, YNP3N highly expresses methyltransferase genes involved in the metabolism of this substrate (*mtmB*, *mtbC*) and expresses genes for using other methylamines (*mtbB*, *mttB*) at a lower level (Figure 3B). While there was not an associated *mtbA* gene, an *mtaA*-like gene is co-located and highly expressed, likely replacing the activity of *mtbA*, similar to the methyltransferase systems found in Methanomassillicoccales (*26*) and *Ca.* Methanglobus hypatiae (*16*). We also identified pyrrolysine biosynthesis genes, PylSBCD, and a Pyl-tRNA used to recode stop codons, notably found within the methylamine methyltransferase genes (*27-29*). Furthermore, we observed expression of diverse corrinoid transporter genes and cobalamin-binding proteins required for import of vitamin B_{12} and synthesis of cobalamin proteins found in methyltransferase systems (*30, 31*). This indicates the importance of this vitamin in the metabolism of this organism and suggests that in the native hot spring these transporters are likely important in uptake of B_{12} from the environment. Strain YNP3N also expressed both the tetrahydromethanopterin S-methyltransferase (MTR) complex and the methyl branch of the Wood-Ljungdahl Pathway (WLP), further indicating its capability for methyl-dismutating methanogenesis. This is in stark contrast to previously cultured Thermoproteota methanogens, which all lack a full MTR complex and most lack the WLP (*1, 10, 12, 19, 20*). We could not sustain growth of strain YNP3N without addition of MMA as a growth substrate, even with CO_2/bicarbonate present, suggesting that YNP3N, contrary to its genomic potential, does not perform CO_2-reducing methanogenesis *in vitro*. This finding highlights the importance of testing predictions from (meta)genomics through careful experiments and the indispensability of cultivation.

For heterodisulfide reduction, methanogens lacking cytochromes commonly have an electron-bifurcating cytoplasmic Mvh-Hdr complex that allows coupling the oxidation of H_2 to the reduction of ferredoxin and heterodisulfide (*32*). In formate-utilizing methanogens, Mvh-Hdr can also form large complexes with FdhAB, to allow formate to substitute for H_2 as an electron donor (*33*). YNP3N lacks FdhAB or MvhABG associated with a soluble HdrABC operon. Instead, a second gene cluster containing FwdBD subunits are found co-located with HdrABC in the YNP3N genome. The FwdBD subunits typically perform CO_2-reduction to formate in hydrogenotrophic methanogens, with formate then migrating to the FwdA subunit (*34, 35*). Thus, this Fwd-Hdr complex may enable formate oxidation to be coupled with heterodisulfide and ferredoxin reduction. The additional co-located FwdF polyferredoxin and FwdG genes would likely allow to electronically couple the HdrABC and FwdBD subunits. This Fwd-Hdr organization is a new strategy distinct from the characterized Fmd-Hdr-Fdh or Fwd-Hdr-Mvh-Fdh complexes found in Methanosprillium or Methanococcus, respectively (*33, 36*).

Strain YNP3N also expresses a separate FwdABCD gene cluster co-located with other genes of the WLP, and here the FwdBD genes are fused, indicating gene duplication and fusion events enabled a repurposing of the second Fwd complex for involvement in heterodisulfide cycling.

Furthermore, for formate, hydrogen, and heterodisulfide cycling, YNP3N also expressed a membrane associated Group 4 [NiFe]-hydrogenase co-located with HdrABC and formate dehydrogenase (FdhA) subunits, termed Ehg (Fig. 3, Fig. S3) (*37*). Thus, this complex may be able to couple the oxidation of H_2

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.25.661132; this version posted June 25, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

or formate to the reduction of heterodisulfide and ion translocation for energy conservation. The FdhA subunit in YNP3N is similar in sequence (35-38% sequence identity) to those found in formate oxidizing methanogens within the Methanococcales, Methanobacteriales, and Methanocellales (*38-40*). However, with the presence of HdrA, this complex may also be capable of flavin based electron bifurcation and ferredoxin reduction. Internal formate cycling may take place during methyl oxidation through the WLP and may be sensitive to formate concentration. Attempts to grow strain YNP3N with MMA on a range of exogenous formate concentrations (2, 10, and 100 mM) were unsuccessful (not shown) and we were unable to identify a formate-specific transporter as found in other formate-utilizing methanogens (*38-40*).

