Expanding the phylogenetic distribution of cytochrome b-containing methanogenic archaea sheds light on the evolution of methanogenesis

Ya-Fei Ou1,8, Hong-Po Dong1,8, Simon J. McIlroy2, Sean A. Crowe1,3,4, Steven J. Hallam6, Ping Han5, Jens Kallmeyer6, Rachel L. Simister1,4, Aurele Vuillemin6, Andy O. Leu2, Zhanfei Liu7, Yan-Ling Zheng5, Qian-Li Sun1, Min Liu5, Gene W. Tyson2 and Li-Jun Hou1,8

© The Author(s) 2022

Methane produced by methanogenic archaea has an important influence on Earth’s changing climate. Methanogenic archaea are phylogenetically diverse and widespread in anoxic environments. These microorganisms can be divided into two subgroups based on whether or not they use b-type cytochromes for energy conservation. Methanogens with b-type cytochromes have a wider substrate range and higher growth yields than those without them. To date, methanogens with b-type cytochromes were found exclusively in the phylum “Ca. Halobacteriota” (formerly part of the phylum Euryarchaeota). Here, we present the discovery of metagenome-assembled genomes harboring methyl-coenzyme M reductase genes reconstructed from mesophilic anoxic sediments, together with the previously reported thermophilic “Ca. Methylarchaeum tenchongensis”, representing a novel archaeal order, designated the “Ca. Methylarchaeales”, of the phylum Thermoproteota (formerly the TACK superphylum). These microorganisms contain genes required for methyl-reducing methanogenesis and the Wood-Ljundahl pathway. Importantly, the genus “Ca. Methanotowutia” of the “Ca. Methylarchaeales” encode a cytochrome b-containing heterodisulfide reductase (HdrDE) and methanophenazine-reducing hydrogenase complex that have similar gene arrangements to those found in methanogenic Methanosarcinales. Our results indicate that members of the “Ca. Methylarchaeales” are methanogens with cytochromes and can conserve energy via membrane-bound electron transport chains. Phylogenetic and amalgamated likelihood estimation analyses indicate that methanogens with cytochrome b-containing electron transfer complexes likely evolved before diversification of Thermoproteota or “Ca. Halobacteriota” in the early Archean Eon. Surveys of public sequence databases suggest that members of the lineage are globally distributed in anoxic sediments and may be important players in the methane cycle.

The ISME Journal (2022) 16:2373–2387; https://doi.org/10.1038/s41396-022-01281-0

INTRODUCTION
Methane is an important greenhouse gas with an atmospheric concentration that has more than doubled since the start of the industrial revolution [1], which is having a profound influence on Earth’s climate. Carbon isotope studies reveal that biogenic methane production, primarily from wetlands and agricultural sources [1], is responsible for the observed rapid increase. Biological methanogenesis by methanogenic archaea (methanogens) accounts for ~74% of global methane emissions [2]. For many years it was assumed that the methanogens were phylogenetically restricted to the phylum Euryarchaeota, which has recently been reclassified as a superphylum consisting of three separate phyla (“Ca. Halobacteriota”, Methanobacteria and “Ca. Thermoplasmata”), the genome Taxonomy Database (GTDB; Release 95) [3]. Recently, metagenome-assembled genomes (MAGs) from several uncultured lineages within the Thermoproteota (former TACK superphylum) have been inferred to be capable of methanogenesis, greatly expanding the phylogenetic diversity of lineages possessing this metabolism. These lineages include members of the orders “Ca. Methanomethylicales” (former phylum “Ca. Verstraetearchaeota”), “Ca. Nezhaarchaeales” (former phylum “Ca. Nezhaarchaeota”), the classes “Ca. Korarchaeia” (former phylum “Ca. Korarchaeota”) and Nitrososphaeria [7] (former phylum “Ca. Korarchaeota”) and Nitrososphaeria [7] (former phylum “Ca. Bathyarchaeota”). Members of the class “Ca. Bathyarchaeia” (former phylum “Ca. Bathyarchaeota”) and the order “Ca. Helarchaeales” (former phylum “Ca. Helarchaeota”) also contain methyl-coenzyme M reductase (Mcr) complex, which is the key enzyme for methane metabolism, but are suggested to more likely oxidize short-chain alkanes [8–10].

1State Key Laboratory of Estuarine and Coastal Research, East China Normal University, Shanghai 200241, China. 2Centre for Microbiome Research, School of Biomedical Sciences, Queensland University of Technology (QUT), Translational Research Institute, Woolloongabba, QLD 4102, Australia. 3Ecosystem Services, Commercialization Platforms, and Entrepreneurship (ECOSCOPE) Training Program, University of British Columbia, Vancouver, BC, Canada. 4Department of Microbiology and Immunology, University of British Columbia, Vancouver, BC, Canada. 5Key Laboratory of Geographic Information Science, Ministry of Education, East China Normal University, Shanghai 200241, China. 6GFZ German Research Centre for Geosciences, Section Geomicrobiology, Potsdam, Germany. 7Marine Science Institute, The University of Texas at Austin, Port Aransas, TX 78373, USA. 8These authors contributed equally: Ya-Fei Ou, Hong-Po Dong.

Received: 23 February 2022 Revised: 20 June 2022 Accepted: 22 June 2022
Published online: 9 July 2022
Methanogens can be divided into three groups based on substrate use: hydrogenotrophic, aceticlastic, methylotrophic, and methyl-reducing [11]. Hydrogenotrophic methanogens reduce CO2 to CH4 using electrons from H2 [11]. They are the most widely distributed methanogens and have been discovered in most methanogenic lineages of the Methanobacteriota and “Ca. Haloarchaeota” [12, 13]. Aceticlastic methanogens generate CH4 and CO2 by disproportionation of acetate, in which the carbonyl group is oxidized to provide electrons for reduction of methyl group to methane [12]. They have been observed only in the class “Ca. Methanosarcina” [14]. Methylotrophic methanogens use methylated compounds such as methylamines, methanol and methyl sulfides as carbon and energy sources. Based on studies of cultured representatives, only members of the Methanosarcinales are found to be capable of performing methylotrophic methanogenesis [12].

As for methyl-reducing methanogens, methyl compounds cannot be oxidized to CO2 but are reduced to methane using electrons derived from H2 or formate [12, 15]. The cultivated representatives from the Methanomassiliicoccales, the Methanotropharchaeota and Methanophaga have been shown to utilize this methyl-reducing pathway for methanogenesis [16–18]. The recently discovered “Ca. Methanomethylicales” and “Ca. Methanospiridiosa” based on metagenomic assembly are inferred to be also likely to depend on this pathway [4, 19]. Based on the difference in energy-conserving systems, all methanogenic archaea can also be classified into two main subgroups: methanogens with and without b-type cytochromes [12, 13]. To our knowledge, within cultivated organisms, cytochrome b-containing methanogens have a wider substrate range, and are able to use CO2 plus H2, acetate or methylated compounds as substrates, whereas methanogens without b-type cytochromes are either hydrogenotrophic or methyl-reducing [12, 13]. In addition, cytochrome b-containing methanogens also have higher growth yields than methanogens without b-type cytochromes owing to use of membrane-bound electron transport chains [12, 13]. As methanogens with b-type cytochromes have been exclusively found in the “Ca. Halobacteriota” of the Euryarchaeota superphylum, it has been suggested that the metabolism originated within this phylum.

