REVIEW ARTICLE

Bathyarchaeota: globally distributed metabolic generalists in anoxic environments

Zhichao Zhou[1,2], Jie Pan[1], Fengping Wang[3,4], Ji-Dong Gu[2] and Meng Li[1,*,†]

[1]Institute for Advanced Study, Shenzhen University, Shenzhen 518060, People's Republic of China, [2]Laboratory of Environmental Microbiology and Toxicology, School of Biological Sciences, The University of Hong Kong, Pokfulam Road, Hong Kong SAR, People's Republic of China, [3]State Key Laboratory of Microbial Metabolism, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, People's Republic of China and [4]State Key Laboratory of Ocean Engineering, Shanghai Jiao Tong University, Shanghai, People's Republic of China

*Corresponding author: Institute for Advanced Study, Shenzhen University, Shenzhen 518060, People's Republic of China. Tel: (+86) 755-26979250; E-mail: limeng848@szu.edu.cn

One sentence summary: This review not only summarizes the recent findings pertaining to the ecological, physiological and genomic aspects of Bathyarchaeota, highlighting the vital role of this phylum in global carbon cycling, but also reveals some new thoughts on the future directions of research.

Editor: Mecky Pohlschroder

†Meng Li, http://orcid.org/0000-0001-8675-0758

ABSTRACT

Bathyarchaeota, formerly known as the Miscellaneous Crenarchaeotal Group, is a phylum of global generalists that are widespread in anoxic sediments, which host relatively high abundance archaeal communities. Until now, 25 subgroups have been identified in the Bathyarchaeota. The distinct bathyarchaeotal subgroups diverged to adapt to marine and freshwater environments. Based on the physiological and genomic evidence, acetyl-coenzyme A-centralized heterotrophic pathways of energy conservation have been proposed to function in Bathyarchaeota; these microbes are able to anaerobically utilize (i) detrital proteins, (ii) polymeric carbohydrates, (iii) fatty acids/aromatic compounds, (iv) methane (or short chain alkane) and methylated compounds, and/or (v) potentially other organic matter. Furthermore, bathyarchaeotal members have wide metabolic capabilities, including acetogenesis, methane metabolism, and dissimilatory nitrogen and sulfur reduction, and they also have potential interactions with anaerobic methane-oxidizing archaea, acetoclastic methanogens and heterotrophic bacteria. These results have not only demonstrated multiple and important ecological functions of this archaeal phylum, but also paved the way for a detailed understanding of the evolution and metabolism of archaea as such. This review summarizes the recent findings pertaining to the ecological, physiological and genomic aspects of Bathyarchaeota, highlighting the vital role of this phylum in global carbon cycling.

Keywords: Bathyarchaeota; distribution pattern; diversity; metabolism; genomics; carbon cycling

INTRODUCTION

The Miscellaneous Crenarchaeotal Group (MCG) archaea were firstly detected from a hot spring (Barns *et al.* 1996) and later proposed with a name in a study surveying 16S rRNA gene se-quences from marine subsurface sediments (Inagaki *et al.* 2003). The group was termed 'miscellaneous' because of its occur-rence in diverse habitats; it is not only abundant in marine sed-iments but is also widely distributed in terrestrial, freshwater,

hot spring, hydrothermal, etc., environments (Kubo *et al.* 2012). It is one of the predominant groups in the marine subsurface archaeal community (Fry *et al.* 2008; Teske and Sørensen 2008; Lloyd *et al.* 2013). Combined with the large amount of carbon deposited in the subseafloor (ca 15×10^{21} g) (Fry *et al.* 2008), the high abundance of MCG archaea in marine sediments (10–100% of total archaeal abundance) (Parkes *et al.* 2005; Biddle *et al.* 2006; Fry *et al.* 2008; Kubo *et al.* 2012; Lloyd *et al.* 2013) and their heterotrophic properties on detrital proteins, acetate, aromatic compounds and/or other organic substrates (Biddle *et al.* 2006; Webster *et al.* 2010; Webster *et al.* 2011; Lloyd *et al.* 2013; Na *et al.* 2015), naturally led to the proposal that this group of archaea may play an important role in global carbon biogeochemical cycling (Kubo *et al.* 2012; Lloyd *et al.* 2013; Fillol *et al.* 2016; He *et al.* 2016). Based on the phylogenetic analysis of concatenated rRNA, ribosome proteins and topomerase IB protein-encoding genes, MCG is phylogenetically distinct from the closely related Aigarchaeota and Thaumarchaeota, and comprises a parallel lineage that has perhaps evolved from a common ancestor (Meng *et al.* 2014). A new phylum name for this group was proposed, i.e. Bathyarchaeota, reflecting its phylogenetic position as deeply branching with Aigarchaeota and Thaumarchaeota, and its prevalence in subsurface sediments (Meng *et al.* 2014). Recent genomic evidence suggests that Bathyarchaeota might potentially be involved in methane metabolism, a property that had only been confirmed to date in the Euryarchaeota domain (Evans *et al.* 2015; Lloyd 2015). This suggests that methane metabolism might have evolved before the divergence of the ancient archaeal lineages of Bathyarchaeota and Euryarchaeota, in agreement with the assumption that methanogenesis might represent one of the earliest metabolic transformations (Battistuzzi, Feijao and Hedges 2004; Ferry and House 2006; Evans *et al.* 2015; Lloyd 2015). More recently, acetogenesis, a metabolic process deemed to be restricted to the domain bacteria, was also suggested to take place in some lineages of Bathyarchaeota (He *et al.* 2016; Lazar *et al.* 2016), expanding the metabolic potential of archaea. These physiological, ecological and evolutionary features place Bathyarchaeota in the spotlight of current microbial ecology studies, encouraging further explorations of their impact on global and local biogeochemical carbon cycling.

GLOBAL DISTRIBUTION AND HIGH DIVERSITY OF BATHYARCHAEOTA

The significance of global distribution

Bathyarchaeota was initially proposed to form a distinct cluster closely related to Aigarchaeota and hyperthermophilic Crenarchaeota; because of their terrestrial origin (Barns *et al.* 1996) (such as freshwater lakes and hot springs), the name Terrestrial MCG was temporarily proposed (Takai *et al.* 2001). Later on, members of Bathyarchaeota were also found to be abundant in deep marine subsurface sediments (Reed *et al.* 2002; Inagaki *et al.* 2003), suggesting that this group of archaea is not restricted to terrestrial environments, and the name has been changed to MCG archaea (Inagaki *et al.* 2003). According to the meta-analysis of archaeal sequences available in the ARB SILVA database (Kubo *et al.* 2012), Bathyarchaeota was further recognized as a group of global generalists dwelling in various environments, including marine sediments, hydrothermal vents, tidal flat and estuary sediments, hypersaline sediments, terrestrial subsurface, biomats, limnic water and sediments, underground aquifers, hot springs, soils, municipal wastewaters, animal digestive tract, etc. (Fig. 1) (for details see Kubo *et al.* 2012).

In addition to the global distribution, expanding prokaryotic community investigations of deep ocean drilling sediments revealed that members of Bathyarchaeota occupy considerable fractions of the archaeal communities (Teske 2006). Considering the total fractions within all horizons from the sediment cores, members of Bathyarchaeota accounted for 92% of the archaeal community in the Peru Margin Site 1229; 48% in the Peru Margin Site 1227; 71% in volcanic ash layers in the Okhotsk Sea; 47.5% in the forearc basin in the Nankai Trough; 20.6% in the accretionary wedge at the Nankai Trough ODP site 1173; and 83.3% in all layers of the Mediterranean Pleistocene sapropel (Coolen *et al.* 2002; Reed *et al.* 2002; Inagaki *et al.* 2003; Newberry *et al.* 2004; Parkes *et al.* 2005; Inagaki *et al.* 2006; Teske 2006). Fry *et al.* (2008) further summarized 47 clone libraries of 16S rRNA genes from the marine subsurface, with Bathyarchaeota accounting for 33% of all archaea. In addition, the catalyzed reporter deposition-fluorescent *in situ* hybridization (CARD-FISH) studies for the detection and quantification of bathyarchaeotal cells suggest that they are abundant in the center and marine invertebrate-inhabited layers in the Haakon Mosby Mud Volcano, and in the marine subsurface sediments in the Equatorial ODP site 1125 and Peru Basin ODP site 1231 (Kubo *et al.* 2012). It has been suggested that Bathyarchaeota is one of the cosmopolitan groups frequently detected in the freshwater and marine sediments (68% of all sediments analyzed), accounting for a large proportion of the sediment microbial communities (average $36 \pm 22\%$) (Fillol *et al.* 2016). Species abundance distribution analysis indicates that Bathyarchaeota is one of the persistent and abundant core lineages of the sediment archaeal communities, showing, to some extent, habitat-specific distribution (Fillol *et al.* 2016).

