Hydrogenotrophic methanogenesis in archaeal phylum Verstraetearchaeota reveals the shared ancestry of all methanogens

Bojk A. Berghuis[a], Feiqiao Brian Yu[a,b], Frederik Schulz[c], Paul C. Blainey[d,e], Tanja Woyke[c], and Stephen R. Quake[a,b,f,1]

[a]Department of Bioengineering, Stanford University, Stanford, CA 94305; [b]Chan Zuckerberg Biohub, San Francisco, CA 94158; [c]Department of Energy Joint Genome Institute, Walnut Creek, CA 94598; [d]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139; [e]Broad Institute of Harvard and MIT, Cambridge, MA 02142; and [f]Department of Applied Physics, Stanford University, Stanford, CA 94305

Methanogenic archaea are major contributors to the global carbon cycle and were long thought to belong exclusively to the euryarchaeal phylum. Discovery of the methanogenesis gene cluster methyl-coenzyme M reductase (Mcr) in the Bathyarchaeota, and thereafter the Verstraetearchaeota, led to a paradigm shift, pushing back the evolutionary origin of methanogenesis to predate that of the Euryarchaeota. The methylotrophic methanogenesis found in the non-Euryarchaota distinguished itself from the predominantly hydrogenotrophic methanogens found in euryarchaeal orders as the former do not couple methanogenesis to carbon fixation through the reductive acetyl-CoA [Wood–Ljungdahl pathway (WLP)], which was interpreted as evidence for independent evolution of the two methanogenesis pathways. Here, we report the discovery of a complete and divergent hydrogenotrophic methanogenesis pathway in a thermophilic order of the Verstraetearchaeota, which we have named Candidatus Methanohydrogenales, as well as the presence of the WLP in the crenarchaeal order Desulfurococcales. Our findings support the ancient origin of hydrogenotrophic methanogenesis, suggest that methylotrophic methanogenesis might be a later adaptation of specific orders, and provide insight into how the transition from hydrogenotrophic to methylotrophic methanogenesis might have occurred.

methanogenesis | archaea | evolution

All known methanogenic organisms belong exclusively to the archaeal domain of life. Methanogens are typically found in the oxygen-depleted environments of soils, sediments, and the intestinal tract of humans and animals (1). With an estimated combined annual production of 500 million tons of the greenhouse gas methane, methanogenic archaea are key contributors to the global carbon cycle and play an important role in climate change (2, 3). Until recently, all known methanogens belonged to the Euryarchaeota and were categorized into two classes (class I and class II). The hypothesis that methane metabolism originated early in the evolution of the Euryarchaeota (4) has since been challenged following the recent discovery of a putative methane metabolism in the archaeal phyla Bathyarchaeota (formerly the miscellaneous Crenarchaeota group) (5, 6) and Verstraetearchaeota (7).

Three major pathways of methanogenesis are known (8, 9): hydrogenotrophic, methylotrophic, and acetoclastic (Fig. 1*A*). The only enzyme that is present in all types of methanogenesis is methyl-coenzyme M reductase (Mcr), a Ni-corrinoid protein catalyzing the last step of methyl group reduction to methane (1, 10, 11). Hydrogenotrophic methanogenesis is the most widespread pathway (1) and has been suggested to represent the ancestral form of methane production (12). Class I methanogens (Methanopyrales, Methanococcales, and Methanobacteriales), as well as most class II methanogens (Methanomicrobiales, Methanocellales, and Methanosarcinales, with the exception of Methanomassiliicoccales) are hydrogenotrophs. They reduce CO_2 to CH_4 in six steps via the reductive acetyl-CoA or Wood–Ljungdahl pathway (WLP). The WLP is one of the most important processes for energy generation and carbon fixation (13). Here, H_2, or sometimes formate, is used as an electron donor (1, 11). To conserve energy, hydrogenotrophs couple the WLP to methanogenesis. This coupling is established by N5-methyl-tetrahydromethanopterin:coenzyme M methyltransferase (Mtr; also known as tetrahydromethanopterin S-methyltransferase), which transfers the methyl group from the WLP to coenzyme M. Mtr uses the free energy of methyl transfer to establish a Na^+-motive force across the membrane (14). Methyl coenzyme M reductase then reduces methyl-coenzyme M to methane, using coenzyme B as an electron donor. The established disulfide bond between these coenzymes is then broken again by heterodisulfide reductase (HdrABC/mvhADG). This cytoplasmic electron bifurcating complex concomitantly generates the reduced ferredoxin required for CO_2 reduction in the process (15, 16).

The Bathyarchaeota, Verstraetearchaeota, and Methanomassiliicoccales are, together with some Methanobacteriales and

Significance

Methane-producing microorganisms are thought to be among the earliest cellular life forms colonizing our planet, and are major contributors to the past and present global carbon cycle. Currently, all methanogens belong to the archaeal domain of life, and there is compounding evidence for a variety of methanogenic metabolisms among a wide distribution of archaeal phyla. However, the predominantly hydrogenotrophic (CO_2-fixing) Euryarchaeota are distinct from the recently discovered methylotrophic (biomass-degrading) noneuryarchaea, making the shared ancestry and origins of all methanogens unclear. We discovered hydrogenotrophic methanogenesis in a thermophilic order of the Verstraetearchaeota, a noneuryarchaeote. The Verstraetearchaeota, hitherto known as methylotrophs, unify the origins of methanogenesis and shed light on how organisms can evolve to adapt from hydrogenotrophic to methylotrophic methane metabolism.