Here, we present the discovery of seven MAGs containing mcr genes recovered from anoxic sediments that belong to novel genera within the family “Ca. Methylocalarchaea” of the phylum Thermoproteota. Importantly, these putative methanogenic archaea encode cytochrome b-containing complexes and are predicted to conserve energy via membrane-bound electron transport chains, which expands the known phylogenetic diversity of cytochrome b-containing methanogens and enhances our understanding of their evolutionary history.

RESULTS AND DISCUSSION
Discovery of a novel archaeal lineage in wetland sediments
To examine archaeal community composition and function in a mangrove ecosystem, we analyzed metagenomic data from 13 sediment samples taken from mangrove wetlands in Techeng Island of Zhanjiang and Dongzhai Harbour of Haikou, China (Supplementary Fig. 1). De novo assembly of these sequencing data (60–120 Gbp for each sample) and genome binning resulted in 242 archaeal MAGs (>70% complete; <10% contamination) (Supplementary Table 1). Five MAGs (H03B1, HK01M, HK01B, HK02M1, and HK02M2) were found to contain genes encoding a complete methyl-coenzyme M reductase complex (mcrABCDG) (Table 1). Based on the Genome Taxonomy Database Toolkit (GTDB-Tk) [3, 20], these MAGs were classified as a novel order within the class Nitrosophaerina (former phylum Thaumarchaeota) of the phylum Thermoproteota (former TACK superphylum) (Fig. 1 and Supplementary Fig. 2).

Comparative analyses revealed that the McrA sequences from these MAGs are distantly related to extant sequences in the NCBI...
Two metagenomes in IMG database generated from sediments of Lake Towuti, Indonesia (Supplementary Fig. 1 and Supplementary Table 2). Two additional related MAGs (TDP8 and TDP10, Table 1) encoding the complete Mcr complex were subsequently recovered from these metagenomes. For these MAGs (with exception of HK01M), the mcrABG operon and other genes related to methane
metabolism were located on long contigs \( \geq 11,476 \text{ bp} \) whose sequence composition features were consistent with their corresponding genomes (Supplementary Fig. 3), supporting the accurate assignment of these contigs to each MAG. The estimated genome size range for the seven MAGs recovered was \( 1.06–2.55 \) Mbp with total number of coding sequences ranging from 1151 to 3291. We examined vertical distribution of these MAGs in sediment cores of two sampling sites and found that their relative abundance increased gradually as depth increased from 15 to 100 cm (Supplementary text; Supplementary Fig. 4). Subsequent searches of public sequencing databases using the 16S rRNA and mcrA gene sequences annotated in these MAGs identified related species in freshwater lake sediments, hot springs, mangrove wetlands, rice paddy soils, hydrothermal vents, and deep-sea sediments distributed in different regions of the world (Supplementary text; Supplementary Table 3 and Supplementary Fig. 5).

Phylogenetic analysis using 122 concatenated archaeal-specific marker proteins revealed that the seven MAGs and “Ca. M. tengchongensis” formed a distinct lineage that is sister to the order *Nitrososphaeraeae* (Fig. 1a and Supplementary Fig. 2b). Phylogenetic analyses of the 16S and 23S rRNA genes recovered from these MAGs supported the novelty of this lineage (Supplementary Table 4 and Supplementary Fig. 2a), with pairwise nucleotide comparisons of 16S rRNA genes revealing an identity of 79.1–87.3% to publicly available *Nitrososphaerinae* genera (Supplementary Table 5). The seven MAGs belonging to the novel lineage had an AAI of 44.0–52.3% to all other genomes of the *Nitrososphaerinae* (Supplementary Table 6), further supporting their classification as a separate order [21, 22]. Collectively, these phylogenetic analyses indicate that these MAGs represent four different genera of the recently described family “Ca. Methylarchaeaeae” within a novel order—designated here as “Ca. Methylarchaeaeae” (Fig. 1a and Supplementary Fig. 2) and Supplementary Tables 5 and 6). H03B1, HK01M, HK01B, and HK02M1 represent one genus (69.7–80% AAI to other MAGs). HK02M2 represents the second (68.9–80% AAI to other MAGs), TDP8 and TDP10 represent the third (70.2–82.5% AAI), and “Ca. M. tengchongensis” represents the fourth (68.9–82.5% AAI); the former three genera are named here “Ca. Methanoinusilarius”, “Ca. Methanoprotocila”, and “Ca. Methanotowutia”, respectively.

The “Ca. Methylarchaeaeae” are potentially methyl-reducing methanogens with b-type cytochromes

Annnotation of the eight “Ca. Methylarchaeaeae” MAGs confirmed genes involved in archaeal methane metabolism (Supplementary Table 7 and Fig. 2), including those encoding the Mcr complex (*mcrABDG* and auxiliary genes *mcrCD*), and the ATP-binding protein AtWA (component A2) required for Mcr activation [23]. The “Ca. Methylarchaeaeae” harbor genes for methane production from methanol and methylamines (*mtaA*, *mtbA*, *mttB*, *mtbB*, and *mtmB*) (Supplementary Table 7 and Fig. 2), suggesting that the “Ca. Methylarchaeaeae” have potential to perform methyl-reducing methanogenesis, as previously suggested for “Ca. M. tengchongensis” [7], and members of the orders *Methanomassiliicoccales* [15], “Ca. Methanofastidiosaeae” [19] and “Ca. Methanomethylaeae” [4]. All of the “Ca. Methylarchaeaeae” MAGs encoded a tetrahydrodromethanopterin (H$_4$MPT) S-methyltransferase subunit H (MtrH), and either a MtrX or MtrA, that are homologous to those of *Methanosarcina Barkeri* (Supplementary Table 7). Phylogenetic analysis revealed that the “Ca. Methylarchaeaeae” MtrH subunits are more closely related to a MtrH (BP07_RS03240) of *Methemococcus shengiensis* than to the MtrH subunits of *Methanosarcina* (Supplementary Fig. 6). It is likely that the “Ca. Methylarchaeaeae” MtrH may be involved in methyl transfer directly to H$_4$MPT, as previously shown in *M. shengiensis* for utilization of methoxylated aromatic compounds [24]. The absence of a complete gene operon for Mtr complex suggests that the “Ca. Methylarchaeaeae” cannot use the CO$_2$ reduction or aceticlastic pathway for methanogenesis.