Along with the widespread distribution of Bathyarchaeota, i.e. their relatively high abundance in the global marine subsurface ecosystem (Kubo *et al.* 2012; Lloyd *et al.* 2013), they are also metabolically active in the subsurface sediments across geological time scales. Bathyarchaeota dominate 16S rRNA clone libraries of transcribed RNA constructed for the Peru Margin ODP site 1229 (Parkes *et al.* 2005; Biddle *et al.* 2006) and the upper 35 m of the subsurface sediments at the Peru Margin ODP site 1227 (Inagaki *et al.* 2006; Sorensen and Teske 2006). Furthermore, both FISH labeling and intact polar lipid quantification suggest the presence of highly abundant and active bathyarchaeotal cells in the Peru offshore subsurface sediments collected during the Ocean Drilling Program Leg 201 (Biddle *et al.* 2006; Lipp *et al.* 2008). Similarly, rRNA slot blot hybridization indicates the existence of functionally active Bathyarchaeota not only in the surface and subsurface sediments from the Nyegga site 272-02, Cascadia Margin, Gulf of Mexico, Hydrate Ridge ODP site 1245 and Janssand (North Sea), but also in the oxic mats in the Arabian Gulf and subsurface White Oak River sediments (Kubo *et al.* 2012). Considering that the marine subseafloor environment is one of the largest reservoirs of the prokaryotic biomass on Earth, with an estimated microbial abundance of 2.9×10^{29} cells and harboring ca 9.1–31.5% of all prokaryotes on Earth (Kallmeyer *et al.* 2012), the predominance and activity of Bathyarchaeota in the marine subsurface sediments indicates that these microbes might play a crucial role in global biogeochemical nutrient cycling.

Bathyarchaeota subgroups and diversity

A detailed knowledge of the phylogenetic structure of the Bathyarchaeota phylum is crucial for the understanding of their
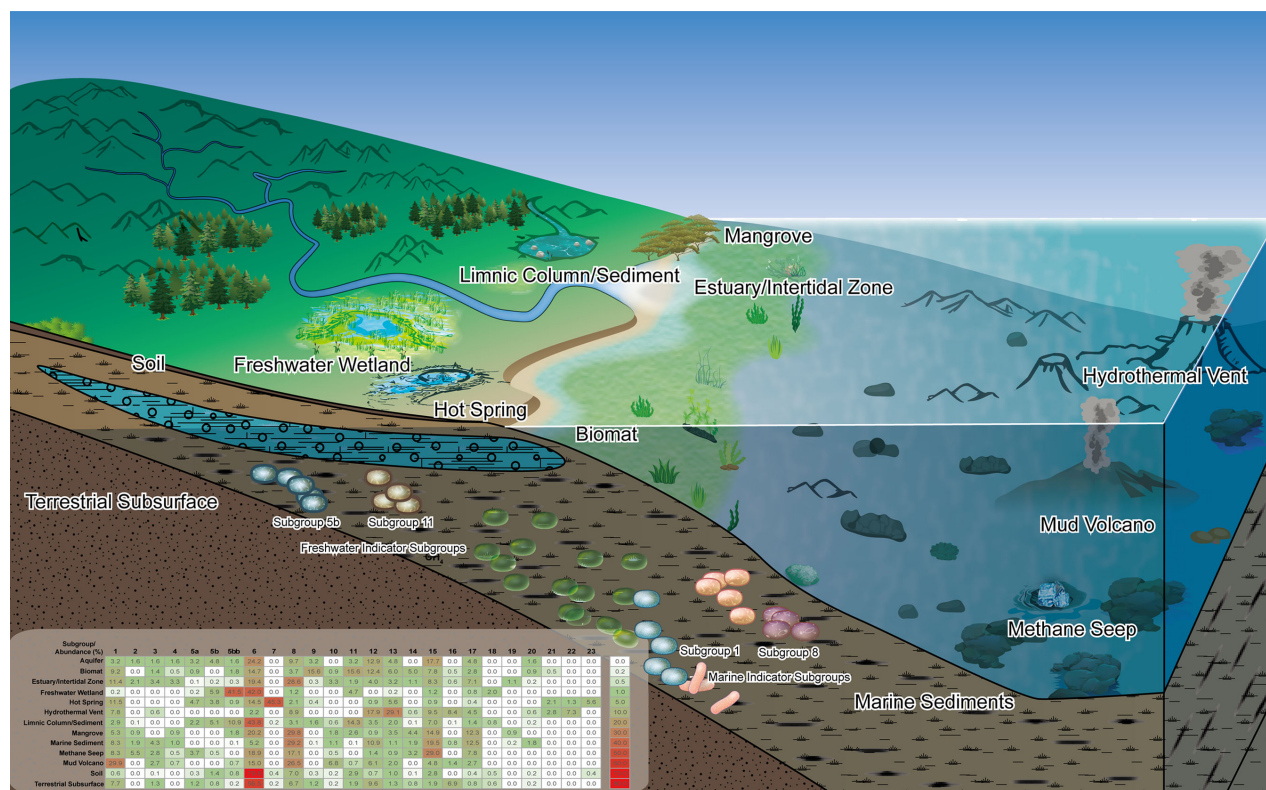
Figure 1. Schematic figure representing major eco-niches of Bathyarchaeota. The indicator subgroups in saline and freshwater sediments were depicted accordingly. The inset table shows the distribution of subgroups in major environmental categories. The isolation source information was parsed from gbk files of bathyarchaeotal 16S rRNA gene sequences. The percentages in every row stand for the proportions of subgroups in each environmental category.

ecological significance in global sedimentary processes. The first comprehensive phylogenetic tree of Bathyarchaeota was constructed in 2012 (Kubo *et al.* 2012); it was based on 4720 bathyarchaeotal sequences from the SILVA database (SSU Ref NR106 and SSU Parc106). Sequences longer than 940 bp were first used to construct the backbone of the tree, and additional sequences were then added without altering the general tree topology. The assignment of bathyarchaeotal subgroups was made based on either having been formerly defined or being monophyletic, using both distance and maximum-likelihood estimations (Kubo *et al.* 2012). In total, 17 subgroups with 76% similarity shared by the most remote sequences were designated; however, 12% of all sequences remained ungrouped. The branching order of Subgroups-13 to -17 was unstable when analyzed by different tree-construction methods, and they were presented as multifurcated branches. Recently, another meta-analysis using newly acquired global sediment bathyarchaeotal sequences resulted in the addition of two more subgroups, Subgroups-18 and -19, with high bootstrap supporting values (96% and 86%, respectively) (Fillol *et al.* 2016). Subgroup-5 is divided into Subgroups-5a and -5b, each with intragroup similarity >90% according to a maximum-likelihood estimation. Subgroup-5b was further split into 5b and 5bb, as additional sequences were added. Bathyarchaeota is characterized by high intragroup diversity, with most subgroups showing within-sequence similarity <92% (Kubo *et al.* 2012; Fillol *et al.* 2016). As suggested by the classification of uncultured archaea based on nearly full-length 16S rRNA gene sequences, the bathyarchaeotal sequence boundary falls into the minimum sequence identity range of phylum level (74.95–79.9%), and each subgroup generally falls into the median sequence identity range of

family and order levels (91.65–92.9% and 88.25–90.1%, respectively) (Yarza *et al.* 2014). It was proposed that the high diversity of Bathyarchaeota implies a high metabolic diversity among its subgroups (Kubo *et al.* 2012).

Because of the high diversity of Bathyarchaeota and various independent analyses of samples from diverse environments, the nomenclature for this archaeal group in previous reports was very complex. For instance, a study into the stratification of the archaeal community from a shallow sediment in the Pearl River Estuary defined bathyarchaeotal subgroups from MCG-A to -F (Jiang *et al.* 2011), including the NT-A3 group, which is predominantly isolated from the hydrate stability zone in the deep subsurface hydrate-bearing marine sediment core in the Nankai Trough (Reed *et al.* 2002); meanwhile, an investigation of archaeal composition in ca 200 m deep sub-seafloor sediment cores at the offshore Peru Margin ODP sites 1228 and 1229 listed Bathyarchaeota subgroups PM-1 to -8 (Webster *et al.* 2006).