EVOLUTION

Fig. 1. WLP coupled to methanogenesis in the Methanohydrogenales. (*A*) Different configurations for the associated or independent functioning of the archaeal version of the WLP and methanogenesis. Missing enzymatic complexes or pathways are shaded in gray. The following are shown: CO_2-reducing methanogenesis as present in the Methanohydrogenales as well as class I and class II methanogens without cytochromes (*Left*); methanogenesis by reduction of methyl compounds using H_2 as inferred in Bathyarchaeota BA1, and a potential link with the WLP in the absence of Mtr (*Left Center*); methanogenesis by reduction of methyl compounds using H_2 as present in the Methanomassiliicoccales and Methanomethyliales (*Right Center*); and carbon fixation using the archaeal WLP in the absence of methanogenesis, and proposal of a mechanism to generate low potential ferredoxin during sulfate reduction in the case of the Archaeoglobales (*Right*). Fd, ferredoxin. (*B*, *Left*) Genome-based phylogeny of the TACK superphylum genomes found in the OP dataset. The tree was inferred using a concatenated set of 56 marker genes, utilizing the superphylum containing the Diapherotrites, Parvarchaeota, Aenigmarchaeota, Nanoarchaeota, and Nanohaloarchaea (DPANN) as the outgroup. OP genomes found in the Korarchaeota, Thaumarchaeota, Bathyarchaeota, Geoarchaeota, Verstraetearchaeota, and Marsarchaeota phyla are shown in red alongside members or clades containing publicly available genomes; the Crenarchaeota contain only OP genomes. The ancestry of the WLP in the TACK superphylum (red branches) inferred through clades that retained the complete set of WLP homologs, as was shown previously (thin line) or in this study (thick line), is illustrated. The scale bar indicates substitutions per site, and Dataset S1 contains the concatenated alignment file used to calculate the tree. (*B*, *Right*) Presence (filled circle) or absence (empty circle) of key genes in the TACK clades. Hco, Hco subunits I + II; Mtr, Mtr subunits A–H; Rgy, reverse gyrase.

Methanosarcinales, methylotrophic methanogens. Here, methylated C_1 compounds, including methanol, methylamines, and methylsulfides, are first activated by specific methyltransferases (17). Typically, this reaction involves four C_1 compounds: one

in which the methyl group is oxidized to CO_2, while the remaining three methyl groups are reduced to methane by Mcr. As the noneuryarchaeal methanogens have been found to be exclusively methylotrophic, it was hypothesized that methylotrophic

methanogenesis is an independently evolved ancient pathway (6, 18).

Alongside the discovery of the key methanogenesis gene cluster Mcr outside the euryarchaeal phylum, the key WLP gene CO dehydrogenase/acetyl-CoA synthase (Cdh) and associated genes have been found to occur independent of the methanogenesis pathway (19). While first thought to uniquely exist in the nonmethanogenic euryarchaeal order Archaeoglobales as a remnant of its ancestral methane-cycling lifestyle (12, 20), the presence of the archaeal WLP in the absence of the Mcr and Mtr complexes has now been reported for an increasing number of archaeal lineages within the Bathyarchaeota (21, 22), the Altiarchaeales (23), the Hadesarchaea/MSBL-1 (24, 25), the Lokiarchaeota (26), and the Thorarchaeota (27), with likely more awaiting discovery. While mechanisms to generate reduced ferredoxin for CO_2 reduction by the archaeal WLP in the absence of methanogenesis remain to be elucidated (6), the hypothesis that WLP-coupled methanogenesis is unique to the Euryarchaeota seemed to hold true (28).

In this study, we collected and studied five samples from the obsidian pool (OP) hot spring in Yellowstone National Park (*SI Appendix*, Fig. S1 and Table S1). We were able to assemble 111 metagenome-assembled genomes (MAGs) through minimetagenomic (29) and metagenomic analysis of these samples. These genomes are predominantly archaeal (98 of 111, named OP bin 000–110) (30), and most of the MAGs discovered in these samples were affiliated with the superphylum originally containing the Thaumarchaeota, Aigarchaeota, Crenarchaeota, and Korachaeota (TACK) (Fig. 1B). A number of MAGs belonging to the strictly anaerobic crenarchaeal genus *Ignisphaera* contained all WLP genes, suggesting that some of the Crenarchaeota are, in fact, capable of carbon fixation through this pathway. One of the Verstraetearchaeota (OP bin 008) formed a deeply branching clade, together with two recently published genomes from the Jinze (JZ) hot spring sediment sample in Tengchong, Yunnan, China [JZ bins 37 and 38 (31)]. Surprisingly, this clade contained all genes coding for the full pathway of hydrogenotrophic methanogenesis coupled to carbon fixation through the WLP. This finding places the origin of hydrogenotrophic methanogenesis outside the euryarchaeal phylum, suggests that hydrogenotrophic and methylotrophic methanogenesis are more closely linked than once thought, and sheds light on how a transition from hydrogenotrophic to methylotrophic methanogenesis might have occurred.

Expansion of the TACK Superphylum by Archaeal MAGs. A genome tree comprising the 98 OP archaeal genomes from our data and 1,427 publicly available archaeal genomes was inferred using a concatenated set of 56 or 122 archaeal-specific marker genes (Fig. 1B and *SI Appendix*, Fig. S1). The majority ($n = 49$) of genomes fall within the TACK superphylum, of which 42 are

Crenarchaeota. Our analysis extends the archaeal tree of life by adding lineages to the Bathyarchaeota, Geoarchaeota, Marsarchaeota, and Verstraetearchaeota (Fig. 1B and Table 1). A comparative analysis based on the Kyoto Encyclopedia of Genes and Genomes orthology (32), Clusters of Orthologous Groups (COGs) (33), and Pfam (34) was performed on the OP archaeal bins and other members of clades containing lineages of interest (Fig. 1B and *SI Appendix*, Figs. S2–S4). OP bin 61 groups within the newly reported Marsarchaeota (35) (Fig. 1B). This placement is supported by functional analysis of the respective MAGs showing high similarity to traits typically found in the Marsarchaeota (*SI Appendix*, Fig. S2). In contrast, OP MAGs 107 and 108, as well as UBA168, which branches basally to the Marsarchaeota, lacked both heme copper cytochrome oxidase (Hco) subunits observed in other Marsarchaeota, suggesting that anaerobic metabolism in this subclade might exist. OP bin 46, an early branching member of the Geoarchaeota, contained both Hco subunits, whereas all other (20) clade members contained only subunit 1, suggesting aerobic metabolism. OP bin 54 groups together with the Verstraetearchaeota. Although the genome is incomplete (Table 1), it contains subunits A, B, G, and D of the Mcr gene cluster characteristic of this methylotrophic methanogenic phylum. OP bin 008 groups monophyletically with JZ bins 37 and 38 (31) as an early branching clade of the Verstraetearchaeota (Fig. 1B). While OP bin 008 lacks the 16S rRNA gene and JZ bin 38 contains an incomplete (502-bp) 16S rRNA sequence, JZ bin 37 full-length 16S rRNA sequences show 84% identity to Verstraetearchaeota V1–V3, 85% identity to V4, and 86% identity to JZ bin 30 (*SI Appendix*, Fig. S10). Partial 16S rRNA gene alignment of JZ bins 37 and 38 shows 90% identity, and JZ bin 38 shows 83% and 81% identity to JZ bin 30 and Verstraetearchaeota V1–V3, respectively. Based on these observations, we propose two new thermophilic orders in the Verstraetearchaeota, grouping JZ bins 38 and 37, as well as OP bin 008, as representatives of the first (Candidatus Methanohydrogenales) and JZ bin 30 as the first representative of the second (Candidatus Methanomediales; naming details are provided in *SI Appendix*).