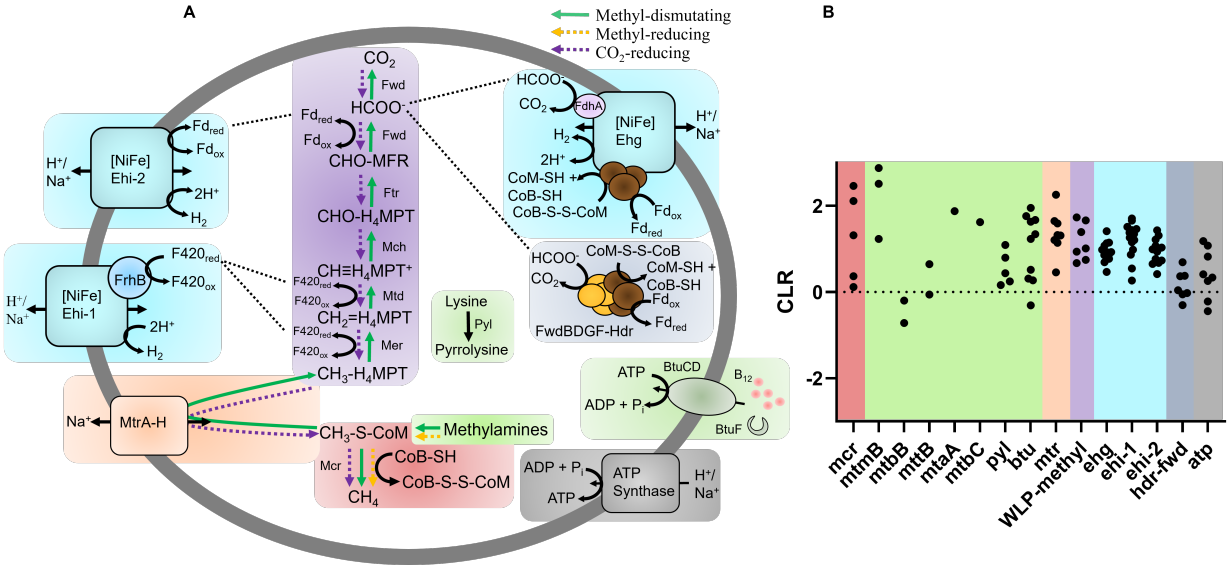


Figure 3: Proposed methane, hydrogen, and formate metabolism in *Ca.* M. fastidiosum YNP3N. (A) Metabolic reconstruction of methanogenesis, energy conservation, and vitamin B12 import in YNP3N. A list of genes used to construct this figure can be found in Supplementary Table 4. **(B)** Gene expression of YNP3N during growth on monomethylamine (n=1 due to low RNA yields) suggests a methyl-dismutating methanogenic lifestyle. Dotted line represents the geometric mean expression of all genes. Unless a specific subunit is listed, each dot represents the expression of a subunit in a multi-subunit enzyme or functional class, or expression of multiple copies of a particular gene. Explicit values can be found in Supplementary Table 5.

Additionally, YNP3N encodes several other [NiFe] hydrogenases, suggesting that internal hydrogen cycling mechanisms could fuel methyl-dismutating methanogenesis. One expressed membrane bound group 4 [NiFe]-hydrogenase complex (Ehi-1, Fig. 3, Fig. S3) is flanked by *frhB*, suggesting this complex is involved in the oxidation of reduced F_{420} that could be produced by oxidizing methyl groups via the WLP during methyl-dismutating methanogenesis. Strain YNP3N encodes the biosynthetic pathway for producing F_{420}, and we could not identify any other genes outside of *frhB* involved in recycling the F_{420} cofactor. A second expressed membrane bound group 4 [NiFe]-hydrogenase may deal with ferredoxin recycling produced by Fwd during methyl oxidation. This reduced ferredoxin could be oxidized by a second group 4 [NiFe] hydrogenase (Ehi-2, Fig. 3, Fig. S3) coupled with transport of ions across the membrane and the production of H_2. Recycling of hydrogen produced through methyl oxidation may be recaptured by the Ehg hydrogenase complex, allowing for the direct coupling of heterodisulfide reduction and energy conservation.

The overall flexible metabolism of YNP3N suggests the use of multiple substrate and electron donor pools (methyl groups, H_2, formate) may enable resilience during changing environmental conditions, as found in dynamic geothermal systems. Consistent with this, Methanonezhaarchaea have been observed to be stable populations and are often the dominant MCR-encoding organism in several hot spring systems (*24, 25*).

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.25.661132; this version posted June 25, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

Evolution and diversity of Methanonezhaarchaeia

Our surprising discovery of methylotrophic methanogenesis in the Methanonezhaarchaea, which previously had been predicted to only perform CO_2-reducing methanogenesis, led us to reassess the wider metabolic diversity of this lineage. We observed 13 different species clusters, comprising three families within the Methanonezhaarchaeales, which we propose as 'Methanonezhaarchaeceae' (Mnzh.; f__WYZ-LMO8), Muzhaarchaeceae' (Mza.; f__B40-G2), and 'Jinzhaarchaeceae' (Jza.; f__JAWCJE01) (Figure 4A, Supplementary Table 5).

Ca. M. fastidiosum YNP3N was the only genome found to contain genes for biosynthesis of pyrrolysine and methyltransferases for methylamines, indicating its unique capability for methylotrophic methanogenesis. We found that while nearly all Methanonezhaarchaeia species encode a complete or partial McrABGCD, all MAGs affiliated with its sister groups, Muzhaarchaeceae and Jinzhaarchaeceae, lacked these genes. This suggests a transition from a methanogenic to non-methanogenic lifestyle within the class Methanonezhaarchaeia. Consistent with this, all members of the Muzhaarchaeceae and Jinzhaarchaeceae, in addition to two species clusters within the Methanonezhaarchaeceae, encoded only a limited subset of the 38 methanogenesis marker proteins, including the one most closely related to strain YNP3N (87.4% ANI) (Figure 4B, Supplementary Table 6).