To alleviate the nomenclature confusion, we constructed an updated RAxML tree (Fig. 2) based on currently available bathyarchaeotal 16S rRNA gene sequences from SILVA SSU 128 by adding the information from pervious publications (Kubo *et al.* 2012; Lazar *et al.* 2015; Fillol *et al.* 2016; He *et al.* 2016; Xiang *et al.* 2017). In this tree, the Subgroups-1 to -17 were the same as Kubo's tree (Kubo *et al.* 2012), and Subgroup-5 was divided into Subgroups-5a, -5b and -5bb as suggested in Fillol *et al.*'s research (Fillol *et al.* 2016). The subgroups MCG-18, -19 and -20 were firstly named in Lazar *et al.*'s study, but only MCG-19 was represented in the phylogenetic tree (Lazar *et al.* 2015). On the other hand, the subgroups MCG-18 and MCG-19 were also named in Fillol *et al.*'s research (Fillol *et al.* 2016). To avoid the confusion, Subgroups-18 and -19 were named to be consistent with subgroups
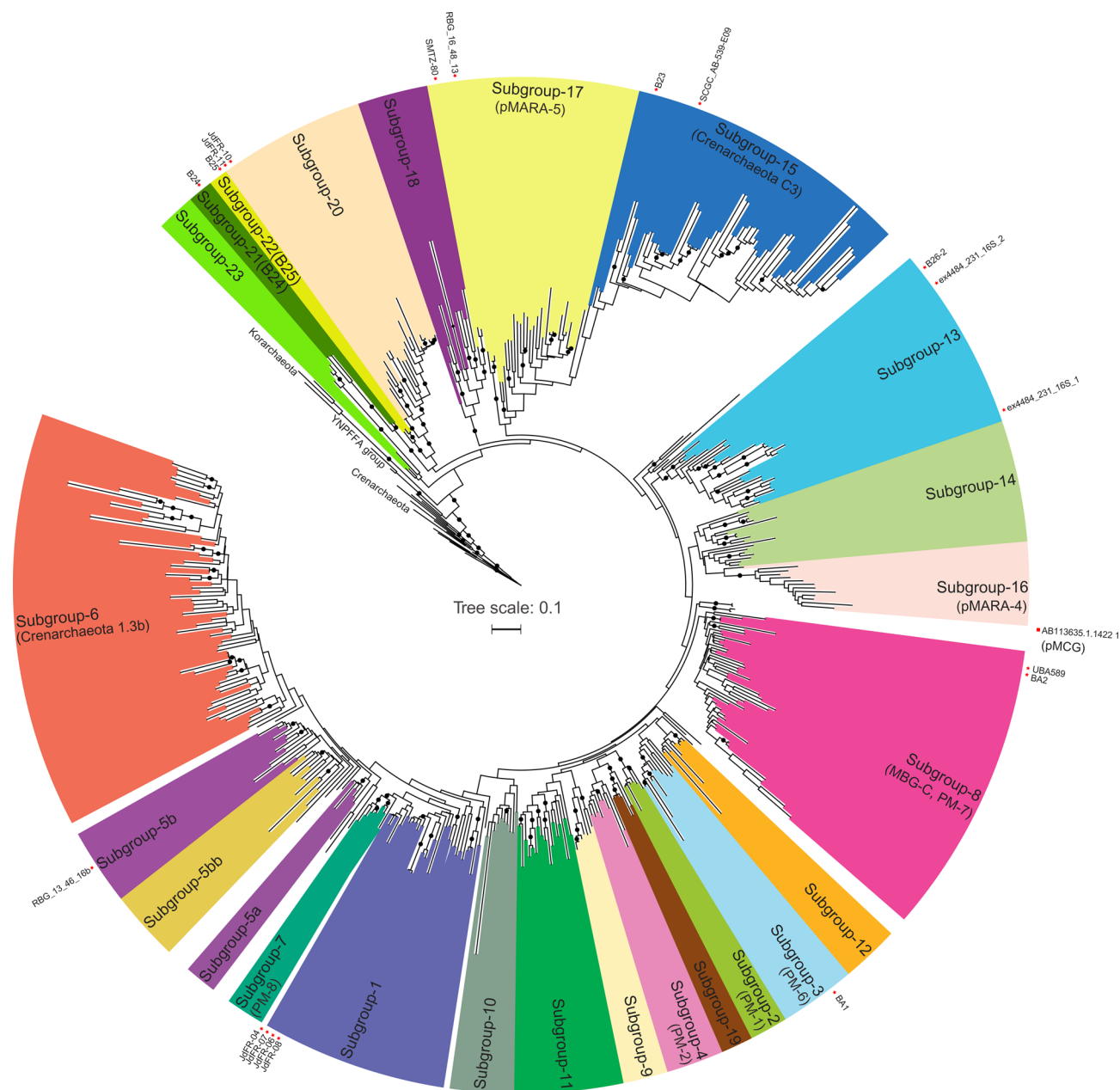
Figure 2. Phylogenetic tree of bathyarchaeotal 16S rRNA genes. Bathyarchaeotal 16S rRNA gene sequences were collected from SILVA SSU database version 128 (sequences of Bathyarchaeota and Group C3; >750 bp) and sequences from pervious publications (Kubo *et al.* 2012; Lazar *et al.* 2015; Fillol *et al.* 2016; He *et al.* 2016; Xiang *et al.* 2017). All sequences were clustered at 90% identity using Usearch v10.0.240 (https://www.drive5.com/usearch/), then the 16S rRNA gene sequences from available bathyarchaeotal genomes in public database, the anchor sequences from Kubo *et al.* (Kubo *et al.* 2012), and the outgroup sequences of Crenarchaeota, YNPFFA group and Korarchaeota were added. All sequences were aligned using SINA v1.2.11 (vision 21 227) with SSU ARB database version 128, and poorly aligned columns (gaps in 50% or more of the sequences) were deleted by using trimAl v1.4.rev15 (Ludwig *et al.* 2004; Capella-Gutiérrez, Silla-Martínez and Gabaldón 2009; Pruesse, Peplies and Gloeckner 2012). Phylogenetic analyses of 16S rRNA gene sequences were inferred by Maximum Likelihood implemented in RAxML 8.0 on the CIPRES Science Gateway using the GTR+GAMMA model and RAxML halted bootstrapping automatically (Miller, Pfeiffer and Schwartz 2010; Stamatakis 2014). Currently available bathyarchaeotal genomes (from GenBank, 29 November 2017 updated) with 16S rRNA gene sequences were labeled in the tree. Peat MCG group was represented with one sequence at 90% cutoff level (Xiang *et al.* 2017). While Subgroups-18 and -19 were named to be consistent with subgroups MCG-18 and MCG-19 as proposed in two previous reports (Lazar *et al.* 2015; Fillol *et al.* 2016), Subgroup-20 was renamed to replace the subgroup MCG-19 in Fillol *et al.*'s tree (Fillol *et al.* 2016). All assigned subgroups have minimum intra-group >90%, and are clustered into one clade with previously reported anchor sequences (Kubo *et al.* 2012). Tree building intermediate files are publicly available (https://github.com/ChaoLab/Bathy16Stree).

MCG-18 and MCG-19 as proposed in two previous reports (respectively Lazar *et al.* 2015; Fillol *et al.* 2016), while Subgroup-20 was renamed to replace the subgroup MCG-19 in Fillol *et al.*'s tree (Fillol *et al.* 2016). The groups of B24 and B25 (He *et al.* 2016) were added into the tree representing Subgroups-21 and -22,