Hydrogenotrophic Methanogenesis Genes in the Verstraetearchaeota, WLP in the Crenarchaeota. Comparative genomics of the two Verstraetearchaeota orders to the existing order Methanomethyliales revealed that the three orders share 438 of the 834 Methanomethyliales homologs (of COG genes, >30% global identity; *SI Appendix*, Fig. S6). The orders, contrary to the Methanomethyliales, have been sampled in hot springs; as such, all encode a reverse gyrase. The three Methanohydrogenales genomes share 477 gene homologs, 75 of which were not observed in the Methanomethyliales (*SI Appendix*, Fig. S6). The orders did not possess the Hco gene cluster, an indication of anaerobic metabolism. In addition, the Methanohydrogenales possessed Mcr complex (McrABGCD), all subunits of

Table 1. Summary statistics of archaeal phylogeny in the OP dataset

| Genomic bin	Phylogeny	Assembly size, bp	Contig N50, bp	No. of Contigs	Guanine-cytosine content	Completeness, * %	Contamination, * %
OP bin 008	Verstraetearchaeota	1,665,326	28,180	82	0.55	97.06	0.74
OP bin 010	Bathyarchaeota	1,422,446	19,095	88	0.42	84.11	0.93
OP bin 021	Geoarchaeota	1,147,812	12,080	106	0.32	55.23	0.00
OP bin 042	Bathyarchaeota	1,286,079	16,585	96	0.42	87.2	1.87
OP bin 046	Geoarchaeota	1,804,512	157,431	23	0.32	94.49	0.74
OP bin 107	Early Marsarchaeota	84,454	6,214	13	0.55	21.99	0.00
OP bin 054	Verstraetearchaeota	337,598	7,976	43	0.49	17.76	0.00
OP bin 061	Marsarchaeota	2,310,778	51,573	84	0.44	96.32	4.41
OP bin 108	Early Marsarchaeota	1,536,417	19,661	99	0.55	85.44	3.74

*As estimated by CheckM (47).

Mtr (MtrABCDEGH), and Hdr (HdrABCD), all three of which are methanogenesis gene clusters (Fig. 2A). The Methanohydrogenales, as well as the Methanomediales, possessed Cdh (CdhABGDE), formylmethanofuran dehydrogenase (Fwd; FwdABCDE), and coenzyme F420-reducing hydrogenase (Frh; FrhABDG), all of which are genes constituting the WLP. In contrast to the methylotrophic methanogenic Methanomethyliales (7), the Methanohydrogenales have retained all enzymes of the WLP, as well as all subunits of Mtr and HdrABC (Fig. 2A). This suggests that the Methanohydrogenales are able to perform hydrogenotrophic methanogenesis (Fig. 3A), a metabolism that has eluded discovery outside the Euryarchaeota until the present time.

The WLP genes were also observed in *Ignisphaera*, a genus in the crenarchaeal family Desulfurococcaceae (Figs. 1B and 2A). The Desulfurococcales, a predominantly anaerobic hyperthermophilic order, are known to have the capacity for autotrophic carbon fixation through the dicarboxylate/hydroxybutyrate cycle (13). OP bins 20, 22, 23, 30, and 31 all contained Cdh, Fwd, Ftr, and Fhr (Fig. 1B and *SI Appendix*, Fig. S4). This suggests that they are

capable of autotrophic carbon fixation through the WLP in a manner similar to the Archaeoglobales and other lineages. Contrary to most members of the Desulfurococcales, *Ignisphaera* lacked several enzymes associated with the dicarboxylate/hydroxybutyrate cycle (*SI Appendix*, Fig. S2), such as pyruvate phosphate dikinase, succinyl-CoA synthethase, and malate/lactate dehydrogenase, as well as fumarate reductase and dehydrogenase. The key enzyme 4-hydroxybutyryl–CoA dehydratase was also not found. These enzymes are observed in the other families within this order. Interestingly, several of these genes were observed in all Methanohydrogenales (*SI Appendix*, Fig. S9), suggesting that this order might be capable of autotrophic carbon fixation in more than one way.