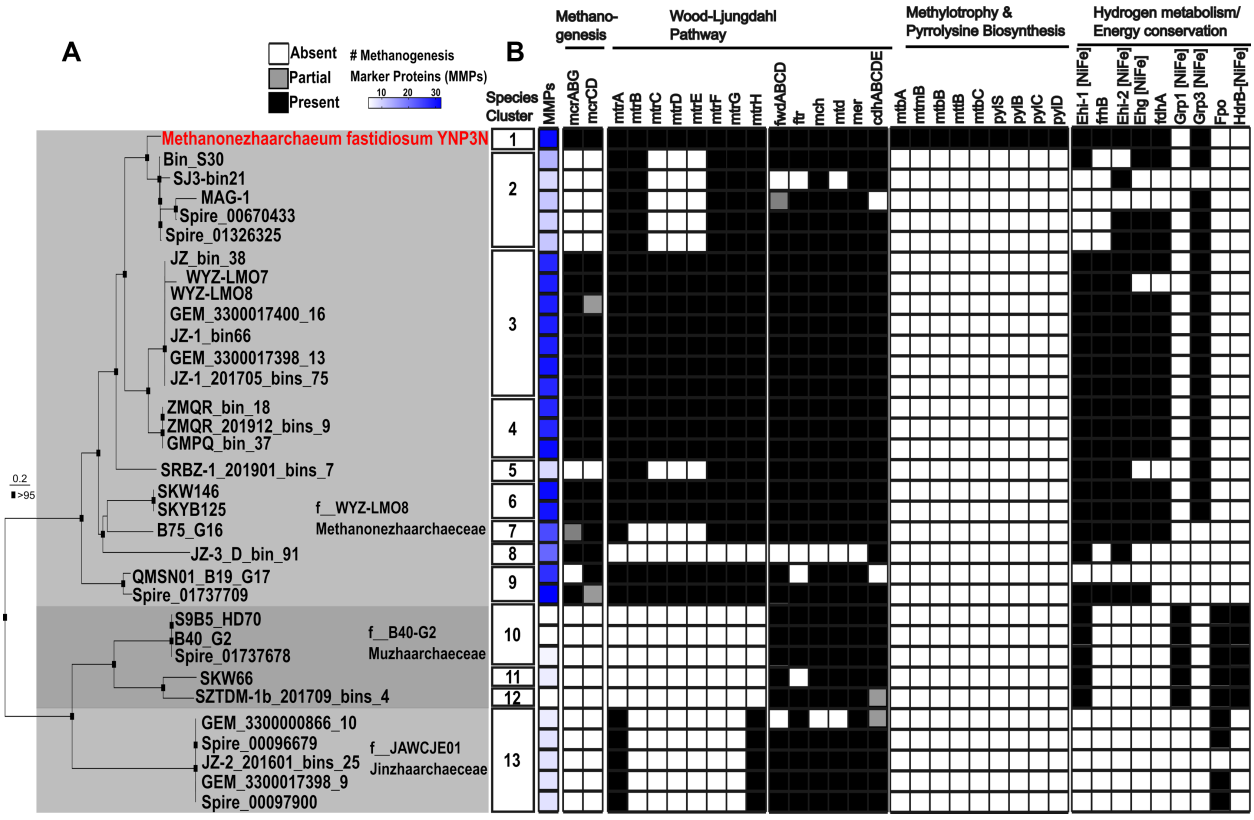


Figure 4. Comparative genomics of three families within Methanonezhaarchaeia. (A) Genome tree illustrating the diversity of Methanonezhaarchaeia. The phylogenetic tree shown here is a subset of the tree shown in Figure 2A. **(B)** Presence and absence of key metabolic genes and methanogenesis marker proteins.

Furthermore, different patterns of MTR loss were observed, with Muzhaarchaeceae seemingly having lost this complex altogether, while members of the Jinzhaarchaeceae only encode the MtrA and MtrH subunits. Five of nine species clusters within the Methanonezhaarchaeceae encode a full MTR complex, while one lacked it completely, and three species clusters were missing MtrCDE. MtrCDE subunits interact in sodium ion pumping, and their loss could indicate an important intermediate stage in the loss of methanogenic

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.25.661132; this version posted June 25, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

potential, by uncoupling MTR from energy conservation in species that have already lost MCR (Figure 4) (*41*). Furthermore, the *mtrA* gene found in Jinzhaarchaeceae has a separate evolutionary history from those found in Methanonezhaarchaeceae, as it branches with sequences from the Asgardarchaeota while Methanonezhaarchaeceae sequences branch with Bathyarchaeia (Figure S4). Together, this is further evidence for a complicated history of loss and horizontal gene transfer of the MTR subunits in addition to repurposing of at least the MtrAH subunits in diverse non-methanogenic lineages (*21, 22, 42*).

Compared to the pattern observed for MTR, the methyl and carbonyl branches of the WLP were largely conserved across all three families. This is consistent with the highly flexible nature of this pathway that is connected to methanogenesis through MTR, but also serves multiple separate catabolic and anabolic roles in methanogenic and non-methanogenic microbes, including the biosynthesis of purines, thymidylate, methione, and acetyl-CoA (*43-45*).