respectively. Furthermore, one new subgroup (Subgroup-23) was proposed in this study (Fig. 2). A group called Peat MCG (pMCG) (Xiang *et al.* 2017) was also listed on the tree; however, because there was only one represented sequence after dereplication at 90% similarity of all bathyarchaeotal 16S rRNA gene sequences,
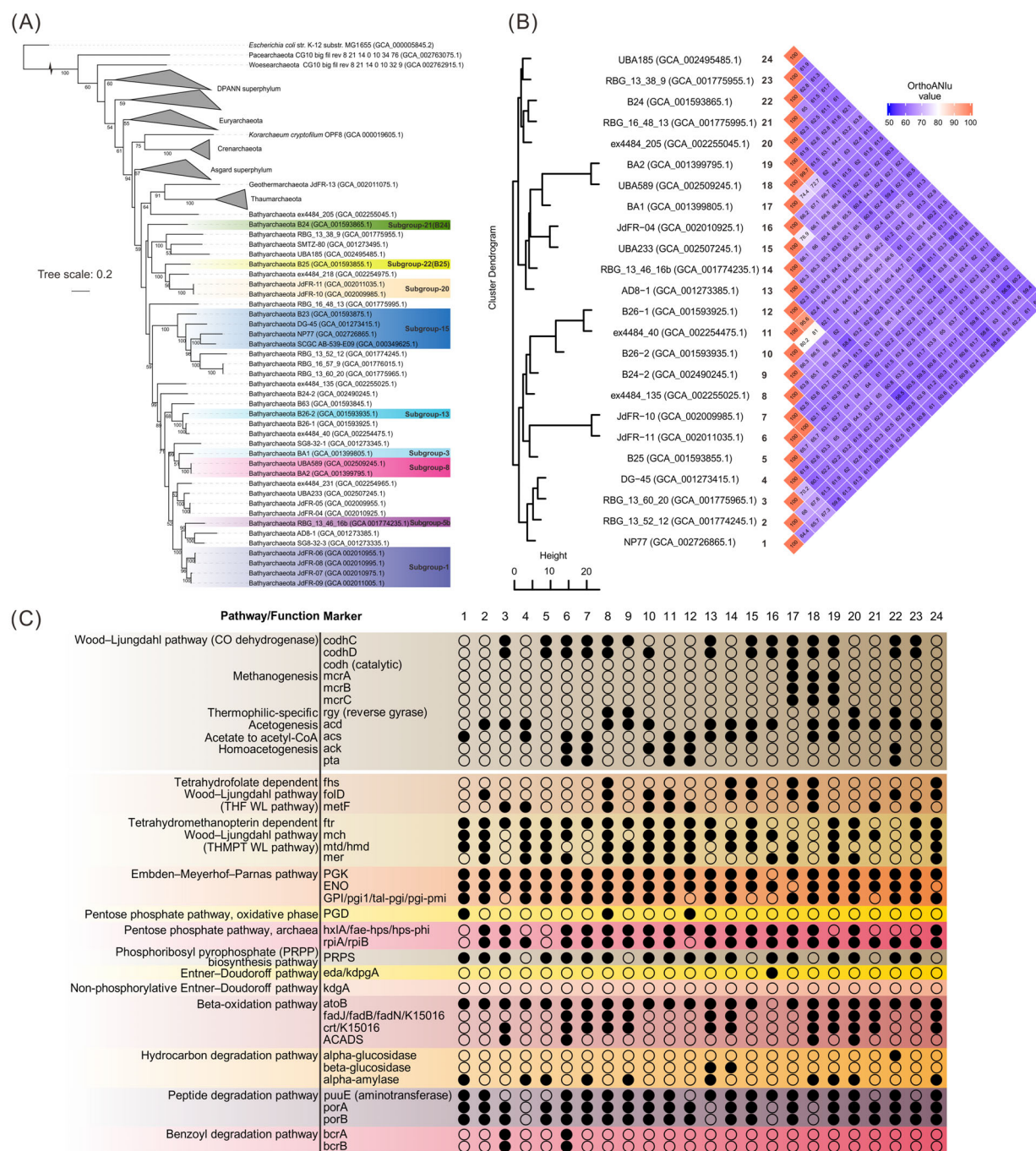
Figure 3. Genomic characterization and metabolic potentials of Bathyarchaeota. (**A**) Phylogenetic tree of ribosomal proteins obtained from currently available bathyarchaeotal genomes (from GenBank, 29 November 2017 updated). The concatenated ribosomal protein (RP) alignment contained 12 RPs, and those genomes with <25% RPs were excluded from tree construction. Subgroups were assigned from the corresponding 16S rRNA gene phylogenic tree (Fig. 2). (**B**) The dendrogram and genome similarity heatmap based on pairwise OrthoANIu values of 24 bathyarchaeotal genomes (Yoon *et al.* 2017). The picked genomes are of high completeness (>70%) and good quality (excluding genomes with numerous long breaking parts with 'N'). (**C**) The metabolic properties of 24 bathyarchaeotal genomes. Markers for individual pathway/function were scanned against genomes using the HMM and KEGG databases (Anantharaman *et al.* 2016; Kanehisa, Sato and Morishima 2016; Spang, Caceres and Ettema 2017). Details of markers refer to Supplementary Table S1 available online. Genome labels are according to panel (B).

we did not list pMCG as a separate subgroup in this tree (Fig. 2). In summary, there are a total of 25 subgroups of Bathyarchaeota based on all available 16S rRNA gene sequences at this moment, and the former names for each subgroup are also labeled in the tree (Fig. 2). Furthermore, the phylogeny of concatenated alignments constituting 12 ribosomal proteins obtained from

currently available bathyarchaeotal genomes (from GenBank, 29 November 2017 updated) was also reconstructed, which showed a similar topology to those of 16S rRNA genes with a few exceptions in Subgroup-17 (Fig. 3A). Thus, this systematic nomenclature based on clear monophyletic or phylogenetically stable subgroups not only facilitates further sequence assignment, but

also provides useful information for understanding the evolutionary separation of specific lineages subjected to natural selection (Fillol *et al.* 2016).

DISTRIBUTION PATTERN AND MOLECULAR DETECTION

Salinity/freshwater separation during diversification

Given the high phylogenetic diversity within the 25 subgroups of Bathyarchaeota, many efforts have been made to understand the key factors that control their distribution and evolution. The phylogenetic species variability index, which reflects the phylogenetic relatedness of sequences originating from specific environments, suggests a non-random distribution of Bathyarchaeota assemblages in natural environments (Fillol *et al.* 2016). A meta-analysis of the distribution of sediment archaeal communities towards environmental eco-factors (7098 archaeal operational taxonomic units from 207 sediment sites worldwide) was performed and a multivariate regression tree was constructed to depict the relationship between archaeal lineages and the environmental origin matrix (Fillol *et al.* 2016). This approach revealed that the separation of subgroups according to saline and anoxic levels could explain 13% of the phylogenetic lineage variance. The first two separation nodes representing the hypersaline, saline and fresh environments accounted for 9.1% of the total phylogenetic lineage variance. Considering the relative abundance of lineages in the separated leaves, Bathyarchaeota accounted for the greatest proportion of lineage variance in the freshwater and saline environments. Further, the IndVal index, which reflects the level of relative abundance and frequency of occurrence, suggests that selective bathyarchaeotal subgroups are bio-indicator lineages in both freshwater and saline environments, as determined by a multivariate regression tree analysis (Fillol *et al.* 2016)

Furthermore, a principal coordinate analysis also clearly separates the bathyarchaeotal community into freshwater and saline sediment groups. This was confirmed by a permutational analysis of variance, with salinity as the best explanatory variable for the variance within the bathyarchaeotal community ($R^2 = 0.04$, $P < 0.001$) (Fillol *et al.* 2016). Eight subgroups were delineated based on the freshwater/saline segregation, as suggested by the significant IndVal values ($P < 0.01$) pointing to freshwater/marine sediment distribution. These indicative subgroups are the dominant ones in the environment, as evaluated by relatively abundant fraction of Bathyarchaeota in corresponding archaeal communities (on average 44% among all studies). Among these are Subgroups-1 and -8 with high IndVal values in marine sediments, and Subgroups-5 and -11 with high IndVal values in fresh sediments (Fillol *et al.* 2016).

The marine/freshwater segregation is a distribution pattern widely shared by diverse microorganisms, including archaea, bacteria, viruses and eukaryotes (Logares *et al.* 2009). Based on the lineage distribution pattern analysis of the archaeal community of seven major eco-niches, it is also evident that the different evolutionary lineages are habitat-specific, and salinity rather than temperature is the primary driving force of the variation of global archaea distribution, with a similar pattern also evident for the global bacterial distribution (Lozupone and Knight 2007; Auguet, Barberan and Casamayor 2010). Ancestral state reconstruction was used to estimate the diversification of bathyarchaeotal lineages previously subjected to the saline/freshwater transition. It has been proposed that the deduced last common ancestor was most likely a saline-adapted

organism, and the evolutionary progression occurred most likely in the saline-to-freshwater direction, with few environmental transitional events. The emergence of freshwater-adapted lineages, including freshwater-indicative Subgroups-5, -7, -9 and -11, occurred after the first saline–freshwater transition event (Fillol *et al.* 2016).