The Methanohydrogenales do not contain genes for methanogenesis from methylated compounds (Fig. 2A), unlike several euryarchaeal orders, as well as the Bathyarchaeota, Archaeoglobales, and Desulfurococcales. JZ bin 30, the only representative of the Methanomediales, does contain trimethylamine methyltransferase (MttB), although it lacks Mtr and Mcr. The genetically encoded
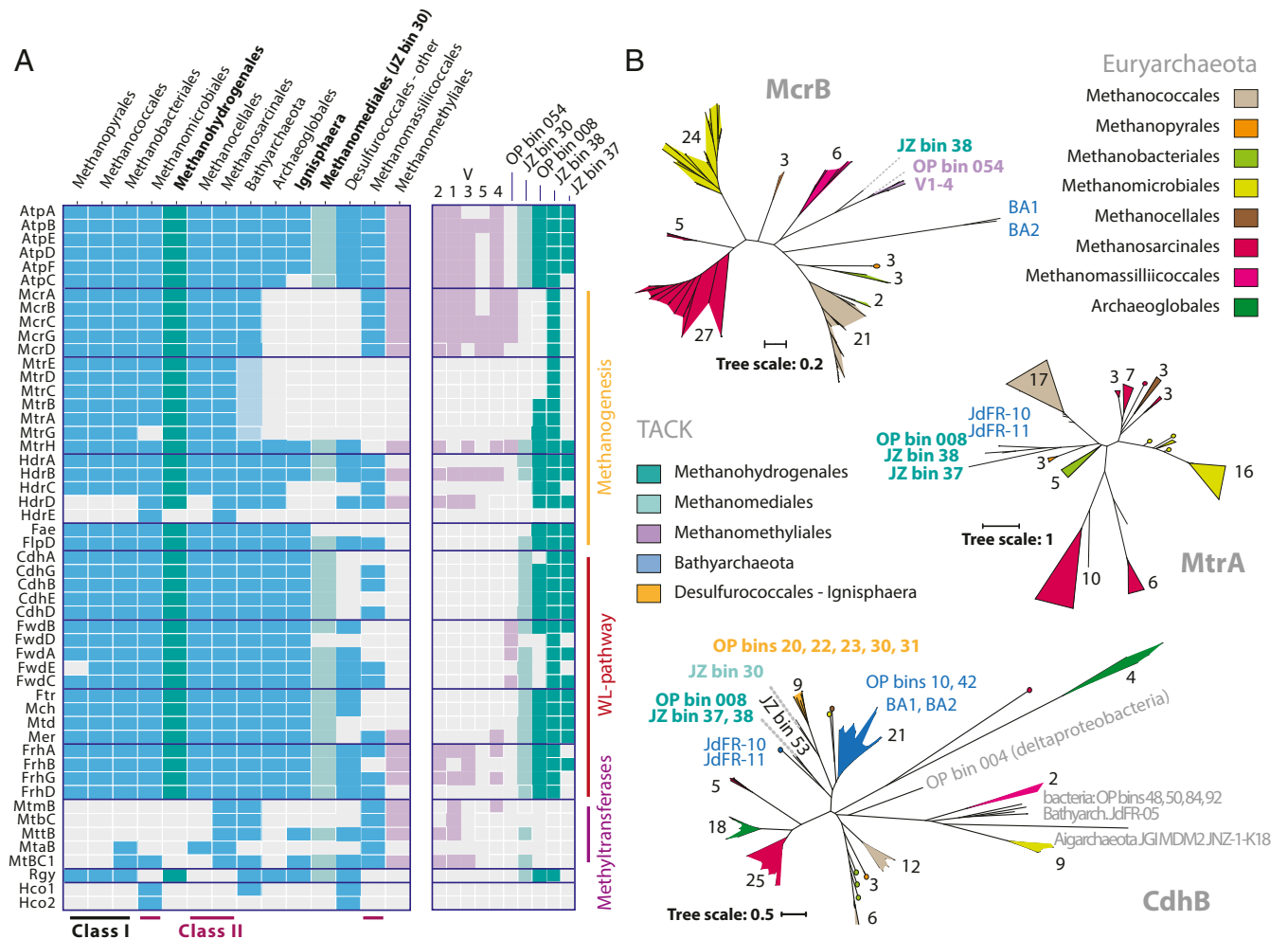
Fig. 2. Key WLP and methanogenesis genes in OP genomes and clades, compared with other methanogens. (A, Left) Presence of genes in the clades of the seven euryarchaeal methanogen groups [the Methanomethyliales (purple), the Bathyarchaeota, the Archaeoglobales, *Ignisphaera*, other Desulfurococcales, the Methanomediales (JZ bin 30, light green), and the Methanohydrogenales (green)], ordered by similarity (gene details are provided in *SI Appendix*, Table S2). The full Mtr gene cluster was observed in two Bathyarchaeota (light blue), but the presence of the full hydrogenotrophic methanogenesis pathway in a single genome has not been observed (*SI Appendix*, Fig. S4). (A, Right) Single-genome presence of genes in the Methanomethyliales (purple), the Methanomediales (light green), and the Methanohydrogenales (green). Additional genes are shown in *SI Appendix*, Fig. S3. (B) Gene trees showing the phylogeny of CdhB, McrB, and MtrA genes found in the OP MAGs relative to the seven existing methanogenic euryarchaeal orders, the Archaeoblobales, the Desulfurococcales, and the Bathyarchaeota (additional subunit trees are shown in *SI Appendix*, Figs. S8–S13). Numbers indicate the number of genomes present in a branch.

Fig. 3. Metabolism of the hydrogenotrophic methanogen Methanohydrogenales and the evolution of methanogenesis in archaea. (*A*) Proposed methanogenesis-related pathways and other metabolism of the three verstraetearchaeotal orders. The presence (colored circle) or absence (white circle) of a gene (cluster) is shown separately for the Methanohydrogenales (green), the Methanomediales (light green), and the Methanomethyliales (purple). The full pathway for autotrophic carbon fixation via the archaeal-type WLP (enzymes shown in cyan), methanogenesis (green) and associated ferredoxin regeneration by hdrABC cocomplexes mvh (purple, H_2 as an electron donor) and fdh (red, formate as an electron donor), energy conservation-related processes (orange), and respiratory complexes (blue), as well as related metabolic processes (brown) or nucleotide and amino acid synthesis capacities (brown background), are shown. (*B*, *Top*) Hydrogenotrophic (or CO_2-reducing) methanogenesis (blue) was previously exclusively observed in the Euryarchaeota, while methylotrophic methanogenesis (purple) was also found in the TACK superphylum. (*B*, *Bottom Left*) Discovery of the full hydrogenotrophic pathway in the Verstrae-tearchaeota, a member of the TACK superphylum, suggests that this pathway was present in the last common ancestor of the Verstraetearchaeota and Euryarchaeota. (*B*, *Bottom Right*) This observation also supports the hypothesis that methylotrophic methanogenesis is a later adaptation [including possible horizontal transfer of methyltransferases (green dashed line)] of the more basal CO_2-reducing form, and further strengthens the notion that the last universal common ancestor (LUCA) was a CO_2-reducing methanogen. *t*, time.

amino acid pyrrolysine was only observed in the methyltransferases (MtmB, MtbC, and MttB; *SI Appendix*, Fig. S19) of the Methanosarcinales and Methanomassilliicoccales, as previously

reported (36). The Methanohydrogenales, together with *Ignisphaera* and many Bathyarchaeota and Euryarchaeota from all classes, contain a formaldehyde-activating enzyme (Fae)

necessary for methanogenesis, while the Methanomethyliales lack this gene. As in the Methanomethyliales, homologs of the membrane-bound respiratory complex NADH-ubiquinone oxidoreductase (Nuo; *SI Appendix*, Fig. S3) are present and capable of reoxidizing reduced ferredoxin as occurs in the Methanosarcinales (37), with concomitant translocation of protons or sodium ions across the membrane (Fig. 3*A*). This, in turn can drive ATP synthesis via the archaeal type ATP synthase (AtpA–AtpF). The use of ferredoxin as an electron donor is further supported by the fact that the subunits required for the binding and oxidation of NADH (NuoEFG) are missing (*SI Appendix*, Fig. S3).