In contrast to the “Ca. Methanomethylicaeae”, all genes for the Wood-Ljundahl pathway (WLP) and acetyl-CoA decarboxylase/ synthase: CO dehydrogenases (ACDS/CODH) are also present in all the genomes (Supplementary Table 7 and Fig. 2). However, we did not identify the energy-converting hydrogenase complex and F$_{242}$CO reducing hydrogenase complex, both of which are required for the oxidation of the methyl groups to CO$_2$ via the WLP [12]. This suggests that the “Ca. Methylarchaeaeae” cannot utilize the methylotrophic pathway for methanogenesis. Similar to methyl-reducing methanogens of the *Methanopteronarchaeaeae* [17], function of the defective WLP remains a mystery. The “Ca. Methylarchaeaeae” MAGs contain one or two copies of a gene encoding heterodisulfide reductase subunit D (Hd2D) (Supplementary Fig. 7 and Supplementary Table 7), one of which was located with a b-type cytochrome gene (Fig. 3a and Supplementary Fig. 7), which is similar to the *hdrDE* operon of *Methanosarcina Barkeri* [25]. The b-type cytochromes in the Hd2D-like complex of the “Ca. Methylarchaeaeae” are integral membrane proteins with five transmembrane helical segments that harbor a nitrate reductase gamma subunit domain (PF02665) (Fig. 3c and Supplementary Figs. 7 and 8). Sequence analysis of these b-type cytochromes revealed two histidine residues located in Helix 2 of these proteins in all the “Ca. Methanarchaeaeae” genomes, two histidine residues located in Helix 5 for H03B1, and single histidine and methionine residues located in Helix 5 for “Ca. Methanotowutia” and “Ca. Methanoinusilarius” (Supplementary Fig. 7b and Fig. 3c). These residues are suggested to be involved in the binding of two heme groups [26], similar to the NarI of *E. coli* [27] and HdrE of *M. Barkeri* [25]. It is assumed that the two heme groups ligated to histidine or methionine residues of Helix 1 and Helix 5 are on the periplasmic and cytoplasmic side of the membrane bilayer respectively, and are responsible for electron transfer. In addition, the *hdrDE* operon is adjacent to the *mcrABDG* operon in all the “Ca. Methylarchaeaeae” MAGs (Fig. 3a), suggesting their role in methanogenesis for these microorganisms. Collectively, these findings strongly indicate that members of the “Ca. Methylarchaeaeae” are b-type cytochrome-containing methanogens that use the Hd2D complex to reduce the heterodisulfide CoM-S-S-CoB of Coenzymes M and B generated in the final step of methanogenesis [28] (Fig. 2).

We identified a homolog of a 11-subunit NADH-quinone oxidoreductase complex in each “Ca. Methylarchaeaeae” genome (Supplementary Table 7) whose gene cluster resembles to the
F$_420$H$_2$ dehydrogenase (Fpo) found in Methanosarcina [29] (Supplementary Fig. 9b). Phylogenetic analysis of the large subunit revealed that the "Ca. Methylarchaeales" complex is more closely related to the Fpo and Fpo-like complexes of Methanosarcinales and Methanomassiliicoccales than to group 4 [NiFe] hydrogenases (Supplementary Fig. 10). The absence of the typical [NiFe]-binding motifs in the catalytic subunit excludes the possibility that the complex is a group 4 [NiFe] hydrogenase (Supplementary Fig. 9a). In addition, the complex also lack the FpoF subunit required for binding and oxidation of F$_{420}$H$_2$ [15]. This suggests that this Fpo-like complex is unable to interact with F$_{420}$H$_2$ and instead may use reduced ferredoxin as an electron donor, similar to its proposed role for the Methanomassiliicoccales [15] and Methanosaeta thermophila [30]. In six MAGs from "Ca. Methanoinulosilans", "Ca. Methanoporticola", and "Ca. M. tengchongensis", genes for soluble methyl viologen-reducing hydrogenase/heterodisulfide reductase complex (MvhADG/HdrABC) and methanophenazine-reducing hydrogenase complex (VhtAGC) are missing. It is extremely unlikely that genes encoding all MvhADG/HdrABC and VhtAGC complex subunits are present in these near-complete genomes but were missed by sequencing. Thus, it is proposed that these microorganisms may use the Fpo-like complex directly to accept electrons from reduced ferredoxin, and subsequently channel these electrons to the HdrDE complex coupled to the reduction of CoM-S-S-CoB (Fig. 2), as shown previously for Methanosaeta thermophila [30]. The reduced ferredoxin may be produced by some unidentified hydrogenases or an unknown pathway. The H03B1 MAG also encodes a formate dehydrogenase subunit A gene (fdhA) co-located with a fdhB gene (Supplementary Table 7) and a putative b-type cytochrome with five transmembrane helices and a prokaryotic b561 domain (PF01292) binding two heme groups (Supplementary Fig. 11c) that is similar to FdhC of E. coli. "Ca. M. tengchongensis" contained fdhAB genes, with the fdhB gene collocated with a gene for a cytochrome b561 with four transmembrane helices and two heme groups (Supplementary Fig. 11b). It is likely that these microorganisms may be able to use formate dehydrogenase to reduce methanophenazine pool which could then transfer electrons to the membrane-bound HdrDE complex (Fig. 2). We identified a geranylfernsyl diphosphate synthase homolog in each “Ca. Methylarchaeales” genome. Phylogenetic analysis revealed that these enzymes cluster together with the geranylfernsyl diphosphate synthase of M. mazei, likely suggesting that the “Ca. Methylarchaeales” may be able to synthesize methanophenazine, as previously shown in M. mazei [31] (Supplementary Fig. 12).