The subgroup distribution pattern

The archaeal community structure, including Bathyarchaeota, is not correlated with a general geochemical categorization, but with the depth and sulfate concentration, subsequently linking to the redox potential, age and the (increasing) degree of organic matter recalcitrance. In the White Oak River estuary, the abundance of Bathyarchaeota decreases with decreasing reductive redox conditions of the sediment (Lazar *et al.* 2015). Subgroup-6 persists in such suboxic, sulfide-depleted shallow sediment layers, while Subgroups-1, -5 and -8 preferentially occur in deeper, more reducing subsurface layers (Lazar *et al.* 2015). Considering the bathyarchaeotal community structure, depth is the first variable responsible for the high degree of absolute subgroup separation, followed by sulfide concentration (reflecting the redox conditions), which is responsible for a low degree of subgroup separation (Lazar *et al.* 2015). A segregated distribution of bathyarchaeotal subgroups was also observed in the water column and sediments in freshwater karstic lakes (Fillol *et al.* 2015). Subgroup-5 thrives in the euxinic bottom water layer, characterized as anoxic and sulfide-rich, with accumulated inorganic and organic reduced compounds; Subgroup-6 is a group of generalists that are adapted to both planktonic and sediment habitats with a wide range of sulfidic conditions. In contrast, Subgroup-15 (Crenarchaeota group C3) organisms dominate cDNA libraries from all sediment layers, albeit with minor contribution to the corresponding DNA libraries; this indicates that this group is metabolically active in the benthic euxinic, organic-rich sediments of karstic lakes (Fillol *et al.* 2015).

Recently, Subgroup-15 was widely detected in both freshwater and marine benthic sediments; its persistent distribution along the sediment depth profile, with higher abundance within active archaeal communities, provides additional hints linking its members' physiological traits to habitat preferences (Liu *et al.* 2014). The members of the Bathyarchaeota are the most abundant archaeal components of the transitional zone between the freshwater and saltwater benthic sediments along the Pearl River, with a central position within the co-occurrence network among other lineages (Liu *et al.* 2014). The members of Bathyarchaeota were positively and strongly correlated especially with the acetoclastic *Methanosaeta*; however, the second most abundant archaeal group, MG-I (subordinate to Thaumarchaeota) is negatively correlated with other groups, probably indicating segregation corresponding to two distinct lifestyles in this case (Liu *et al.* 2014). Furthermore, in contrast to the consistent vertical distribution of all archaeal lineages in freshwater sediments with almost no abundance changes, the total abundance of all Bathyarchaeota and the fraction of Subgroup-15 increase along with the depths of sediments, with significantly high abundance within the archaeal community (Liu *et al.* 2014). In surface and shallow subsurface sediments (surficial to 10 cm deep) of an intertidal mudflat of Brouage in the Bay of Marennes-Oléron, however, the abundances of Subgroup-15 and other bathyarchaeotal subgroups are stable, while the total abundance of Euryarchaeota sequences increases in the same depth range (Hélène *et al.* 2015).

On the other hand, the proportion of bathyarchaeotal sequence in the total archaeal community sequence increases with depth, and they may favor anoxic benthic sediments with iron-reducing conditions. For example, Bathyarchaeota dominates the archaeal community within Louisiana continental shelf (LCS) surface sediment, in both hypoxic and oxic covering water conditions in two distinct seasons (Devereux *et al.* 2015). The major bathyarchaeotal community comprises Subgroups-1, -8, -12 and -15, and is relatively stable during the hypoxic/oxic change, thus being independent of the sedimentary chemistry change, such as manganese and iron redox cycling during different seasons (Devereux *et al.* 2015).

Yu *et al.* (2017) investigated the bathyarchaeotal community in two sediment cores from the South China Sea; the authors revealed a direct strong positive correlation between bathyarchaeotal 16S rRNA gene abundance and total organic carbon content along the core depth, suggesting an overall heterotrophic lifestyle of Bathyarchaeota in the South China Sea. High-throughput sequencing of the archaeal communities and the analysis of the relationship between the distribution pattern of bathyarchaeotal subgroups and the physicochemical parameters of study sites revealed that sediment depth and sulfate concentration were important environmental factors that shape the distribution of bathyarchaeotal subgroups; Subgroup-8 was shown to be predominantly distributed in the reducing and deeper sediment layers, while Subgroup-10 was preferentially distributed in the relatively more oxidizing and shallow sediment layers (Yu *et al.* 2017). However, because of the high intragroup diversity and potential heterogeneous metabolic properties and adaptive strategies within the bathyarchaeotal subgroups, investigation into the subgroup distribution patterns at a fine-sorted phylotype level was recommended. The results also revealed that some operational taxonomic units affiliated with Subgroups-2 and -15 are dominant in all surface and bottom sediment layers in these two cores, suggesting that these operational taxonomic units might be adaptive to redox changes (Yu *et al.* 2017).

In a recent global evaluation of the archaeal clone libraries from various terrestrial environmental settings, permutational analysis that tested the relationship between Bathyarchaeota and environmental factors suggested that salinity, total organic carbon and temperature are the most influential factors impacting community distribution across different terrestrial habitats (Xiang *et al.* 2017). Co-occurrence networks in the archaeal clone libraries indicated the role of Bathyarchaeota as keystone species, and suggested their function in maintaining the stability and adaptability of the archaeal community (Xiang *et al.* 2017). Further, a close co-occurrence of Bathyarchaeota and Methanomicrobia hinted at a syntrophic association between them; the acetate production/consumption relationship between the two might be responsible for such a scenario, as proposed by metabolic predictions (He *et al.* 2016; Xiang *et al.* 2017). The IndVal species with statistical support in terrestrial environments indicated by this study were pMCG and Subgroup-5b in peat; Subgroup-5a in hot springs; Subgroup-6 in the soil; Subgroups-3, -4, -13 and -16 in estuaries; and Subgroup-15 in mangroves. The current genomic and physiological information of these subgroups also suggests their potential ecological strategies and functions in specific habitats, further highlighting their important roles in global biogeochemical cycling (Xiang *et al.* 2017).

Although the accumulated information paves the way for further clarification of the adaptation of different lineages to various environments, systematic understanding of the distribution pattern of bathyarchaeotal subgroups and influential factors is still needed. Given the diverse and complex phylogeny of the Bathyarchaeota (Kubo *et al.* 2012; Fillol *et al.* 2016), the occurrence of commonly shared physiological and metabolic properties in different lineages seems unlikely, with the evolutionary diversification of bathyarchaeotal lineages largely driven by the adaptation to various environmental conditions and available carbon and energy sources, etc. Moreover, with the rapid development and application of 16S rRNA-based high-throughput sequencing techniques for microbial ecological profiling, and 16S rRNA-independent microbial metagenomic profiling that avoids the issue of polymerase chain reaction (PCR) primer bias, a much clearer distribution pattern of diverse bathyarchaeotal subgroups can be expected; at the same time, higher resolution of local physicochemical characteristics will facilitate classification of ecological niches of bathyarchaeotal subgroups into more detailed geochemical categories.

Molecular detection methods

Several sets of PCR primers and probes have been developed to detect and quantify Bathyarchaeota in natural community (Table 1). To cover all bathyarchaeotal subgroups that are characterized by high intragroup diversity while retaining bathyarchaeotal sequence specificity is necessary but challenging. Kubo *et al.* (2012) demonstrated that the developed primers and probes result in poor coverage of Subgroups-13 to -17. The *in silico* tests revealed that primers MCG528, MCG493, MCG528 and MCG732 cover 87, 79, 44 and 27% of sequences of Subgroups-1 to -12 on average, respectively. Because of their high sequence coverage and bathyarchaeotal sequence specificity, MCG528 and MCG732 primers are recommended for the detection and quantification of Bathyarchaeota (Kubo *et al.* 2012); nevertheless, this primer pair is not suitable for quantifying Bathyarchaeota in freshwater columns and sediments (Fillol *et al.* 2015). Hence, the primer pair MCG242dF and MCG678R was developed based on a collection of bathyarchaeotal sequences of freshwater origin (Fillol *et al.* 2016). The use of MCG242dF resulted in an adequate coverage of almost all subgroups with 0/1 nucleotide mismatches, except for Subgroups-10 and -17, which showed low coverage efficiency with no nucleotide mismatches. However, after allowing for a single nucleotide mismatch, the coverage efficiency markedly increased, to around 80–90%. In the case of Subgroup-15, which branched away from other groups, MCG242dF would be associated with a relatively low coverage efficiency in the absence of nucleotide mismatches, but high (above 80%) coverage efficiency with 1 or 2 nucleotide mismatches; similarly, MCG678R would be associated with a limited coverage efficiency in the absence of nucleotide mismatches, but the coverage efficiency increases considerably with 1 or 2 nucleotide mismatches. To compare the coverage and specificity of analysis using the qPCR primer pairs MCG242dF/MCG678R and MCG528F/MCG732R for freshwater and marine sediment samples, amplicons obtained with these two primer pairs were analyzed and community structures compared (Fillol *et al.* 2015). Similar community structures across different bathyarchaeotal subgroups were revealed using the two primer pairs; however, both pairs performed poorly with respect to indicating the prevalence of Subgroup-15 in cDNA libraries from freshwater sediments (Fillol *et al.* 2015). Combinations of MCG242dF with MCG678R or MCG732R were recommended for targeting relatively long 16S rRNA gene fragments to obtain more phylogenetic information; these might be used in clone library construction or for denaturing gradient gel

Table 1. Primers and probes for molecular detection and quantification of Bathyarchaeota subgroups.