The hdrABC genes in the Methanohydrogenales are colocated with homologs for the Ni,Fe hydrogenase mvhADG (*SI Appendix*, Fig. S8). This electron bifurcating complex is known to facilitate the reduction of heterodisulfide (CoM-SS-CoB) as well as the production of reduced ferredoxin (16) (Fig. 3*A*). In addition, the presence of formate dehydrogenase genes in the same region (*SI Appendix*, Fig. S8) suggests that the Methanohydrogenales might be capable of using formate as an electron donor in a manner similar to *Methanococcus maripaludis* (38). As in the Methanomethyliales, hdrD is present and colocated with FAD-containing hydrogenase glcD (*SI Appendix*, Fig. S7). HdrD/glcD might also be capable of heterodisulfide reduction as has been suggested for the Methanomethyliales, the Bathyarchaeota, and *Archaeoglobus fulgidis* (7). The subunits belonging to an energy-conserving hydrogenase (eha/ehb) were not observed in the Methanohydrogenales and Methanomediales, contrary to the Methanomethyliales (*SI Appendix*, Fig. S3). As such, an alternative mechanism for the production of reduced ferredoxin needs to be present. The two orders contain the electron transferring flavoprotein subunits A and B (etfAB), and butyryl-CoA dehydrogenase (bcdA) was found in JZ bin 37. While this cytoplasmic electron bifurcating complex has been shown to generate reduced ferredoxin through the reduction of crotonyl-CoA to butyryl-CoA in *Clostridium kluyveri* (39), this mechanism has not been associated with energy conservation in methanogens. Another possibility would be the existence of an ech-like complex that is capable of producing reduced ferredoxin with the concomitant translocation of protons or sodium across the membrane. Several hydrogenases and ech-like genes are found in the hdrABC region of the Methanohydrogenales and Methanomediales. The Methanohydrogenales lack several transporters of complex molecules that are found in the Methanomethyliales. The Methanomethyliales contain genes for importing complex sugars, lipopolysaccharides, and oligopeptides (Fig. 3*A* and *SI Appendix*, Fig. S5). The ability of the Methanomethyliales to carry out complex fermentation, while the Methanohydrogenales do not, is likely a result of substrate diversification due to the organic compound-rich environment of the latter (active sludge) compared with the former (geothermal springs).

Evolutionary History of Hydrogenotrophic Methanogenesis in the Verstraetearchaeota. The full hydrogenotrophic methanogenesis pathway has not been previously identified in microorganisms outside the Euryarchaeota. To further explore the distribution and diversity, phylogenetic trees of Mcr, Mtr, and Cdh (or subunits thereof) were calculated (*SI Appendix*, Figs. S12–S18). Phylogenetic analysis of MtrA, MtrB, and MtrG, the three subunits in addition to MtrH that are present in OP bin 008 and JZ bin 38, resulted in a robust grouping of these two genomes in a monophyletic clade (Fig. 2*B* and *SI Appendix*, Figs. S12–S14), together with the two Bathyarchaeota genomes containing Mtr [Juan de Fuca Ridge 10 (JdFR-10) and JdFR-11; Fig. 2*B* and *SI Appendix*, Fig. S4] branching off at the root of the Euryarchaeota. This noneuryarchaeal clade is closely affiliated with the Methanobacteriales and (in the case of MtrA and MtrB) the Methanopyrales. The trees of McrA and McrB show that OP bin 54 branches off closely to the Verstraetearchaeota (V1–V5; Fig.

2*B* and *SI Appendix*, Fig. S15). JZ bin 38, the only genome in Methanohydrogenales that possesses the Mcr gene cluster, is present on the same branch as the other Verstraetearchaeota, suggesting a similar evolutionary trajectory of this gene cluster. JZ bin 38 branches in a basal position to OP bin 54 for McrB, and for McrA, OP bin 54 branches basally to JZ bin 38, together with the McrA genes of Bathyarchaeota BA1 and BA2. The CdhB, CdhD, and CdhG gene trees show more diversity in the TACK superphylum and the Bathyarchaeota as a result of the more common presence of the Cdh gene cluster throughout these clades (Fig. 2*B* and *SI Appendix*, Figs. S16 and S17). For McrB, the three Methanohydrogenales are closely related to JdFR-10 and JdFR-11, the two Bathyarchaeota with the full Mtr gene cluster (Fig. 2*B*). This clade forms the first branch of a group containing all Bathyarchaeota, *Ignisphaera*, and JZ bin 30. Methanohydrogenales have more divergent CdhD, yet form a single clade. JZ bin 30 branches ancestral to a clade containing *Ignisphaera* in CdhB, CdhD, and CdhG, showing divergent evolution of the acetyl-CoA synthase gene cluster in the Crenarchaeota, as well as branches leading toward the Verstraetearchaeota. CdhG shows that JZ bin 37 and Bathyarchaeota JdFR-10 and JdFR-11 diverge from the position of the other Methanohydrogenales as an early branch of the TACK superphylum and the Bathyarchaeota, showing a closer affiliation to the Methanopyrales.

The presence of subunit H of the Mtr gene cluster correlates with the presence of the gene clusters associated with autotrophic carbon fixation through the WLP. As such, MtrH shows a distinctly different evolutionary pattern than subunits A–G (*SI Appendix*, Fig. S14). It has been suggested that it plays a role in methylamine/coenzyme M methyl transfer activity in the Methanomethyliales (7), yet it is unclear what its role is in nonmethanogens where it is observed, such as the Archaeoglobales, *Ignisphaera*, and many Bathyarchaeota. Other genes typically associated with hydrogenotrophic methanogenesis are found throughout members of the Bathyarchaeota (Mtr in JdFR-10 and JdFR-11; *SI Appendix*, Figs. S4–S11), the Archaeoglobales (Fae, HdrA–HdrD, and MtrH), and *Ignisphaera* (Fae, HdrAB, MtrH, and EhbCDHI; Fig. 2*A* and *SI Appendix*, Fig. S11). In addition, OP bin 54, a member of the Methanomethyliales, contains an almost intact Fwd complex (FwdABCD; Fig. 2*A*), a key enzyme complex in the WLP. It is plausible that the presence of these genes might constitute remnants of ancestral hydrogenotrophic methanogenesis, as the various lineages described here adopted to other modes of energy production and metabolism.