The “Ca. Methanotowutia” (TDP8 and TDP10) MAGs encode the small and large subunits for a [NiFe] active site-containing hydrogenase co-located with a gene for membrane-spanning b561 domain (PF01292) cytochrome b (Fig. 3b), which is similar to the operon of VhtAGC complex found in Methanosarcina with cytochromes [12]. The b-type cytochrome harbors five transmembrane helices, with histidine or methionine residues located in
Fig. 3 Gene composition and structural model of HdrDE and VhtAGC complexes in the “Ca. Methylarchaeales”. a Gene composition of contigs/scaffolds containing the gene cluster of heterodisulfide reductase (HdrDE) complex. Genes related to methane metabolism are highlighted with red, blue, yellow, and cyan. The hdrDE complex gene cluster is always adjacent to mcrABDG operon. b Gene composition of methanophenazine-reducing hydrogenase (VhtAGC) complex. Genes for VhtAGC were collocated on the same contig/scaffold, forming a transcriptional unit. c Structural model of b-type cytochromes in HdrDE and VhtAGC complexes showing the proposed heme ligation.
Helix 1, 2, 5 for the ligation of two heme groups (Supplementary Fig. 11a). It has been proposed that the VhtA is guided to the cell membrane with the help of twin-arginine signal peptide of VhtG and its [NiFe] active site faces periplasmic side [32, 33]. As a result, two H⁺ ions generated by H₂ oxidation are released into the periplasm while two electrons are transferred to heme groups of VhtC through Fe-S clusters of VhtG [12, 34]. Furthermore, the electron carrier methanophenazine connects VhtAGC with HdrDE, and its reduction and reoxidation results in the release of two additional H⁺ ions into the periplasm (Fig. 2) [34, 35]. Altogether, four energetic protons are generated in the system, which can be used to drive the synthesis of one ATP via an archaeal A-type ATP synthase. The HdrDE complex that receives electrons from the methanophenazine can be used to reduce CoM-S-S-CoB (Fig. 2), enabling the coupling of methane production with energy conservation. This is the first report of a VhtAGC complex and an HdrDE complex found in an mcr-containing archaeal lineage outside the Euryarchaeota superphylum (Fig. 1) and indicates that “Ca. Methanotowutia” may be capable of performing H₂-dependent methyl-reducing methanogenesis. The membrane-bound electron transport chain is more efficient than electron bifurcation that is used by methanogens without cytochromes [12].

Sequence analysis revealed that key conserved residues of the McrA sequences of the “Ca. Methylarchaeales”, including the binding sites for F₄₃₀ cofactors, coenzyme M, and coenzyme B [36], are the same as those in McrA sequences of members of the Euryarchaeota superphylum, with exception that the cysteine at site α452 is replaced with an alanine or serine (Supplementary Fig. 13 and Supplementary Table 8). Phylogenetic trees of concatenated and individual McrABG were reconstructed, showing that the “Ca. Methylarchaeales” encode canonical Mcr complexes that cluster with those of putative methane-metabolizing archaea and are divergent from those of short-chain alkane-oxidizing archaea (Fig. 4 and Supplementary Fig. 14). These results support the view that the “Ca. Methylarchaeales” metabolize methane.

We also explored the possibility that the “Ca. Methylarchaeales” may be able to oxidize methane. In reported anaerobic methanotrophic archaea (ANME), methane oxidation is coupled to the reduction of several electron acceptors (nitrate, sulfate or metal oxides). Known ANME are predicted to utilize canonical

---

**Fig. 4** Phylogeny of the Mcr/Mcr-like complex showing the relationship with their species tree. a Maximum-likelihood tree (IQ-TREE, LG + C60 + F + G) based on an alignment of concatenated McrABG/McrABG-like subunits from 167 archaeal genomes. The Mcr-like complex is found in short-chain alkane-oxidizing archaea. b Maximum-likelihood tree (IQTREE, LG + C60 + F + G) based on concatenated 122 archaeal-specific marker proteins using the same genomes with those of Mcr/Mcr-like tree. Ultrafast bootstraps values ≥95 are indicated with green filled squares.
terminal respiratory reductases or multi-heme c-type cytochromes (MHCs) to transfer electrons to a syntrophic partner microorganism [37], metal oxides [38, 39] or humics [40]. We could not identify any terminal reductases or MHCs in the “Ca. Methylothermaceae” genomes. Previous studies have hypothesized that formate or acetate might act as potential syntrophic electron carriers between methane-oxidizing archaea and their partners [41, 42], and members of the “Ca. Methylothermaceae” possesses the genetic potential for the production of these electron carriers. However, to our knowledge, these electron-transferring mechanisms have never been experimentally verified for ANME. Collectively, these analyses suggest that these “Ca. Methylothermaceae” are more likely methanogens, although empirical studies are required to confirm this.

Similar to all described methanogens [15], the “Ca. Methylothermaceae” do not encode a complete tricarboxylic acid cycle, with citrate synthase, fumarase and succinate dehydrogenase absent from these MAGs. The “Ca. Methylothermaceae” lack a canonical pyruvate kinase for glycolysis (Supplementary Fig. 15 and Supplementary Table 7). However, pyruvate-water dikinase or pyruvate phosphate dikinase in glucoseogenesis may replace pyruvate kinase to catalyze the reversible interconversion of phosphoenolpyruvate and pyruvate, as shown in cultivated methanogens Methanomassiliicoccales [15]. The identification of sugar transport proteins and a variety of extracellular and intracellular carbohydrate-active enzymes (CAZymes) including glycoside hydrolases (EC 3.2.1.1 and 5.4.99.16) and glycosyltransferases (EC 2.4.1, 2.4.1.83, and 2.4.99.18, etc.) in the “Ca. Methylothermaceae” (Supplementary Fig. 15) suggests that they may be able to utilize sugars as an alternative carbon and energy source, as previously hypothesized for the “Ca. Methanomethyli- cales” and “Ca. Batharchaeia” [4, 8]. However, comparative genomics revealed that cultured methanogens that do not utilize sugars also encode similar proteins (Supplementary Fig. 15) [12, 13], and they may instead be involved in biosynthetic pathways. In addition, peptide and amino acid transporters, and enzymes related to peptide fermentation including extracellular peptides, endopeptidases, 2-oxoglutarate ferredoxin oxidoreductase (kor), 2-ketoisovalerate ferredoxin oxidoreductase (vor), indolepyruvate ferredoxin oxidoreductase (ior), and pyruvate ferredoxin oxidoreductase (por) are present in both the “Ca. Methylothermaceae” and cultured methanogens (Supplementary Fig. 15). Nevertheless, to our best knowledge, peptide fermentation has never been reported in these isolated methanogens to date. Thus, the genes may be involved in assimilation and metabolism of amino acids in the “Ca. Methylothermaceae” and other newly discovered uncultured methanogens [4, 8, 12].