Primer	Primer/probe sequence	*Escherichia coli* position	qPCR T_a (°C)	Slot-bolt T_d (°C)	CARD-FISH FA conc.	Reference
MCG410F	TCCGCTGAGGATGGCTTTT	409–427	60	–	–	(Kubo *et al.* 2012)
MCG528F	CGGTAATACCAGCTCTCCGAG	527–547	60	–	–	
MCG528R	CTCGGAGAGCTGGTATTACCG	527–547	60	–	–	
MCG732R	CGCGTTCTAGCCGACAGC	730–748	60	–	–	
			qPCR $T_{a,e}$ (°C)			
MCG242dF	TDACCGGTDCGGGCCGTG	279–296	68	–	–	(Fillol *et al.* 2015)
MCG678R	CACCGTCGGGCGCGTTCT	740–757	68	–	–	
Bathy-442F	AGACTGGTGTCAGCCGCC	509–526	58	–	–	
Bathy-644R	CGCCACTGGTGGTCCTC	710–726	58	–	–	(Yu *et al.* 2017)
Probe						
MCG493	CTTGCCCTCTCCTTATTCC	492–510	–	55	20–30	(Kubo *et al.* 2012)
MCG528	CGGAGAGCTGGTATTACC	528–545	–	–	30	

T_a stands for qPCR annealing temperature, $T_{a,e}$ stands for annealing and extension temperature of two-step qPCR. T_d stands for dissociation temperature for RNA slot-bolt. FA conc. stands for formamide concentration in the hybridization buffer (%, vol/vol).

electrophoresis-based community fingerprinting analysis. The primer pair MCG242dF/MCG528R may potentially be used for the determination of the bathyarchaeotal community abundance, with relatively high subgroup coverage and specificity *in silico*; however, experimental tests are needed to confirm this. A pair of primers (Bathy-442F/Bathy-644R) was recently designed to target Subgroups-15 and -17; the *in silico* primer testing indicates that Bathy-442F can also adequately cover Subgroups-2, -4, -9 and -14, with Bathy-644R covering nearly all subgroups, except for Subgroups-6 and -11 (Yu *et al.* 2017). This primer pair shows good specificity toward Bathyarchaeota; it allowed amplification of 10–100 times more bathyarchaeotal 16S rRNA gene sequences from the sediment samples from the South China Sea, and the Atlantic and Antarctic Oceans than the MCG242dF/MCG678R primers (Yu *et al.* 2017).

RNA slot blot hybridization can also be used for the quantification of functionally active bathyarchaeotal 16S rRNA. The total RNA is blotted onto nylon membranes and subsequently hybridized with ^{33}P-labeled Bathyarchaeota-specific probes (Table 1). The presence and relative abundance of bathyarchaeotal rRNA can then be estimated based on the hybridization intensity (Stahl *et al.* 1988; Kubo *et al.* 2012).

CARD-FISH can be utilized for the detection, identification and enumeration of microorganisms in various environments, independently of culturing (Kubota 2013). This method has been used to target the bathyarchaeotal 16S rRNA gene with specific probes, providing information on the active bathyarchaeotal community without culturing (Table 1). To increase the permeability of the cell wall and obtain a good amplification signal, a 10-min 0.01 M HCl treatment may be employed (Kubo *et al.* 2012). The three methods described above may be used for the quantification of bathyarchaeotal abundance based on DNA and RNA targets. Their results agree well and reflect the relatively higher bathyarchaeotal fraction in marine sediments with sulfate penetration (>0.15 m below seafloor) (Kubo *et al.* 2012).

Membrane lipids are an informative indicator of the distribution and activity of living microbial cells, independently of their culturing (Sturt *et al.* 2004; Jacquemet *et al.* 2009; Lipp, Liu and Hinrichs 2009). It is well known that isoprenoid glycerol dialkyl glycerol tetraether lipids are specifically synthesized by archaea. Specific lipids, exclusively synthesized by certain archaea, can serve as a supplementary biomarker for tracing the existence and abundance of targeted archaeal groups; their isotopic com-

position can be used to indicate specific carbon acquisition pathways (Schouten, Hopmans and Damste 2013). In a recent study exploring the stratified distribution of archaeal groups in a tropical water column, the analysis of archaeal 16S rRNA community distribution was combined with isoprenoid glycerol dialkyl glycerol tetraether lipid abundance information to reveal that glycerol dibiphytanyl glycerol tetraether lacking the cyclopentane rings [GDGT(0)] likely originated from the Bathyarchaeota-enriched layer in the water column (Buckles *et al.* 2013). Because of the wide distribution of this lipid in many other archaea, it cannot be used for the detection of Bathyarchaeota and its carbon stable isotopic composition cannot be used for metabolic property deductions. However, in a study investigating the archaeal lipidome in the White Oak River estuary, the presence of the recently discovered butanetriol dibiphytanyl glycerol tetraethers correlated well with bathyarchaeotal abundance along the sediment depth (Meador *et al.* 2015). This group of lipids has not been found in natural environments or microorganism enrichments dominated by methanotrophic archaea before (Rossel *et al.* 2008; Kellermann *et al.* 2012), nor have they been detected after re-analyzing lipid extracts from the above two studies using the same method in the study (Meador *et al.* 2015). The ^{13}C-depleted nature of butanetriol dibiphytanyl glycerol tetraethers found in the study implied that members of Bathyarchaeota might be autotrophs or fueled by ^{13}C-depleted organic substrates (Meador *et al.* 2015). Further membrane lipid characterization of enriched or pure bathyarchaeotal cultures will help to validate this discovery. This will have a profound impact not only on deciphering the metabolic properties of Bathyarchaeota, by using butanetriol dibiphytanyl glycerol tetraethers as biomarkers to trace carbon acquisition by isotopic labeling, but also by representing their pivotal contribution, associated with their global abundance, to biogeochemical carbon cycling on a large ecological scale.

PHYSIOLOGICAL AND GENOMIC CHARACTERIZATION

Pre-/non-enriched sediment cultures and trophic properties

No bathyarchaeotal species have as yet been successfully cultured in pure cultures, despite their widespread distribution

in the marine, terrestrial and limnic environments (Kubo *et al.* 2012), which hampers their direct physiological characterization. The wide phylogenetic coverage increases the difficulty of inferring the general metabolic properties across whole lineages. Several pre-/non-enriched sediment cultures afforded preliminary evidence for the trophic properties and metabolic capacities of Bathyarchaeota. In one study, small amounts of stable isotope-labeled substrates, including glucose, acetate and CO_2, were introduced multiple times into slurries from different biogeochemical depths of tidal sediments from the Severn estuary (UK) to better reflect the *in situ* environmental conditions (Webster *et al.* 2010). Together with evidence of few phylogenetic changes throughout the incubation, it was suggested that the microbial community detected by stable isotopic probing could serve well in reflecting the metabolically active components. After incubation with ^{13}C-acetate, the archaeal population within a sulfate reduction zone, detected on the basis of ^{13}C-DNA, was almost entirely dominated by Bathyarchaeota (65% by Subgroup-8 and 30% by Subgroup-15) (Webster *et al.* 2010). Thaumarchaeota MG-I was present in the ^{12}C-DNA library in the corresponding zone but was not detected in the ^{13}C-DNA library, suggesting that these microbes are not able to use ^{13}C-acetate (Webster *et al.* 2010). Furthermore, another study demonstrated that the archaeal communities of the sulfate–methane transition zone at diffusion-controlled sediments of Aarhus Bay (Denmark) contain considerable amounts of Bathyarchaeota; the overall archaeal community structure did not change greatly during the experiment—its diversity was lower after 6 months of incubation under heterotrophic conditions, with periodic modest sulfate and acetate additions (Webster *et al.* 2011).