For Methanonezhaarchaeceae, the membrane associated energy conservation and redox modules are largely conserved across the 8 identified species clusters, with numerous group 4 [NiFe] hydrogenases, including the HdrB and FdhA associated Ehg complex and the FrhB-associated Ehi-2 complex (Figures 3, 4). For Muzhaarchaeceae, we observed a large remodeling of membrane associated complexes, with a group 1 [NiFe] hydrogenase, Fpo, and a group 4 HdrB-associated [NiFe] hydrogenase similar to the Ehd complex found in the Methanomethylicia and Culexarchaeia (Figure 4, Figure S3) (*5, 37*). In contrast, Jinzhaarchaeceae exhibit a more minimalistic electron transport chain, encoding a single Fpo-like complex. Further cultivation and high-quality genomes of diverse Methanonezhaarchaeia are required to further untangle the stepwise modifications of MTR, hydrogenases, and Fpo-like complexes that lead to a loss of methanogenic potential and the repurposing of the enzymatic remnants of methanogenesis into a new lifestyle.

Conclusion

Together, our results confirm methanogenesis in a third lineage of Thermoproteota, the Methanonezhaarchaeia. We expand the metabolic diversity of this group by cultivating a methylotrophic species from a group that hitherto had been assumed to only perform CO_2-reducing methanogenesis based on metagenomic predictions. These findings highlight the need for continued sampling of non-traditional methanogen genomes and cultures, which likely harbor further yet-unidentified diversity in their methanogenic pathways, and to test (meta)genomic prediction by experimentation. Multiple novel gene arrangements of methanogenesis genes and associated redox and energy conserving complexes were identified in diverse Methanonezhaarchaeia genomes and found to be expressed in the YNP-3N culture. This continues the trend of novel discoveries in methanogenic members of the Thermoproteota and identifies gaps in our understanding of the biochemistry and diversity of methanogens.

Methanogenesis was independently lost multiple times in archaea, leading to a patchy distribution of this metabolism (*5, 21, 22*). We further find that methanogenesis likely is not a feature of all members of the Methanonezhaarchaeia, and that this metabolic transition was driven by loss of methanogenesis genes (MCR, MTR) and remodeling of membrane associated energy conservation and redox modules. Obtaining more diverse cultures will aid in resolving the full metabolic diversity of these non-traditional methanogens and their close, yet uncultured, non-methanogenic relatives.

Etymology

YNP3N is a methanogen affiliated with the phylum Thermoproteota (*9*) for which we propose the candidate status Methanonezhaarchaeum fastidiosum sp. nov. For a protologue, please see the supplementary online information.

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.25.661132; this version posted June 25, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

Etymology: Me.tha.no.ne.zha.ar.chae'eum. N.L. pref. methano- pertaining to methane metabolism; N.L. neut. pl. n. archaeum, "ancient", referring to the cell's affiliation to the archaeal domain; N.L. neut. n. Methanonezhaarchaeum, methanogen named after Nezha, an immortal in Chinese mythology (*9*). fas.ti.di.o'sum L. neut, adj. fastidiosum, referring to the fastidious growth of this organism.
Locality: Sediment from a hot spring, identified as feature LCB003 (*25*).
Diagnosis: a thermophilic methyl-dismutating methanogen of the phylum Thermoproteota that grows as coccoid cells.

Methods

Source of inoculum and cultivation

Sediment was retrieved from hot spring LCB003 (44.57763, −110.78957; November 2022; 77 °C; pH 6.5) as previously described and stored at room temperature (*19, 25*). Enrichment cultures were established using anoxic medium as previously described (*19, 20*), with minor modifications. The medium had a pH of 6.5, containing a base of (per liter): KH_2PO_4, 0.5 g; $MgSO_4 \cdot 7H_2O$, 0.4 g; NaCl, 0.5 g; NH_4Cl, 0.4 g; $CaCl_2 \cdot 2H_2O$, 0.05 g; 2-N-morpholino-ethanesulfonic acid (MES), 2.17 g; yeast extract, 0.1 g; and 0.002% (w/v) $(NH_4)_2Fe(SO_4)_2 \cdot 6H_2O$, 5 mM $NaHCO_3$, 1 mL trace element solution SL-10, 1 mL Selenite-Tungstate solution, 1 mL CCM vitamins, 0.0005% (w/v) resazurin, 10 mg of coenzyme-M, 2 mg sodium dithionite, 1 mM dithiothreitol and 1 mM $Na_2S \cdot 9H_2O$. Sediment slurry was dispensed into serum bottles (10% v/v) and then sealed with thick butyl-rubber stoppers and aluminum crimps in an anoxic glove bag (Coy). The headspace of the enrichment was exchanged with N_2:CO_2 (90:10) gas for 5 minutes and set to 200 kPa. Monomethylamine (MMA) was added to a final concentration of 10 mM. Cultures were incubated at 77 °C upright in the dark and without shaking. While attempting to optimize our media and obtain a stable culture, we made the following modifications for the final media: (i) in the original incubation isopropanol was also added at a concentration of 10 mM, but was subsequently omitted from the medium, because it did not affect the rate of growth or final methane concentration (ii) replaced $MgSO_4$ with $MgCl_2 \cdot 6H_2O$ at 0.33 g (per liter), and (iii) removed choline from the CCM vitamins. Transfers of these batch cultures were made with 20% v/v inoculum into fresh medium when methane concentrations reached 8-12%. Dimethylamine (DMA) and trimethylamine (TMA) were also tested for methane production with and without isopropanol. However, these higher amines led to more inconsistent methane production profiles relative to MMA, and we focused on MMA.