The discovery of two orders of thermophilic Verstraetearchaeota sheds light on how the transition from autotrophic hydrogenotrophic methanogenesis to heterotrophic methylotrophic methanogenesis might have occurred. Decoupling of carbon fixation and methane production might have been achieved through the acquisition of methyltransferases. This could render the Mtr complex redundant, while maintaining the methanogenic mode of energy conservation, as has been proposed before (40). Signatures for this hypothesis can be found in the observation of two methyltransferases in the Methanomediales (JZ bin 30), the order more closely related to the Methanomethyliales, while the Methanohydrogenales lack methyltransferases (Fig. 2*A*). In addition, the preservation of the WLP in *Ignisphaera* further expands the presence of this ancient autotrophic carbon fixation pathway to the Crenarchaeota, suggesting that the WLP might have been present in the last common ancestor of this phylum, or the TACK superphylum as a whole. While other scenarios remain possible, this, in turn, would suggest that transitions from WLP-mediated autotrophic carbon fixation to other forms of autotrophic or heterotrophic energy metabolisms have occurred readily throughout the evolutionary history of the TACK superphylum.

The congruent topology of the species and gene trees, as well as the presence of the entire collection of genes of the

hydrogenotrophic methanogenesis pathway (Fig. 3*A*), suggests that these complexes have evolved as functional units. In addition, our findings suggest that hydrogenotrophic methanogenesis was present in the last common ancestor of the Verstraetearchaeota and Euryarchaeota (Fig. 3*B*). The presence of the hydrogenotrophic pathway in the TACK superphylum makes it plausible that methylotrophic methanogenesis is the later adaptation. While the methylotrophic pathway requires only a subset of the genes from the hydrogenotrophic pathway, the hydrogenotrophic pathway enables metabolism from a much simpler carbon source. This leaves us with genetic evidence for two possible evolutionary origins of methanogenesis (Fig. 3*B*). In the first scenario, hydrogenotrophic methanogenesis evolved first, supporting primitive cellular life in an environment that required using only CO_2 as a source of carbon. The simpler methylotrophic pathways then evolved through gene deletion as more complex nutrient environments became available. In the second scenario, methylotrophic methanogenesis evolved first as a means for cellular energy production. This would have taken place in an environment where complex organic molecules were readily available, after which the metabolic potential expanded through further gene addition as organisms eventually colonized ever more extreme environments. The presence of both methanogenesis pathways in a single noneuryarchaeal phylum, as well as the WLP in the Crenarchaeota, illustrates the rich metabolic capacity of the TACK superphylum. Our findings suggest that additional metabolic versatility awaits further discovery.

Methods

Environmental Sample Collection and Storage. The environmental samples used in this study were collected from the OP (44.610000°N; 110.438833°W) in Yellowstone National Park under permit number YELL-2009-SCI-5788. Five sediment samples were collected from different sections of the OP in the mud volcano area, placed in 2-mL tubes without any filtering, and soaked in 50% ethanol onsite (or glycerol). Details of each sampling site are included in *SI Appendix*, Table S1. After mixing with ethanol, samples were kept frozen until returning from Yellowstone National Park to Stanford University; at that time, tubes containing the samples were transferred to −80 °C for long-term storage.

Minimetagenomic Sequencing.
Microfluidic-based minimetagenomic sample preparation. Each sample was processed using a previously developed microfluidic-based minimetagenomic sequencing method (29). Briefly, after removing large particles and debris from each sample, cells were pelleted and resuspended in 1% NaCl. Cell concentration was adjusted to ~2 million cells per milliliter before loading onto a Fluidigm C1 microfluidic integrated fluidic circuit. We performed the same set of microfluidic lysis and multiple displacement amplification (MDA) steps as described in our previous work, with one modification. Instead of using Qiagen's REPLI-g Single Cell Kit for MDA on-chip, we adapted the SYGNIS TruePrime MDA chemistry to the microfluidic minimetagenomic method. We hoped that the primase-based whole-genome amplification (WGA) method would result in less bias compared with the random hexamers used in Qiagen's WGA kit (41). We also carried out microfluidic-based minimetagenomics using the Qiagen REPLI-g Single Cell Kit for one of the OP samples for the sake of comparison. The associated Fluidigm C1 instrumental control scripts and protocols are available through Fluidigm's ScriptHub. Following WGA via TruePrime MDA, amplified DNA from all minimetagenomic samples was harvested and Nextera XT libraries were created and sequenced on the Illumina HiSeq 2500 platform.
Shotgun metagenomics. Bulk genomic DNA was extracted from OP samples using MoBio's PowerSoil kit. Nextera XT V2 libraries were constructed and sequenced on Illumina's HiSeq 2500 platform under high-output mode. A total of 24, 20, 49, 6.6, and 104 million reads were obtained from OP samples 1 to 5, respectively (*SI Appendix*, Table S1), and trimmed using the same parameters as the minimetagenomic sequencing reads. Finally, assembly was performed using MEGAHIT (42), with default options and kmer values of 21, 31, 41, 51, 61, 71, and 99, as well as MetaSPAdes (43).

Sequence Assembly and Annotation. A previously developed custom bioinformatic pipeline was used to generate combined biosample contigs (29). The contigs were uploaded to the Joint Genome Institute's Integrated Microbial Genomes' Expert Review (IMG/ER) online database for annotation (44). Briefly, structural annotations were performed to identify CRISPRs (pilercr), tRNA (tRNAscan), and rRNA (hmmsearch). Protein coding genes were identified with a set of four ab initio gene prediction tools: GeneMark, Prodigal, MetaGeneAnnotator, and FragGeneScan. Finally, functional annotation was achieved by associating protein coding genes with COGs, Pfam descriptions, KEGG Orthology (KO) terms, and Enzyme Commission (EC) numbers. Phylogenetic lineage was assigned to each contig based on gene assignments.