The analysis of the stable isotopic-probed microcosms from Cheesequake salt marsh sediment revealed that all 'Crenarchaeota' groups, which still include Bathyarchaeota and Thaumarchaeota (formerly Crenarchaeota MG 1.a) and other Crenarchaeota groups, are heterotrophic and do not incorporate ^{13}C-bicarbonate (Seyler, McGuinness and Kerkhof 2014). Some of these 'Crenarchaeota' were able to assimilate all ^{13}C-organic compounds tested, including acetate, glycine, urea, simple biopolymers (extracted algal lipids) and complex biopolymers (ISOGRO), while others were only detected in specific substrates (acetate or urea). Furthermore, analysis of clone libraries retrieved after ^{13}C-DNA amplification combined with matched terminal fragment length polymorphism peaks suggested that the heterotrophic bathyarchaeotal community possibly comprised Subgroups-6 and -8 (Seyler, McGuinness and Kerkhof 2014). Subgroup-15 was recently found to be enriched in ^{13}C-labeled DNA after a 3-month incubation experiment using sulfate-reducing sediments from Aarhus Bay, but was not present in the corresponding total DNA library or in a control incubation sample (i.e. with ^{12}C-acetate added); this indicated that the acetate might participate in microbial biosynthesis rather than being used for energy production (Na *et al.* 2015).

In experiments towards cultivating Bathyarchaeota from the White Oak River estuary sediments, the abundance of Bathyarchaeota in control groups (basal medium) and in experimental groups containing various substrate additives and submitted to various culture processing steps were compared (Gagen *et al.* 2013). That approach revealed an order of magnitude increase in bathyarchaeotal abundance in both the control and experimental groups compared with time zero; however, no significant increase of bathyarchaeotal abundance was observed in experimental groups with substrate additives and various cultivation processing steps, compared with control groups with

basal medium alone. The diversity of bathyarchaeotal community turns out to be similar in the four cultivation treatments (basal medium, addition of an amino acid mix, H_2-CO_2 headspace and initial aerobic treatment). Following the four treatments, the viable bathyarchaeotal communities mainly comprised Subgroups-4 and -8, thus indicating that these two subgroups could tolerate the initial aerobic conditions (Gagen *et al.* 2013).

A recent study found that the refractory aromatic polymer lignin stimulated the growth of Bathyarchaeota (Subgroup-8) and they incorporated CO_2 as a carbon source autotrophically and utilized lignin as an energy source (Yu *et al.* 2018). During the enriching process with lignin addition, the Subgroup-8 abundance climbed over 10 times compared with the initial stage and became the most dominant archaeal species. The incorporation of ^{13}C-bicarbonate into the archaeal lipids (potential bathyarchaeotal-specific biphytanes) was significantly observed only with lignin addition. The clear growth stimulus and lignin-related ^{13}C-bicarbonate incorporation into lipids strongly suggests that Bathyarchaeota (Subgroup-8) may be able to use the second-most abundant biopolymer lignin on Earth (Yu *et al.* 2018). Furthermore, genomic features of Subgroup-8 resolved from the metagenome of lignin-added enrichments evidence the putative lignin and aromatics degrading genes, thus it is hypothesized that Subgroup-8 catalyzes methoxy-groups of lignin, and combines the resulting methyl-group with CO_2 to acetyl-coenzyme A (CoA) through the Wood–Ljungdahl pathway for either biosynthesis or acetogenesis in downstream pathways (Yu *et al.* 2018).

Genomic characterization of fosmid fragments, single amplified genomes and genomic bins

Fosmid clone 37F10 containing a genome fragment originating from a bathyarchaeotal member was isolated from a metagenomic library constructed from Pearl River sediment samples (Meng *et al.* 2009); its G + C content indicated that this genomic fragment had two portions: an archaeon-like portion (42.2%) and a bacterium-like portion (60.1%) (Meng *et al.* 2009; Li *et al.* 2012). More importantly, the first-ever bacteriochlorophyll a synthase (BchG) of archaeal origin was identified in the archaeal portion of the genomic fragment, and its function confirmed by producing BchG in a heterologous expression system (Meng *et al.* 2009). A phylogenetic tree based on the sequences of UbiA prenyltransferase superfamily proteins, including ChlG/BchG and additional five subfamilies of this superfamily, revealed that this unique BchG of archaeal origin groups within the ChlG/BchG family; however, it diverged earlier than the bacterial BchG proteins. With respect to its function, the protein might be responsible for photosynthesis in archaea; this suggests that photosynthesis may have evolved before the divergence of the bacteria and archaea domains (Meng *et al.* 2009; Li *et al.* 2012). Three fosmid clones harboring bathyarchaeotal genomic fragments were screened from the South China Sea sediments (0–5 cm depth) (Li *et al.* 2012). Low collinear regions were found between bathyarchaeotal and reported archaeal genomic fragments, suggesting that the gene arrangement of Bathyarchaeota is distinct from that of sequenced archaea. Open reading frames encoded by the three fosmid clones comprised genes related to lipid biosynthesis, energy metabolism and resistance to oxidants. In addition, some regions of the bathyarchaeotal genome might have been acquired from bacteria because of the aberrant tetranucleotide frequency in the genomic fragments of Bathyarchaeota and bacterial phylogenetic origins of these genomic

fragments (Li *et al.* 2012). Recently, two more bathyarchaeotal fosmid clones were screened from estuarine mangrove sediments (Meng *et al.* 2014). Gene arrangement in these two fosmid clones, together with the previously recovered bathyarchaeotal fosmid sequences, confirmed low collinearity with other known archaeal genomes. Genomic fragments of the fosmid clone 75G8 harbor a putative methyl-accepting chemotaxis protein- and 4-carboxymuconolactone decarboxylase-encoding genes, suggesting that this bathyarchaeotal member (Subgroup-8) is able to utilize aromatic compounds. The production of a putative 4-carboxymuconolactone decarboxylase was evident when the mangrove sediments were supplemented with protocatchuate, further suggesting the capacity of certain bathyarchaeotal members to degrade aromatic compounds (Meng *et al.* 2014).

Single amplified genomes (SAGs) of a Subgroup-15 bathyarchaeotal member from the Aarhus Bay sediments harbor genes for predicted extracellular protein degrading enzymes, such as clostripain (Lloyd *et al.* 2013). A complete set of active sites and signal sequences for extracellular transport is also encoded by bathyarchaeotal SAGs (Lloyd *et al.* 2013). Peptidases targeting D-amino acids, which are highly enriched in the peptidoglycan of bacterial cell walls, are encoded as well, indicating that Bathyarchaeota may have acquired the capacity to degrade recalcitrant components of bacterial cell walls, i.e. the most persistent detrital matter in marine sediments (Lomstein *et al.* 2012; Lloyd *et al.* 2013). Bathyarchaeotal SAGs also encode pathways for the intracellular breakdown of amino acids. Considering the ubiquity and frequent predominance of Bathyarchaeota in marine sediments, as well as the high abundance and potential activity of extracellular peptidases that they encode, it has been proposed that Bathyarchaeota may play a previously undiscovered role in protein remineralization in anoxic marine sediments.

Recently, two bathyarchaeotal genome bins (BA1 and BA2) were recovered from the formation waters of coal-bed methane wells within the Surat Basin (Evans *et al.* 2015). BA1 (Subgroup-3) genome contains many genes of the reductive acetyl-CoA (Wood–Ljungdahl) pathway and key genes of the methane metabolism pathway. It harbors methyl-coenzyme M reductase (MCR)-encoding genes, and many identified and unidentified methyltransferase-encoding genes for the utilization of various methylated compounds, but lacks most of the genes encoding the subunits of Na$^+$-translocating methyl-H$_4$MPT:coenzyme M methyltransferase, suggesting that the organism does not engage in hydrogenotrophic methanogenesis. It also contains typical methane metabolism genes (*hdrABC* and *mvhADG*) but lacks *hdrE*, similar to Methanomassiliicoccales genomes (Evans *et al.* 2015). The gene for cytoplasmic flavin adenine dinucleotide-containing dehydrogenase (*glcD*) co-located with *hdrD*, indicating that BA1 uses lactate to reduce heterodisulfide in methanogenesis. BA1 also lacks other genes for energy-conserving complexes, including F420H$_2$ dehydrogenase, energy-converting hydrogenases A and B, *Rhodobacter* nitrogen fixation complex and V/A-type ATP synthase. The BA2 (Subgroup-8) genome contains MCR-encoding genes and additional genes of typical methane metabolism, like BA1, reflecting a similar methylotrophic methanogenesis activity. However, it has lost the majority of genes involved in the methyl branch of the Wood–Ljungdahl pathway and also lost energy-conserving complexes, similar to BA1. Regarding the functional properties, metabolic pathway analysis revealed that BA1 is a peptide and glucose fermenter, while BA2 is a fatty-acid oxidizer (Evans *et al.* 2015). Furthermore, both BA1 and BA2 lack ATP-synthase, indicating that they are restricted to substrate-level phosphorylation for energy, which was first found in methanogenic archaea (Evans *et al.* 2015).