**Evolution of the b-type cytochrome-containing methanogens**

The rapid increase in the number and diversity of MAGs has greatly expanded the known diversity and distribution of Mcr genes in archaea. To investigate the evolutionary history of the Mcr complexes in methanogens, we inferred the phylogeny of concatenated McrABG subunits based on all mcr-containing archaeal genomes available in public databases. In accordance with previous studies [43, 44], lineages in Class I and Class II methanogens within the Euryarchaeota superphylum appear congruent between McrABG and species trees while H2-dependent methylothermophilic methanogens Methanomassiliicoccales and Methanobacteriaceae, and methanotroph “Ca. Methanoflobi- gales” (ANME-1) are not (Fig. 4). The results were further supported by the phylogeny of the six conserved markers (m4–m9) in this (Supplementary Fig. 16) and previous studies [44]. These markers are solely present in archaea containing Mcr or Mcr-like complexes and suggested to be involved in activation, folding and assembly of Mcr subunits [44]. The Mcr genes of “Ca. Methanomethylicales” and “Ca. Korarchaeia” within the phylum Thermoproteota were previously suggested to be acquired via HGTs, since they are closely related with those of methylothetic methanogens of the Euryarchaeota superphylum in McrABG tree [44]. However, analyses including our “Ca. Methylothermaceae” MAGs and several others with an Mcr complex revealed good congruence between the concatenated McrABG, m4-m9 genes, and the genome-based trees for the lineages within the Thermoproteota (including the “Ca. Methanomethylicales”, “Ca. Korarchaeia”, “Ca. Nezhaarchaeales”, and our “Ca. Methylothermaceae”; Fig. 4 and Supplementary Fig. 16) suggest vertical inheritance and evolution independent of the Euryarchaeota superphylum. Wide distribution of mcr genes in archaea (Supplementary Fig. 17 and Supplementary Table 9) and their congruence with the genome-based tree for many lineages within the Euryarchaeota superphylum and the Thermoproteota suggest that these genes likely have originated before the divergence of these two major archaeal lineages.

Recently, amalgamated likelihood estimation (ALE) has been used to estimate presence probability of McrA in each internal node in a rooted archaeal species tree, supporting the presence of McrA with high confidence in the common ancestor of Class I and Class II methanogens, “Ca. Methanofastidiosi-diales”/“Ca. Niuwarchaeales” in Euryarchaeota superphylum, as well as “Ca. Methano- methylcalicales”, “Ca. Korarchaeia”, and “Ca. Nezhaarchaeales” in the Thermoproteota [45]. Compared to the previous study [45], our ALE results support the presence of McrA with high confidence [presence probability (pp) >0.9] at the basal node of “Ca. Methanomethylicales”, “Ca. Nezhaarchaeales”, “Ca. Korarchaeia”, and “Ca. Methylothermaceales” in the Thermoproteota (Supplementary Fig. 17), suggesting an earlier origin of Mrc complex in Thermoproteota. The difference is likely attributed to the addition of “Ca. Methylothermaceae”. Confidence in evolutionary inferences from ALE analyses will require expansion of genome coverage of some of the poorly represented or yet-to-be-discovered Mcr-containing lineages. A previous study showed that an ancestral McrA sequence were more closely related to McrA from “Ca. Methanodesulfofokores washburnensis” in the “Ca. Korarchaeia” compared to any other lineages [6], possibly supporting our inference that methane metabolism may have evolved relatively early in Thermoproteota.

The b-type cytochrome in HdrDE complex belongs to the protein family of nitrate reductase gamma subunit (PF02265, NarI). Using all publicly available archaeal genomes, we found that the NarI domain-containing cytochromes (NarI-Cyt) are primarily used in three electron transfer complexes: HdrDE, dissimilatory nitrate reductase (NarGHI) [46], and sulfate reductase (DsrABC/KJ/MOP). For the HdrDE and NarGHI complexes, the genes encoding the subunits are co-localized in archaeal genomes, each forming a transcriptional unit. However, in the Dsr complex, only a DsrK is co-localized with a DsrM (b-type cytochrome) while other subunits are usually not adjacent to the DsrKM but separated by few genes [6]. We examined distribution of the three complexes in archaea. A total of 101 genomes were found to encode these complexes (66 for HdrDE, 16 for Nar, 23 for Dsr), and they are distributed across the Euryarchaeota superphylum, Thermoproteota, and Asgardarchaeota (Supplementary Fig. 17 and Supplementary Table 9). Among these genomes, the HdrDE is found in methanogens and methanotrophs belonging to the class “Ca. Methanosarcinicia”, the orders Methanococcales and Methanobacterales, and in alkane-oxidizing archaea belonging to the orders Archaeoglobales, “Ca. Syntropharchaeales”, and Methanosarciniales (GoM-Arc1) (Supplementary Fig. 17). In Mcr-containing archaea outside of the Euryarchaeota superphylum, the complex is exclusively found in the “Ca. Methylothermaceae” (Fig. 1 and Supplementary Fig. 17).

Phylogenetic analyses of the NarI-Cyt were conducted to investigate the evolution of these genes in archaea (Fig. 5a). The
results showed that these cytochromes have experienced frequent horizontal gene transfer, especially DsrM. The DsrM sequences annotated in members of the Thermoproteota form a distinct cluster. In the cluster, Archaeoglobi and “Ca. Hydrothermarchaeota” DsrM branch far from their Euryarchaeota superphylum relatives, and have potentially gained their cytochromes from a member of the “Ca. Korarcheia”. Similarly, the “Ca. Methanoperedenaceae” and Archaeoglobi might have
acquired their NarI genes from a member of *Thermoproteia*. Congruence between the cytochrome and genome-based trees for members of the *Thermoproteota* suggest that these cytochromes might have evolved before the diversification of this phylum. We further inferred a gene tree using concatenated HdrDE complex (Fig. 5b). The topological structure of this tree exhibits high congruence with the genome-based tree for all lineages except the *Methanoratronarchaeia*, supporting an early presence of the complex in archaea. This suggestion is supported by ALE analyses which indicate the presence of NarI-Cyt with high confidence in the common ancestor of *Thermoproteota* (pp = 0.69) and in the common ancestor of “Ca. Halobacteriota” (pp = 0.70) (Supplementary Fig. 17).

As mentioned above, *b*-type cytochromes are classified into different protein families, and form part of many membrane-bound electron transfer complexes in bioenergetic pathways [47, 48]. Aside from HdrDE, Nar, and Dsr, such complexes also include Vht, Fdh, *b*6f complex, *bc*1 complex, and succinate dehydrogenase (Sdh). We examined the distribution of different families of *b*-type cytochromes in 416 representative archaea covering 41 orders or phyla of the Euryarchaeota superphylum, *Thermoproteota*, and *Asgardarchaeota* (Supplementary Fig. 17 and Supplementary Table 9). A total of 246 genomes contained these *b*-type cytochromes that were distributed across 23 archaean lineages. In total, 11 of the 13 lineages of the *Thermoproteota*, and 11 of the 24 orders in *Euryarchaeota* superphylum, had *b*-type cytochrome, suggesting its pervasiveness in archaea. We conducted phylogenetic analyses of the *b*-type cytochromes from different families (Fig. 6a). The result indicates that cytochromes from Fdh and Sdh complexes form two large clusters. Within each cluster, lineages from *Thermoproteota* or the *Euryarchaeota* superphylum were essentially grouped together, suggesting that these cytochromes may have evolved before the divergence of these major archaean lineages. The cluster of cytochromes of the *b*6f complex is close to those of the *bc*1 complex, consistent with the suggestion that bacterial cytochromes in *bc*1 complex may originate from cytochromes in *b*6f complex [48]. A phylogenetic analysis of concatenated VhtAGC showed clustering of lineages from *Thermoproteota* with *Archaeoglobi* (Fig. 6b), suggesting ancient exchanges of the Vht complex among these lineages. Taken together, these results support an early origin of *b*-type cytochromes in archaea. Previous studies also imply that some core enzymes for bioenergetic pathways, including membrane-integral *b*-type cytochrome, formate dehydrogenase, [NiFe]-hydrogenase, the Rieske/cytb complexes, and NO-reductases, were present in the Last Universal Common Ancestor of Bacteria and Archaea [48, 49].