Binning, Reassembly, Pruning, and Quality Assessment. Contig binning was performed by creating a 2D projection of the contig 5-mer space using the dimensionality reduction algorithm tSNE (45), followed by unsupervised clustering of grouped contigs using HDBSCAN (46). The redundancy of these bins, often containing a mix of contigs from the three different assemblers used (SPAdes for minimetagenomic and MetaSPAdes or MEGAHIT for the metagenomic contigs), was reduced by reassembling all of the reads mapping onto the contigs of a particular bin using SPAdes. The resulting reassembled bin was then further pruned or split based on cluster tightness observed from guanine-cytosine (GC) and shotgun coverage plots. The bin quality was assessed using CheckM; only bins with contamination <5% and strain heterogeneity <1% were used for further analysis (47).

Genome Tree Phylogeny. The 98 archaeal MAGs generated in this study were placed in phylogenetic context with 4,472 publicly available archaeal genomes available in the Integrated Microbial Genomes and Microbiomes (IMG/M) online database (ref. 48, database accessed March 2018) and recently published Marsarchaeota (35) using two sets of conserved phylogenetic marker proteins: 56 universal proteins (29) or 122 archaeal-specific marker genes (gtdb.ecogenomic.org/). In brief, marker proteins were identified with hmmsearch (version 3.1b2, hmmer.org/) using a specific hidden Markov model for each of the markers. For every protein, alignments were built with MAFFT (version 7.294b) (49) and subsequently trimmed with trimAl 1.4 (50), removing sites at which more than 90% of taxa contained missing information. Genomes with less than five marker proteins and genomes that had more than five duplicated marker proteins were removed from the alignment. Based on an initial tree built with FastTree 2 (REF, -spr 4 -mlacc 2 -slownni -lg), a subset of archaeal MAGs and reference genomes affiliated with the Geoarchaeota, Verstraetearchaeota, and Marsarchaeota was selected. These genomes were then used to build the final phylogenetic tree with IQ-TREE and the ultrafast bootstrap option [version 1.5.5; REF, LG+F+I+G4 -alrt 1000 -bb 1000 (51)] on an alignment with 11,134 informative sites for the concatenated alignment of the 56 universal marker proteins and 26,962 informative sites for the concatenated alignment of the 122 archaeal marker proteins.

Genome-Wide and 16S rRNA Average Nucleotide Identity. The 16S rRNA sequence average nucleotide identity (ANI) was calculated using the BLAST Global align Needleman–Wunsch algorithm (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch&PROG_DEF=blastn&BLAST_PROG_DEF=blastn&BLAST_SPEC=GlobalAln&LINK_LOC=BlastHomeLink). Genome-wide ANI was calculated using the ANI calculation tool of the IMG/ER web-based metagenome analysis toolset.

Gene Tree Phylogeny. Extracted gene protein sequences were aligned using MAFFT, using the local pair option (mafft-linsi), and the tree was calculated using IQ-TREE (51). The unrooted trees (*SI Appendix*, Figs. S8–S13 and S15) were visualized using the online web tool from the Interactive Tree of Life (https://itol.embl.de) and subsequently manually colored. The rooted tree of the full cdh gene cluster (*SI Appendix*, Fig. S14) was calculated with IQ-TREE, using the concatenated sequence of separately aligned subunits (mafft-linsi). The cdh gene cluster of OP bin 004 (a deltaproteobacterium) was used as the outgroup.