DNA extraction and metagenome sequencing

Under anoxic conditions, 10 mL of the YNP3N culture was collected aseptically with a needle and syringe and concentrated on a sterile 0.2-micron polycarbonate membrane filter (Nuclepore) with vacuum filtration. The filter was then immediately placed into a Zymo BashingBead tube containing 750 µL of ZymoBIOMICS lysis solution. DNA was extracted using a ZymoBIOMICS DNA miniprep kit following the manufacturer's instructions. Illumina short-read library construction was performed with an Illumina DNA Prep kit following the manufacturer's instructions and was sequenced on an Illumina NovaSeq X Plus platform with 2 x 151 bp paired-end read chemistry performed at SeqCenter (Pittsburgh, PA, USA). Nanopore long-read sequencing was performed in-lab using a MinION platform and a R10.4.1 flow cell (Oxford Nanopore Technologies). Library preparation was performed using a rapid barcoding kit (SQK-RBK114.24) following the manufacturer's instructions and the flow cell was run for 27 hours. Base-calling was performed with dorado (v.0.7.1) using model sup@v.5.0.0. The resulting BAM file was demultiplexed using dorado demux (--emit-fastq).

Metagenome assembly, polishing, binning, and quality assessment

An initial Nanopore assembly was performed with metaFlye (*46*) (v.2.9.4), which produced a circular assembly of the strain YNP3N genome. This genome was further polished using Illumina short reads using

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.25.661132; this version posted June 25, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

one round of Polypolish (v.0.6.0) followed by PyPolca (v.0.3.1; --careful) as recently suggested, given the high sequence coverage of the YNP3N genome (*47, 48*).

A short-read assembly was also generated for analysis of co-enriched microbes, with raw metagenomic reads with Illumina read quality, linker and adaptor trimming, artefact and common contaminant removal performed using the rqcfilter2 pipeline (maxns=3, maq=20) and error correction with bbcms (mincount=2, hcf=0.6) from the BBTools suite v.38.94 (Bushnell B. 2014. BBMap: a fast, accurate, splice-aware aligner. https://sourceforge.net/projects/bbmap). Resulting reads were then assembled with megahit (v.1.2.9; --k-list 27,37,47,57,67,77,87,97,107,117,127,141 –min-count 1) (*49*). Coverage was determined by mapping reads to the assembly using minimap2 (v.2.26-r1175.; -x sr) (*50*). Contigs with a length of 2kb or greater were used for binning with SemiBin2 (v.2.0.2). CheckM (v.1.1.3) was used to assess completeness and redundancy of the resulting MAGs (*51*). GTDB-tk (v.2.4.0) with the R220 release was used for assigning initial taxonomy (*52*). CoverM (v.0.7.0) with default settings was used to generate relative abundances of each MAG (*53*). CoverM was also used to estimate the relative abundance of the YNP3N genome in hot spring LCB003 using quality filtered reads from a previous sediment metagenome (IMG Genome ID 3300028675) (*19*).

Annotation and reconstruction of metabolic potential

MAGs were annotated with Prokka (v.1.14.6), Interproscan (v.5.70-102.0) (*54, 55*). Annotation refinement and confirmation was done by submission of protein sequences to the NCBI conserved domain database and HHpred using default settings. Sequences of [NiFe] hydrogenases (PF00374) were checked for conserved CXXC motifs at the N and C terminus and submitted to HydDB for initial classification (*56*). Hydrogenase classification was further refined with phylogenetic analysis (below).

RNA extraction, sequencing, and analysis

Two replicate cultures were grown on 10 mM monomethylamine under a N_2/CO_2 headspace. RNA was extracted from these cultures during active methane production. Culture bottles were cooled from 77 °C by placing them in an ice bath that was then placed at -20 °C for 15 minutes, and vials were subsequently kept on ice during the following procedure. Cells were rapidly concentrated onto sterivex filters, which were then cut open and placed into a Zymo BashingBead lysis tube (Zymo Research) containing 800 µL of TRIzol (Invitrogen) and vortexed at maximum speed for 20 minutes. The tubes were centrifuged at 14,000x g for 15 minutes and the supernatant was used for RNA purification with the Zymo Direct-Zol kit according to the manufacturer's instructions and including the DNAse I step. RNA extract from the replicate cultures was pooled, and purified RNA was used for library preparation and sequencing at SeqCenter (Pittsburgh, PA, USA). RNA was treated again with DNAse I (Invitrogen) and library preparation was performed with Illumina's Stranded Total RNA Prep Ligation with Ribo-Zero Plus kit and 10 bp unique dual indices. The library was sequenced on a NovaSeq X Plus platform, producing paired 2 x 150 bp reads. Demultiplexing, quality control, and adaptor trimming was performed with bcl-conver (v.4.1.5). Reads were further processed with rqcfilter2 (rna=t, trimrnaadapter=t, qtrim=rl, trimq=10, maq=20, maxns=0, minlen=50, mlf=0.33, removeribo=f) and mapped to reference genomes from the YNP3N culture using salmon (v.1.10.2; --validateMappings --seqBias) (*57*). Reads mapping to rRNA genes were excluded before normalization. Reads mapping to each gene were normalized to gene length before calculating the center log ratio (CLR; $\log_{10}$).

Phylogenetic analysis

A set of 43 single-copy marker proteins was collected from Nezhaarchaeia MAGs and reference archaeal genomes as previously described (*20, 37*). Reference Nezhaarchaeia MAGs were collected from public databases and recent publications and only used if they were estimated at >75% completeness and had at least 37 out of 43 phylogenomic marker proteins (*58-60*). These markers were aligned with MAFFT-linsi, trimmed with trimAL (-gt 0.7) and concatenated to produce a final alignment of 9,913 positions (*61, 62*).