As the heme is indispensable to *b*-type cytochrome [47], we also investigated distribution of its biosynthetic pathway in archaea. Although there are 11 genes involving in the heme biosynthesis,
The three genes (Ahh-NirDH, Ahb-NirJ1, and Ahb-NirJ2), responsible for conversion from precorrin-2 to heme, are the key to this pathway. Thus, these three genes were used as markers denoting the presence of heme biosynthetic pathway. Among 41 archaeal lineages, 32 had this pathway including the "Ca. Methylocarchaeales" (Supplemental text, Fig. 2, Supplementary Fig. 17 and Supplementary Table 9). Phylogenetic analyses reveal that these lineages from Thermoproteota largely cluster together for Ahb-NirDH (Supplementary Fig. 18). However, for Ahb-NirJ1 and Ahb-NirJ2, lineages from the Euryarchaeota superphylum, the Thermoproteota, and Asgardarchaeota are tangled up, suggesting frequent HGTs of these genes between these lineages. The wide distribution of this pathway across the Euryarchaeota superphylum, the Thermoproteota, and Asgardarchaeota (Supplementary Fig. 17 and Supplementary Table 9) suggests that a common ancestor may have been able to synthesize heme. This observation further supports the possibility of the early presence of b-type cytochromes in archaea.

Here we described the discovery of the novel archaean order "Ca. Methylocarchaeales", expanding known methanogen and archaean diversity. Members of the lineage are methyl-reducing methanogens that can conserve energy via membrane-bound electron transport chains. The "Ca. Methylocarchaeales" are globally distributed in anoxic lake and marine sediments, suggesting that they make an important contribution to global methane emissions. Our broader analyses suggest that methanogens who use b-type cytochrome-containing complexes to transfer electrons may have originated before diversification of Thermoproteota or "Ca. Halobacteriota" phyla based on a conservative estimation for the origin of McrA and Narl-Cyt genes in the ALE analysis. A previous study using molecular clock analyses to indicate that the origin of McrA and NarI-Cyt genes in the ALE analysis. A previous study using molecular clock analyses to indicate that the diversification of Thermoproteota likely occurred in the early Archean Eon [45]. Archean oceans are thought to have been anoxic and contain abundant ferrous iron from hydrothermal volcanics [50, 51], which would have provided sufficient raw materials for heme synthesis by methanogens. In addition, CO2, H2, and organic compounds produced by volcanic activity are transported to the early oceans [32], which provides adequate carbon and energy sources for methanogenic growth. Compared to hydrogenotrophic methanogens using electron bifurcation, methanogens using the membrane-bound electron chain have a higher energy production efficiency and growth yield, providing an advantage for members of the "Ca. Methylocarchaeales" described here.

Taxonomic proposals
"Ca. Methanowutia igneequa" (gen. nov., sp. nov.). Methanowutia (Me.tha.no.to.wu.ti.a. N.L. pref. methano-, pertaining to methane; N.L. fem. n. Methanowutia methanogenic organism named after the lake Towuti in Indonesia where members of the genus were first discovered).

Methanowutia igneequa (ig.ne.a'qua. L. masc. adj. igneus, of fire; L. fem. n. aqua, freshwater, pertaining to freshwater habitats; N.L. gen. n. igneequa from/of water of fire, referring to the volcanic lake environment). This organism is deduced to be able to use methylated compounds for methanogenesis. Representative genomes are near-complete bins TDP8 (Accession No. SAMN15658089) and TDP10 (Accession No. SAMN15658091) recovered from freshwater sediments in Lake Towuti in Indonesia where members of the latter type the genome for the species.

"Ca. Methanoinsularis halodrymii" (gen. nov., sp. nov.). Methanoinsularis (Me.tha.no.in.sul.ar.is. N.L. pref. methano-, pertaining to methane; L. fem. adj. insularis, from an island; N.L. fem. n. Methanoinsularis methanogenic organism from an island, specifically referring to Techeng Island in China where these microorganisms were discovered).

Methanoinsularis halodrymii (ha.lo.dry.mi. Gr. masc. n. halos (gen. halos) salt; Gr. masc. n. drymos coppice; N.L. gen. n. halodrymii of salty woodland, referring to the mangrove wetland environment). This uncultivated microorganism is assumed to be able to perform methylotrophic methanogenesis. The type genome for the species is the bin H0381 (Accession No. SAMN15658086) recovered from mangrove wetlands in Techeng Island in China.

"Ca. Methanoinsularis haikouensis" (gen. nov., sp. nov.). Methanoinsularis haikouensis (hai.kou.en'sis. N.L. fem. adj. haikouensis, pertaining to Haikou). This uncultivated microorganism is assumed to be able to perform methylotrophic methanogenesis.

Representative genomes are the bins HK01M, HK01B, HK02M1 (Accession No. SAMN25131447, SAMN25131448, SAMN25131449) recovered from mangrove wetlands in Dongzhai Harbour in Haikou, China.

"Ca. Methanoporticola haikouensis" (gen. nov., sp. nov.). Methanoporticola (Me.tha.no.porti.co.la. N.L. pref. methano-, pertaining to methane; L. masc. n. portus, harbour; L. suff. -cola (from L. masc. or fem. n. incola), inhabitant, dweller; N.L. masc. n. Methanoporticola, a methane-forming dweller of a harbor, specifically referring to Dongzha harbour in China where these microorganisms were discovered).

Methanoporticola haikouensis (hai.kou.en'sis. N.L. masc. adj. haikouensis, pertaining to Haikou). This uncultivated microorganism is assumed to be able to perform methylotrophic methanogenesis. The type genome for the species is the bin HK02M2 (Accession No. SAMN25131450) recovered from mangrove wetlands in Dongzhai Harbour in Haikou, China.

"Ca. Methylocarchaeales" (ord. nov.). Methylocarchaeales (Me.thyl.ar.chae.a.les. N.L. neut. n. Methylocarchaeum (Candidatus) type genus of the order; -ales, ending denoting an order; N.L. fem. pl. n. Methylocarchaeales, the order of the genus "Ca. Methylocarchaeum"); Methylocarchaeaceae (Me.thyl.ar.chae.a.ce.ae. N.L. neut. n. Methylocarchaeum (Candidatus) type genus of the family); -aceae, ending denoting a family; N.L. fem. pl. n. Methylocarchaeaceae, the family of the genus "Ca. Methylocarchaeum").