1. Thauer RK, Kaster AK, Seedorf H, Buckel W, Hedderich R (2008) Methanogenic archaea: Ecologically relevant differences in energy conservation. *Nat Rev Microbiol* 6:579–591.
2. Reeburgh WS (2007) Oceanic methane biogeochemistry. *Chem Rev* 107:486–513.
3. Intergovernmental Panel of Climate Change (2014) *Climate Change 2014: Synthesis Report*, eds Core Writing Team; Pachauri RK, Meyer LA (IPCC, Geneva).
4. Gribaldo S, Brochier-Armanet C (2006) The origin and evolution of archaea: A state of the art. *Philos Trans R Soc Lond B Biol Sci* 361:1007–1022.
5. Evans PN, et al. (2015) Methane metabolism in the archaeal phylum Bathyarchaeota revealed by genome-centric metagenomics. *Science* 350:434–438.
6. Borrel G, Adam PS, Gribaldo S (2016) Methanogenesis and the Wood-Ljungdahl pathway: An ancient, versatile, and fragile association. *Genome Biol Evol* 8:1706–1711.
7. Vanwonterghem I, et al. (2016) Methylotrophic methanogenesis discovered in the archaeal phylum Verstraetearchaeota. *Nat Microbiol* 1:16170.
8. Ferry JG, Kastead KA (2007) Methanogenesis. *Archaea: Molecular and Cellular Biology*, ed Cavicchioli R (ASM Press, Washington, DC), pp 288–314.
9. Conrad R (2009) The global methane cycle: Recent advances in understanding the microbial processes involved. *Environ Microbiol Rep* 1:285–292.
10. Hedderich R, Whitman WB (2013) Physiology and biochemistry of the methane-producing archaea. *The Prokaryotes*, eds Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F (Springer, Berlin Heidelberg), pp 635–662.
11. Liu Y, Whitman WB (2008) Metabolic, phylogenetic, and ecological diversity of the methanogenic archaea. *Ann N Y Acad Sci* 1125:171–189.
12. Bapteste E, Brochier C, Boucher Y (2005) Higher-level classification of the archaea: Evolution of methanogenesis and methanogens. *Archaea* 1:353–363.
13. Berg IA, et al. (2010) Autotrophic carbon fixation in archaea. *Nat Rev Microbiol* 8:447–460.
14. Schlegel K, Müller V (2013) Evolution of Na(+) and H(+) bioenergetics in methanogenic archaea. *Biochem Soc Trans* 41:421–426.
15. Costa KC, et al. (2010) Protein complexing in a methanogen suggests electron bifurcation and electron delivery from formate to heterodisulfide reductase. *Proc Natl Acad Sci USA* 107:11050–11055.
16. Kaster A-K, Moll J, Parey K, Thauer RK (2011) Coupling of ferredoxin and heterodisulfide reduction via electron bifurcation in hydrogenotrophic methanogenic archaea. *Proc Natl Acad Sci USA* 108:2981–2986.
17. Lang K, et al. (2015) New mode of energy metabolism in the seventh order of methanogens as revealed by comparative genome analysis of "Candidatus methanoplasma termitum". *Appl Environ Microbiol* 81:1338–1352.
18. Williams TA, et al. (2017) Integrative modeling of gene and genome evolution roots the archaeal tree of life. *Proc Natl Acad Sci USA* 114:E4602–E4611.
19. Adam PS, Borrel G, Gribaldo S (2018) Evolutionary history of carbon monoxide dehydrogenase/acetyl-CoA synthase, one of the oldest enzymatic complexes. *Proc Natl Acad Sci USA* 115:E1166–E1173.
20. Vorholt J, Kunow J, Stetter KO, Thauer RK (1995) Enzymes and coenzymes of the carbon monoxide dehydrogenase pathway for autotrophic CO2 fixation in Archaeoglobus lithotrophicus and the lack of carbon monoxide dehydrogenase in the heterotrophic A. profundus. *Arch Microbiol* 163:112–118.
21. Lazar CS, et al. (2015) Environmental controls on intragroup diversity of the uncultured benthic archaea of the miscellaneous Crenarchaeotal group lineage naturally enriched in anoxic sediments of the White Oak River estuary (North Carolina, USA). *Environ Microbiol* 17:2228–2238.
22. He Y, et al. (2016) Genomic and enzymatic evidence for acetogenesis among multiple lineages of the archaeal phylum Bathyarchaeota widespread in marine sediments. *Nat Microbiol* 1:16035.
23. Probst AJ, et al. (2014) Biology of a widespread uncultivated archaeon that contributes to carbon fixation in the subsurface. *Nat Commun* 5:5497.
24. Baker BJ, et al. (2016) Genomic inference of the metabolism of cosmopolitan subsurface archaea, Hadesarchaea. *Nat Microbiol* 1:16002.
25. Mwirichia R, et al. (2016) Metabolic traits of an uncultured archaeal lineage–MSBL1–from brine pools of the Red Sea. *Sci Rep* 6:19181.
26. Martin WF, Neukirchen S, Zimorski V, Gould SB, Sousa FL (2016) Energy for two: New archaeal lineages and the origin of mitochondria. *BioEssays* 38:850–856.
27. Seitz KW, Lazar CS, Hinrichs KU, Teske AP, Baker BJ (2016) Genomic reconstruction of a novel, deeply branched sediment archaeal phylum with pathways for acetogenesis and sulfur reduction. *ISME J* 10:1696–1705.
28. Sorokin DY, et al. (2017) Discovery of extremely halophilic, methyl-reducing euryarchaea provides insights into the evolutionary origin of methanogenesis. *Nat Microbiol* 2:17081.
29. Yu FB, et al. (2017) Microfluidic-based mini-metagenomics enables discovery of novel microbial lineages from complex environmental samples. *eLife* 6:e26580.
30. Berghuis et al. (2019) Hydrogenotrophic methanogenesis in archaeal phylum Verstraetearchaeota reveals the shared ancestry of all methanogens: Data files IMG database. Available at https://img.jgi.doe.gov/cgi-bin/m/main.cgi. Deposited January 9, 2019.
31. Eloe-Fadrosh EA, et al. (2016) Global metagenomic survey reveals a new bacterial candidate phylum in geothermal springs. *Nat Commun* 7:10476.
32. Kanehisa M, Goto S (2000) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28:27–30.
33. Tatusov RL, Galperin MY, Natale DA, Koonin EV (2000) The COG database: A tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res* 28:33–36.
34. Punta M, et al. (2012) The Pfam protein families database. *Nucleic Acids Res* 40:D290–D301.
35. Jay ZJ, et al. (2018) Marsarchaeota are an aerobic archaeal lineage abundant in geothermal iron oxide microbial mats. *Nat Microbiol* 3:732–740.
36. Borrel G, et al. (2014) Unique characteristics of the pyrrolysine system in the 7th order of methanogens: Implications for the evolution of a genetic code expansion cassette. *Archaea* 2014:374146.
37. Welte C, Deppenmeier U (2014) Bioenergetics and anaerobic respiratory chains of aceticlastic methanogens. *Biochim Biophys Acta* 1837:1130–1147.
38. Costa KC, Lie TJ, Jacobs MA, Leigh JA (2013) H2-independent growth of the hydrogenotrophic methanogen Methanococcus maripaludis. *MBio* 4:e00062-13.
39. Li F, et al. (2008) Coupled ferredoxin and crotonyl coenzyme A (CoA) reduction with NADH catalyzed by the butyryl-CoA dehydrogenase/Etf complex from Clostridium kluyveri. *J Bacteriol* 190:843–850.
40. Fricke WF, et al. (2006) The genome sequence of Methanosphaera stadtmanae reveals why this human intestinal archaeon is restricted to methanol and H2 for methane formation and ATP synthesis. *J Bacteriol* 188:642–658.
41. Picher AC, et al. (2016) TruePrime is a novel method for whole-genome amplification from single cells based on TthPrimPol. *Nat Commun* 7:13296.
42. Li D, Liu CM, Luo R, Sadakane K, Lam TW (2015) MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31:1674–1676.
43. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA (2017) metaSPAdes: A new versatile metagenomic assembler. *Genome Res* 27:824–834.
44. Huntemann M, et al. (2016) The standard operating procedure of the DOE-JGI Metagenome Annotation Pipeline (MAP v.4). *Stand Genomic Sci* 11:17.
45. Van Der Maaten LJP, Hinton GE (2008) Visualizing high-dimensional data using t-sne. *J Mach Learn Res* 9:2579–2605.
46. McInnes L, Healy J, Astels S (2017) hdbscan: Hierarchical density-based clustering. *J Open Source Softw* 2:205.
47. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW (2015) CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25:1043–1055.
48. Chen IA, et al. (2017) IMG/M: Integrated genome and metagenome comparative data analysis system. *Nucleic Acids Res* 45:D507–D516.
49. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol Biol Evol* 30:772–780.
50. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T (2009) trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25:1972–1973.
51. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ (2015) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 32:268–274.