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.25.661132; this version posted June 25, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

IQtree2 was used to reconstruct a maximum-likelihood phylogenetic tree using the best fit LG+F+R10 model, and 1,000 ultrafast bootstraps (*63*).

The YNP3N McrA protein sequence was extracted and aligned with reference archaeal McrA sequences with MAFFT-linsi, trimmed with trimAL (-gt 0.5), producing a final alignment of 556 positions. IQtree2 was used to reconstruct a maximum-likelihood phylogenetic tree using the LG+C60+F+G model and 1,000 ultrafast bootstraps.

Reference archaeal 16S rRNA sequences ≥1000bp were aligned with the YNP3N 16S rRNA sequence using MAFFT-linsi and trimmed with trimAL (-automated1), producing an alignment of 1,498 positions. IQtree2 was used to reconstruct a maximum-likelihood phylogenetic tree using the best-fit SYM+R8 model and 1,000 ultrafast bootstraps.

Reference MtrA sequences were collected from diverse methanogenic and alkane degrading lineages and aligned with the YNP3N sequence using MAFFT-linsi and trimmed with trimAL (-automated1), producing an alignment of 242 positions. IQtree2 was used to reconstruct a maximum-likelihood phylogenetic tree using the best-fit LG+F+R6 model and 1,000 ultrafast bootstraps.

Reference sequences of catalytic subunits of group 1, 2, 3, and 4 [NiFe] hydrogenases (PF00374) and Fpo-like subunits (PF00346) were collected from HydDB and previous studies and aligned with sequences collected from Methanonezhaarchaeia genomes (*20, 56*). Sequences were aligned using MAFFT-linsi and trimmed with trimAL (-gt 0.5), producing an alignment of 423 positions. IQtree2 was used to reconstruct a maximum-likelihood phylogenetic tree using the LG+C60+F+G model and 1,000 ultrafast bootstraps.

Fluorescence *in situ* hybridization

Subsamples of the YNP3N culture were fixed using 2% paraformaldehyde for 1 hour at room temperature as previously described (*20*). An oligonucleotide probe targeting most 16S rRNA sequences of Nezhaarchaea (equivalent to Terrestrial Miscellaneous Crenarchaeal Group (TMCG) in Silva release 138) was designed using the probe design tool in ARB (Mnez277, 5'-ACGGCCCGTACCCGTTATCG-3'). DOPE-FISH was performed as previously described (*20*) using the Mnez probe doubly labelled with Cy3 fluorophores (obtained from IDT-DNA) and a hybridization time of 3 hours at 10% formamide. Mnez277 was used in combination with the general archaeal probe Arch915, which was doubly labeled with Alexa Fluor 647 and cells were counterstained with the DNA stain 4′,6-diamidino-2-phenylindole (DAPI). Negative controls used double-labeled NON-EUB338 performed in parallel.

Acknowledgments

We thank the US National Park Service for permitting work in YNP under permit number YELL-SCI-8010.

Funding

U.S. Department of Energy, Office of Science, Biological and Environmental Research DE-SC0025661 (RH)

U.S. National Science Foundation, Biological Oceanography OCE-2049445 (RH)

Simons and the Gordon and Betty Moore Foundations 737750 (RH)

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.25.661132; this version posted June 25, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

Author contributions

Conceptualization: AJK, RH

Methodology: AJK, SN

Investigation: AJK, SN

Visualization: AJK

Supervision: RH

Writing - original draft: AJK

Writing - review & editing: AJK, SN, RH

Competing interests: none

Data and materials availability:

Metagenomic and metatranscriptomic data were deposited on NCBI under the Project ID PRJNA1272160.

References

1. P. S. Garcia, S. Gribaldo, G. Borrel, Diversity and evolution of methane-related pathways in archaea. *Annual Review of Microbiology* **76**, 727-755 (2022).
2. R. K. Thauer, A.-K. Kaster, H. Seedorf, W. Buckel, R. Hedderich, Methanogenic archaea: ecologically relevant differences in energy conservation. *Nature Reviews Microbiology* **6**, 579-591 (2008).
3. M. Saunois *et al.*, Global methane budget 2000–2020. *Earth System Science Data* **17**, 1873-1958 (2025).
4. D. Mayumi *et al.*, Methane production from coal by a single methanogen. *Science* **354**, 222-225 (2016).
5. G. Borrel *et al.*, Wide diversity of methane and short-chain alkane metabolisms in uncultured archaea. *Nature microbiology* **4**, 603-613 (2019).
6. Z. Zhou *et al.*, Non-syntrophic methanogenic hydrocarbon degradation by an archaeal species. *Nature* **601**, 257-262 (2022).
7. L. J. McKay *et al.*, Co-occurring genomic capacity for anaerobic methane and dissimilatory sulfur metabolisms discovered in the Korarchaeota. *Nature microbiology* **4**, 614-622 (2019).
8. I. Vanwonterghem *et al.*, Methylotrophic methanogenesis discovered in the archaeal phylum Verstraetearchaeota. *Nature microbiology* **1**, 1-9 (2016).
9. Y. Wang, G. Wegener, J. Hou, F. Wang, X. Xiao, Expanding anaerobic alkane metabolism in the domain of Archaea. *Nature microbiology* **4**, 595-602 (2019).
10. Y.-F. Ou *et al.*, Expanding the phylogenetic distribution of cytochrome b-containing methanogenic archaea sheds light on the evolution of methanogenesis. *The ISME Journal* **16**, 2373-2387 (2022).
11. Z.-S. Hua *et al.*, Insights into the ecological roles and evolution of methyl-coenzyme M reductase-containing hot spring Archaea. *Nature communications* **10**, 4574 (2019).
12. K. Wu *et al.*, Isolation of a methyl-reducing methanogen outside the Euryarchaeota. *Nature* **632**, 1124-1130 (2024).
13. J. Hou *et al.*, Taxonomic and carbon metabolic diversification of Bathyarchaeia during its coevolution history with early Earth surface environment. *Science Advances* **9**, eadf5069 (2023).