**MATERIALS AND METHODS**

**Sample collection and DNA sequencing**

Thirteen sediment samples were obtained from mangrove wetlands on Techeng Island, Zhanjiang, Guangdong, China on November 25, 2018, and in Dongzha Harbour, Haikou, China on September 30, 2021 (Supplementary Fig. 1). In each wetland, the two to three cores (1 m deep and 2–10 m apart) were taken using a peat sampler (two cores for Techeng Island; three cores for Dongzha Harbour). Each sediment core was evenly divided into three parts in an anoxic glove box. Sediments from subsurface (15–20 cm depth), middle (40–45 cm depth), and bottom (95–100 cm depth) layers were put into plastic bags immediately after collection, kept in a sampling box with dry ice, transported to the laboratory and stored at −80 °C for further analysis. The detailed sampling information is shown in Supplementary Fig. 1.

Genomic DNA was extracted from ~10 g of sediment samples with the PowerSoil DNA Isolation Kit (MoBio Laboratories, Carlsbad, CA, USA). Metagenomic sequencing was conducted on HiSeq 2500 platform (Illumina, San Diego, CA, USA) at Guangdong MagiGene Technology Company (Guangzhou, China). Each sample from Techeng Island wetland generated about 60 Gbp of raw sequence data (2 × 150 bp paired-end reads), while 100 Gbp of sequencing data per sample were obtained for mangrove sediment from Dongzha harbour.

**Genome assembly and binning**

Raw reads generated from mangrove wetland sediments were quality filtered and pruned using Trimmomatic [53]. The resulting clean reads were assembled using MEGAHIT [54] with the following parameters:
concatenated set of 122 archaeal-specific single copy marker genes in the GTDB (https://gtdb.ecogenomic.org/) (Supplementary Table 11). The orthologs of these marker genes in the "Ca. Methylarachaeales" MAGs and the reference genomes were identified using GTDB-Tk tool [20] (v1.3.0, https://github.com/Ecogenomics/GTDBTk) based on hidden Markov models. Maximum-likelihood trees were constructed with IQ-TREE [66] using the following command: "-m LG -G+F +G -bb 1000". The trees were edited using ITOL [67], using the DPA NN superphylogeny as an outgroup, and modified using Adobe Illustrator.
Phylogenies of MtrH, group 4 [NiFe] hydrogenases and geranylgeranylsyphosphate synthase. For MtrH, homologs from the "Ca. Methylophilaceae" were identified by searching against arCOGs and nr database using BLASTP. Reference sequences were derived from a previous study [24]. For group 4 [NiFe] hydrogenases, catalytic subunit of group 4 [NiFe] hydrogenases homologs from the "Ca. Methylophilaceae" were identified using arCOGs, and confirmed with HydDB [77]. Reference sequences were downloaded from HydDB. For geranylgeranylsyphosphate synthase, homologs from the "Ca. Methylophilaceae" were annotated with arCOGs while reference sequences refer to a previous study [31]. Sequences were aligned using MAFFT (-auto) and trimmed with BMGE (-m BLOSUM30 -b 3 -g 0.9). IQ-TREE (-m TEST, -bb 1000) was used to infer these trees.

Phylogenies of the key genes for the heme biosynthesis pathway (Ahb-NirDH, Ahb-Nir1 and Ahb-Nir2). These genes from the "Ca. Methylophilaceae" and 408 representative archaeal genomes were identified using eggNOG-mapper. Hits were confirmed by searching against arCOGs and nr databases using BLASTP. Sequences were aligned using MAFFT (-auto) and trimmed with trimal (automated). Maximum-likelihood trees were constructed with IQ-TREE (-m TEST, -bb 1000).

Gene tree-species tree reconciliation. The ALE analyses were performed using the ALEml_undated algorithm of the ALE package [78] (v1.0 https://github.com/ssolo/ALE). A sample of 1000 and 10,000 trees that were produced in IQ-TREE (-tbs: 1000 for Mcra, 10,000 for Narl-Cy) for each gene family were reconciled with their rooted species trees. The presence probability of gene family as well as duplication, transfer and loss events were estimated in each internal node in the rooted species tree.

DATA AVAILABILITY
Genomes are archived in the NCBI database under BioProject ID PRJNA648665. Genome bins can be found at NCBI under the Accession numbers SAMN15658086 (H0381), SAMN15658087 (H0382), SAMN25131447 (H0301M), SAMN25131448 (HK0181), SAMN25131449 (HK02M1), SAMN25131450 (HK02M2), SAMN15658088 (TDP1), SAMN15658089 (TDP2), SAMN15658090 (TDP9), SAMN15658091 (TDP10). Related raw reads have been submitted to Sequence Read Archive under SRA accession PRJNA629047.

REFERENCES
1. Nisbet E, Dlugokencky E, Manning M, Lowry D, Fisher R, France J, et al. Rising atmospheric methane: 2007–2014 growth and isotopic shift. Glob Biogeochem Cycles. 2016;30:1356–70.
2. Dean JF, Middelburg JJ, Röckmann T, Aerts R, Blauw LG, Egger M, et al. Methane feeding on the global climate system in a warmer world. Rev Geophys. 2018;56:207–50.
3. Rinke C, Chuvchalova M, Mussig AJ, Chaumeil PA, Davín AA, Waite DW, et al. The ALE analyses were performed with the genome taxonomy database. Nat Microbiol. 2021;6:1–14.
4. Vanwamveghem L, Evans PN, Parks DH, Jensen PD, Woodcroft BJ, Hugenholtz P, et al. Methanotrophic and methylotrophic methanogenesis discovered in the archaean phylum Paracrassarchaeota. Nat Microbiol. 2016;1:1–9.
5. Wang Y, Wegener G, Hou J, Wang F, Xiao X. Expanding anaerobic alkaline metabolism in the domain of archaea. Nat Microbiol. 2019;4:595–602.
6. McKay LI, Dlakic M, Fields MW, Delmont TD, Eren AM, Jay ZJ, et al. Archaeal lineage and properties of a methyl monotube containing-containing nitro-sulfonate bacteria. Nat Commun. 2019;10:1–11.
7. Hua ZS, Wang YL, Evans PN, Qu YN, Goh KM, Rao YZ, et al. Insights into the ecological roles and evolution of methyl-coenzyme M reductase-containing hot spring archaea. Proc Natl Acad Sci USA 1988;85:60–63.
8. Vanpauel M, Thiery P, Heim S, Thauer RK. Heterodisulfide reductase from methanol-grown cells of Methanosaeta Barkeri is not a flavoenzyme. Eur J Biochem. 1997;244:326–34.
9. Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W, Siefert KH, et al. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. Nat Rev Microbiol. 2014;12:635–45.
10. Hoekstra W, Groenwold RM, de Vos WM, van der Oost J. Methane metabolism in the archaeal phylum Bathyarchaeota revealed by genomic and metagenomic sequences. Nucleic Acids Res. 2014;42:e73.
2386

40. Bai YN, Wang XN, Wu J, Lu YZ, Fu L, Zhang F, et al. Humic substances as electron acceptors for anaerobic oxidation of methane driven by ANME-2d. Water Res. 2019;164:114935.