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.25.661132; this version posted June 25, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

14. P. N. Evans *et al.*, Methane metabolism in the archaeal phylum Bathyarchaeota revealed by genome-centric metagenomics. *Science* **350**, 434-438 (2015).
15. D. Y. Sorokin *et al.*, Discovery of extremely halophilic, methyl-reducing euryarchaea provides insights into the evolutionary origin of methanogenesis. *Nature microbiology* **2**, 1-11 (2017).
16. M. M. Lynes, Z. J. Jay, A. J. Kohtz, R. Hatzenpichler, Methylotrophic methanogenesis in the Archaeoglobi revealed by cultivation of Ca. Methanoglobus hypatiae from a Yellowstone hot spring. *The ISME Journal* **18**, wrae026 (2024).
17. S. Buessecker *et al.*, Mcr-dependent methanogenesis in Archaeoglobaceae enriched from a terrestrial hot spring. *The ISME Journal* **17**, 1649-1659 (2023).
18. M. K. Nobu, T. Narihiro, K. Kuroda, R. Mei, W.-T. Liu, Chasing the elusive Euryarchaeota class WSA2: genomes reveal a uniquely fastidious methyl-reducing methanogen. *The ISME journal* **10**, 2478-2487 (2016).
19. V. Krukenberg, A. J. Kohtz, Z. J. Jay, R. Hatzenpichler, Methyl-reducing methanogenesis by a thermophilic culture of Korarchaeia. *Nature* **632**, 1131–1136 (2024).
20. A. J. Kohtz *et al.*, Cultivation and visualization of a methanogen of the phylum Thermoproteota. *Nature* **632**, 1118–1123 (2024).
21. Y. Wang *et al.*, A methylotrophic origin of methanogenesis and early divergence of anaerobic multicarbon alkane metabolism. *Science advances* **7**, eabj1453 (2021).
22. P. S. Adam, G. E. Kolyfetis, T. L. Bornemann, C. E. Vorgias, A. J. Probst, Genomic remnants of ancestral methanogenesis and hydrogenotrophy in Archaea drive anaerobic carbon cycling. *Science advances* **8**, eabm9651 (2022).
23. N. Dombrowski, A. P. Teske, B. J. Baker, Expansive microbial metabolic versatility and biodiversity in dynamic Guaymas Basin hydrothermal sediments. *Nature communications* **9**, 4999 (2018).
24. J. Wang *et al.*, Evidence for nontraditional mcr-containing archaea contributing to biological methanogenesis in geothermal springs. *Science Advances* **9**, eadg6004 (2023).
25. M. M. Lynes *et al.*, Diversity and function of methyl-coenzyme M reductase-encoding archaea in Yellowstone hot springs revealed by metagenomics and mesocosm experiments. *ISME communications* **3**, 22 (2023).
26. K. Lang *et al.*, New mode of energy metabolism in the seventh order of methanogens as revealed by comparative genome analysis of "Candidatus Methanoplasma termitum". *Applied and environmental microbiology* **81**, 1338-1352 (2015).
27. A. Mahapatra *et al.*, Characterization of a Methanosarcina acetivorans mutant unable to translate UAG as pyrrolysine. *Molecular microbiology* **59**, 56-66 (2006).
28. G. Borrel *et al.*, Comparative genomics highlights the unique biology of Methanomassiliicoccales, a Thermoplasmatales-related seventh order of methanogenic archaea that encodes pyrrolysine. *BMC genomics* **15**, 1-24 (2014).
29. B. Hao *et al.*, A new UAG-encoded residue in the structure of a methanogen methyltransferase. *Science* **296**, 1462-1466 (2002).
30. F. Xie *et al.*, Unraveling the phylogenomic diversity of Methanomassiliicoccales and implications for mitigating ruminant methane emissions. *Genome Biology* **25**, 32 (2024).
31. O. M. Sokolovskaya, A. N. Shelton, M. E. Taga, Sharing vitamins: Cobamides unveil microbial interactions. *Science* **369**, eaba0165 (2020).
32. T. Wagner, J. Koch, U. Ermler, S. Shima, Methanogenic heterodisulfide reductase (HdrABC-MvhAGD) uses two noncubane [4Fe-4S] clusters for reduction. *Science* **357**, 699-703 (2017).
33. K. C. Costa *et al.*, Protein complexing in a methanogen suggests electron bifurcation and electron delivery from formate to heterodisulfide reductase. *Proceedings of the National Academy of Sciences* **107**, 11050-11055 (2010).
34. T. Wagner, U. Ermler, S. Shima, The methanogenic CO2 reducing-and-fixing enzyme is bifunctional and contains 46 [4Fe-4S] clusters. *Science* **354**, 114-117 (2016).