41. Meyerderks A, Kube M, Kostadinov I, Teeling H, Glöckner FO, Reinhardt R, et al. Metagenome and mrRNA expression analyses of methanotrophic archaea of the ANME-1 group. Environ Microbiol. 2010;12:422–39.

42. Yang S, Li Y, Liu X, Wang Y, Fan Q, Yang Z, et al. Generic and enzymatic evidence of acetogenesis by anaerobic methanotrophic archaea. Nat Commun. 2020;11:1–11.

43. Borrel G, Adam PS, Gribaldo S. Methanogenesis and the Wood–Ljungdahl pathway: an ancient, versatile, and fragile association. Genome Biol. Evol. 2016;8:1706–11.

44. Borrel G, Adam PS, McKay LJ, Chen LX, Sierra García IN, Sieber CM, et al. Wide diversity of methane and short-chain alkane metabolisms in uncultured archaea. Nat Microbiol. 2019;4:603–13.

45. Wang Y, Wagenegel G, Williams TA, Xie R, Hou J, Tian C, et al. A methylotrophic origin of methanogenesis and early divergence of archaeal multicore alkane metabolism. Sci Adv. 2021;7:eabj1453.

46. Moreno Vivían C, Cabello P, Martinez Luque M, Blasco R, Castillo F. Prokaryotic nitrate reduction: molecular properties and functional distinction among bacterial nitrate reductases. J Bacteriol. 1999;181:6573–84.

47. Jagow a G, Walter S. b-type cytochromes. Annu Rev Biochem. 1980;49:281–314.

48. Dibrova DV, Cherepanov DA, Galperin MY, Skulachev VP, Mildjianian AY. Evolution of cytochrome bc complexes: from membrane-anchored dehydrogenases of ancient bacteria to triggers of apoptosis in vertebrates. Biochem Biophys Acta Biol. 2011;1817:1397–37.

49. Baymann F, Lebrun E, Brugna M, Schoepf Cothenet B, Giudici Orticini MT, Nitschke W. The redox protein construction kit: pre-last universal common ancestor evolution of energy-conserving enzymes. Philos Trans R Soc B. 2003;358:267–74.

50. Wiechert UH. Earth’s early atmosphere. Science. 2002;298:2341–2.

51. Canfield DE. The early history of atmospheric oxygen: homage to Robert M. Garrels. Ann Rev Earth Planet Sci. 2005;33:1–36.

52. Canfield DE, Rosing MT, Bjerrum C. Early anaerobic metabolisms. Philos Trans R Soc B. 2006;361:1819–36.

53. Bolger AM, Lohse M, Usadell B. Trimomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20.

54. Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics. 2015;31:1674–6.

55. Peng Y, Leung HC, Yiu SM, Chin FY. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. Bioinformatics. 2015;31:1674–6.

56. Diekert V, Gardiner K, Arning L, Chepelev I, Corne D, Dessimoz C, et al. CoreFinder: a new software for selection of phylogenetic informative regions from multiple sequence alignments. BMC Evol Biol. 2010;10:210.

57. Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2013;30:772–80.

58. Cresci A, Gribaldo S, BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. BMC Evol Biol. 2010;10:210.

59. Lu Y, Schmidt HA, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32:2676–64.

60. Letunic I, Bork P. Interactive tree of life (iTOl) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. 2016;44:242–5.

61. Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinf. 2010;11:119.

62. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997;25:3389–402.

63. Jones P, Brins D, Chang HY, Fraser M, Li W, McAnulla C, et al. InterProScan 5: genome-scale protein function classification. Bioinformatics. 2013;30:1236–40.

64. Kanehisa M, Satoh Y, Morishima K. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. J Mol Biol. 2016;428:726–31.

65. Eddy SR. Accelerated profile HMM searches. PLoS Comput Biol. 2011;7:e1002195.

66. Niu Y, Mao X, Yang J, Chen X, Mao F, Xu Y. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucleic Acids Res. 2012;40:445–51.

67. Yu CS, Lin CJ, Hwang JK. Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. Protein Sci. 2004;13:1402–6.

68. Kröninger L, Berger S, Welte C, Deppenmeier U. Evidence for the involvement of two heterodisulfide reductases in the energy-conserving system of Methano- masellicoccus luminyensis. FEBS J. 2016;283:472–83.

69. Capella Gutiérrez S, Silla Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25:192–7.

70. Sandergaard D, Pedersen CN, Greening C. HydroDB: a web tool for hydrogenase classification and analysis. Sci Rep. 2016;6:1–8.

71. Szöllösi GJ, Rosikiewicz W, Boussaou B, Tannier E, Daubin V. Efficient exploration of the space of reconciled gene trees. Syst Biol. 2013;62:901–12.

ACKNOWLEDGEMENTS

HPD and LJM was supported by the National Science Foundation for Distinguished Young Scholars (41725002), the National Natural Science Foundation of China (41971125 and 42030401) and the Chinese National Key Programs for Fundamental Research and Development (Nos. 2016YFA0600904, and 2016YFE013370). GWT and SJM are supported by Australian Research Council (ARC) Future Fellowships (FT170100070 and FT190100211, respectively). JK and AV were supported by the ICDP priority program of the DFG (Project Numbers 258261192, 270921149). AV was also supported through an NSF grant (P3EG2.148621). GFZ Potsdam provided additional support through an expedition grant. The authors would like to thank Maria Chuvchina and Aharon Oren for her help with the naming etymology.

AUTHOR CONTRIBUTIONS

HPD and LJM conceived the study. YFO and HPD recovered genomes from metagenomes and analyzed these genomic data. YFO, YLZ, PH, and ML performed analyses of phylogenies and environmental distribution of genes. HPD, LJM, SJM, GWT, SAC, JK, AOL, RLS and ZL wrote the manuscript and Supplementary Information. SAC, SJM, and JK led the drilling project collecting sediment cores from Lake Towuti, South Sulawesi, Indonesia. RLS did the work to generate metagenomes from sediment cores of Lake Towuti. AV contributed to collecting the sediment cores from Lake Towuti in the field. All authors provided comments on the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41396-022-01281-0. Correspondence and requests for materials should be addressed to Hong-Po Dong or Li-Jun Hou. Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