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.25.661132; this version posted June 25, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

35. O. N. Lemaire, T. Wagner, All-in-One CO2 Capture and Transformation: Lessons from Formylmethanofuran Dehydrogenases. *Accounts of Chemical Research* **57**, 3512-3523 (2024).
36. T. Watanabe *et al.*, Three-megadalton complex of methanogenic electron-bifurcating and CO2-fixing enzymes. *Science* **373**, 1151-1156 (2021).
37. A. J. Kohtz, Z. J. Jay, M. M. Lynes, V. Krukenberg, R. Hatzenpichler, Culexarchaeia, a novel archaeal class of anaerobic generalists inhabiting geothermal environments. *ISME communications* **2**, 86 (2022).
38. M. F. Abdul Halim, L. A. Day, K. C. Costa, Formate-dependent heterodisulfide reduction in a Methanomicrobiales archaeon. *Applied and environmental microbiology* **87**, e02698-02620 (2021).
39. K. C. Costa *et al.*, Effects of H2 and formate on growth yield and regulation of methanogenesis in Methanococcus maripaludis. *Journal of bacteriology* **195**, 1456-1462 (2013).
40. P. Liu, Y. Lu, Concerted metabolic shifts give new insights into the syntrophic mechanism between propionate-fermenting Pelotomaculum thermopropionicum and hydrogenotrophic Methanocella conradii. *Frontiers in microbiology* **9**, 1551 (2018).
41. I. Aziz *et al.*, Structural and mechanistic basis of the central energy-converting methyltransferase complex of methanogenesis. *Proceedings of the National Academy of Sciences* **121**, e2315568121 (2024).
42. Y.-L. Qi *et al.*, Comparative genomics reveals thermal adaptation and a high metabolic diversity in "Candidatus Bathyarchaeia". *Msystems* **6**, 10.1128/msystems. 00252-00221 (2021).
43. W. F. Fricke *et al.*, The genome sequence of Methanosphaera stadtmanae reveals why this human intestinal archaeon is restricted to methanol and H2 for methane formation and ATP synthesis. *Journal of bacteriology* **188**, 642-658 (2006).
44. G. Borrel, P. S. Adam, S. Gribaldo, Methanogenesis and the Wood–Ljungdahl pathway: an ancient, versatile, and fragile association. *Genome biology and evolution* **8**, 1706-1711 (2016).
45. R. K. Thauer, The Wolfe cycle comes full circle. *Proceedings of the National Academy of Sciences* **109**, 15084-15085 (2012).
46. M. Kolmogorov *et al.*, metaFlye: scalable long-read metagenome assembly using repeat graphs. *Nature methods* **17**, 1103-1110 (2020).
47. G. Bouras *et al.*, How low can you go? Short-read polishing of Oxford Nanopore bacterial genome assemblies. *Microbial Genomics* **10**, 001254 (2024).
48. R. R. Wick, K. E. Holt, Polypolish: short-read polishing of long-read bacterial genome assemblies. *PLoS computational biology* **18**, e1009802 (2022).
49. D. Li, C.-M. Liu, R. Luo, K. Sadakane, T.-W. Lam, MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**, 1674-1676 (2015).
50. H. Li, Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094-3100 (2018).
51. D. H. Parks, M. Imelfort, C. T. Skennerton, P. Hugenholtz, G. W. Tyson, CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome research* **25**, 1043-1055 (2015).
52. P.-A. Chaumeil, A. J. Mussig, P. Hugenholtz, D. H. Parks, GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics* **38**, 5315-5316 (2022).
53. S. T. Aroney *et al.*, CoverM: read alignment statistics for metagenomics. *Bioinformatics* **41**, btaf147 (2025).
54. T. Seemann, Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068-2069 (2014).
55. P. Jones *et al.*, InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236-1240 (2014).
56. D. Søndergaard, C. N. Pedersen, C. Greening, HydDB: a web tool for hydrogenase classification and analysis. *Scientific reports* **6**, 34212 (2016).
57. R. Patro, G. Duggal, M. I. Love, R. A. Irizarry, C. Kingsford, Salmon provides fast and bias-aware quantification of transcript expression. *Nature methods* **14**, 417-419 (2017).

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.25.661132; this version posted June 25, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license.

58. T. S. Schmidt *et al.*, SPIRE: a Searchable, Planetary-scale mIcrobiome REsource. *Nucleic Acids Research* **52**, D777-D783 (2024).
59. S. Nayfach *et al.*, A genomic catalog of Earth's microbiomes. *Nature biotechnology* **39**, 499-509 (2021).
60. Y.-L. Qi *et al.*, Analysis of nearly 3000 archaeal genomes from terrestrial geothermal springs sheds light on interconnected biogeochemical processes. *Nature communications* **15**, 4066 (2024).
61. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution* **30**, 772-780 (2013).
62. S. Capella-Gutiérrez, J. M. Silla-Martínez, T. Gabaldón, trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972-1973 (2009).
63. B. Q. Minh *et al.*, IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular biology and evolution* **37**, 1530-1534 (2020).