The metagenomic binning of WOR estuarine sediment DNA led to the reconstruction of draft genomes of four widespread Bathyarchaeota, with the genome completeness in the range of 48–98% (Lazar *et al.* 2016). Four genomes (Subgroups-1, -6, -7 and -15) were recovered from the sediment metagenome. Subgroup-6 genome was reconstructed from the surficial sulfate reduction zone, harboring genes encoding enzymes with predicted functions in the degradation of extracellular plant-derived mono- and polysaccharides. The reconstructed bathyarchaeotal genomes (except for Subgroup-15) also encode proteins with the ability to import extracellular carbohydrates. Bathyarchaeotal subgroups analyzed here acquired an almost complete Embden–Meyerhof –Parnas glycolysis pathway. The Subgroups-1, -6 and -15 genomes also encoded the methyl glyoxylate pathway, which is typically activated when slow-growing cells are exposed to an increased supply of sugar phosphates (Weber, Kayser and Rinas 2005). The Subgroup-15 genome contained genes encoding extracellular peptidases, consistent with previous findings for this subgroup (Lloyd *et al.* 2013); however, other bathyarchaeotal subgroups lack genes responsible for extracellular protein degradation, suggesting that they can only utilize small amino acids or oligopeptides, as suggested by their genomes. Hence, Bathyarchaeota acquired the core heterotrophic metabolic capacity for processing complex carbohydrates, and an additional ability to utilize peptides and amino acids, as suggested before (Seyler, McGuinness and Kerkhof 2014). In terms of energy metabolism, these archaea contain the Wood–Ljungdahl pathway, capable of generating acetyl-CoA autotrophically by CO$_2$ and H$_2$. Acetyl-CoA might be involved in acetate generation in a fermentative pathway; however, genomic evidence suggests that Subgroup-1 cells might rely on both fermentative and respiratory metabolism (a simple respiratory metabolism based on a membrane-bound hydrogenase). The evidence for the presence of respiratory metabolism in other bathyarchaeotal subgroups is ambiguous although it cannot be excluded (Lazar *et al.* 2016). Genes responsible for the dissimilatory nitrite reduction to ammonium (*nirB* and *nrfD*) were identified in Subgroups-1, -17 (formally Subgroup-7/17), -6 and -15, respectively, suggesting the potential existence of a respiratory pathway involving nitrite reduction (Lazar *et al.* 2016). No methane metabolism genes were recovered from bathyarchaeotal genomic bins or any contigs from the WOR estuarine sediments, in contrast to an earlier study (Evans *et al.* 2015).

More recently, He *et al.* (2016) reconstructed six nearly complete bathyarchaeotal genomes (Subgroups-13, -15, -16, -18 and -19) from the Guaymas Basin subsurface sediment. Based on the genomic evidence, the authors concluded that some lineages of Bathyarchaeota are similar to *bona fide* bacterial homoacetogens, with pathways for acetogenesis and fermentative utilization of a variety of organic substrates (He *et al.* 2016). A subsequent heterologous expression and activity assays of the bathyarchaeotal acetate kinase gene *ack* demonstrated the ability of these bathyarchaeotal members to grow as acetogens. These findings expand the metabolic potential of archaea and argue for a revision of the role of archaea in the carbon cycle in marine sediments (He *et al.* 2016). Furthermore, genes encoding ATP sulfurylase, for the reduction of sulfate to adenosine 5′-phosphosulfate, and adenylyl-sulfate reductase, for the reduction of adenosine 5′-phosphosulfate to sulfite, were identified in a metagenomic assembly of Bathyarchaeota TCS49 genome from the Thuwal cold seep brine pool of the Red Sea; this suggests that specific

bathyarchaeotal members might harbor a dissimilatory sulfate reduction pathway, indicating the existence of additional potential metabolic capacities of Bathyarchaeota (Zhang *et al.* 2016).

The currently available bathyarchaeotal genomes shared 63.5% similarity on average, indicating a wide phylogenetic diversity at the genome scale (Fig. 3B). The metabolic properties are also considerably diverse based on genomic analysis (Fig. 3C). Combined with the aforementioned specific heterotrophic metabolic potentials of members within bathyarchaeotal subgroups and their occurrence in sediment layers of distinct biogeochemical properties (Lazar *et al.* 2015), it was proposed that the acquisition of diverse physiological capacities by Bathyarchaeota is driven by adaptation to specific habitats rather than there being a common metabolic capacity.

ECOLOGICAL FUNCTIONS AND EVOLUTION OF BATHYARCHAEOTA

Methanogenesis or anaerobic oxidation of methane?

Genomic inferences from the two reconstructed bathyarchaeotal genomic bins from the coal-bed methane wells suggest that some Bathyarchaeota are methylotrophic methanogens feeding on a wide variety of methylated compounds, possessing an additional ability to ferment peptides, glucose and fatty acids (Evans *et al.* 2015). It was proposed that reduced ferredoxin generated by peptide and/or glucose might be used for the reduction of methyl groups on methylated compounds to subsequently generate methane (Evans *et al.* 2015). The identification of key genes of the MCR complex (*mcrA*, *mcrB* and *mcrG*), and the presence of *hdrABC* and *mcvhADG* responsible for the cycling of coenzyme M (CoM) and coenzyme B (CoB), suggest their role in the methanogenesis machinery that mediates the CoM-S-S-CoB cycling, similar to Euryarchaeota methanogens (Evans *et al.* 2015). Furthermore, the MCR complexes found in the BA1 and BA2 genomes are phylogenetically divergent from traditional MCR and they coevolved as a whole functional unit, indicating that methane metabolism began to evolve before the divergence of the Bathyarchaeota and Euryarchaeota common ancestors (Evans *et al.* 2015). Since these two genomic bins represent only a small fraction of all bathyarchaeotal lineages, and no evidence of methanogenic machinery is apparent in the recent parallel genomic binning data, the ability to metabolize methane might not be shared by all subgroup lineages (Lloyd *et al.* 2013; Meng *et al.* 2014; He *et al.* 2016; Lazar *et al.* 2016). However, the global methane cycle should be reconsidered since the previously unrecognized methane metabolic capacity appears to be present within such a widespread and abundant phylum.

On the other hand, because of the bidirectionality of these enzymes in methane metabolism (Boetius *et al.* 2000; Knittel and Boetius 2009), it is still possible that some members of Bathyarchaeota are involved in anaerobic methane oxidation. Further, based on genomic inferences, Evans *et al.* (2015) presumed the syntrophy between Bathyarchaeota and sulfate-reducing bacteria (SRB) toward anaerobic oxidation of methane (AOM) (Evans *et al.* 2015). Furthermore, the lack of genes for ATPases and membrane-bound electron transport enzymes in the two genomic bins (BA1 and BA2) and the presence of the ion pumping, energy-converting hydrogenase complex (only in BA1), which might allow solute transportation independently of energy-generation mechanisms, suggest that the soluble substrate transportation is solely responsible for energy conservation (Evans *et al.* 2015). Consequently, CO_2 appears to be the only electron acceptor mediating AOM, like

in a reverse acetoclastic methanogenesis (Hallam *et al.* 2004; Wang *et al.* 2014). Methane would be oxidized in a stepwise manner to methyl-tetrahydromethanopterin (CH$_3$-H$_4$MPT); the methyl group of CH$_3$-H$_4$MPT and CO_2 would then be subjected to a CO dehydrogenase/acetyl-CoA synthase (CODH/ACS complex); CO_2 would be fixed by a reverse CO dehydrogenation to CO, and then coupled with a methyl group and CoA to generate acetyl-CoA; ATP would be generated in the course of substrate-level phosphorylation from ADP, with one acetate molecule simultaneously generated by a reverse ADP-forming acetyl-CoA synthase. The product, acetate, would then be used by acetate-consuming SRB to benefit the thermodynamic efficiency of AOM. The syntrophic relationship between Bathyarchaeota and SRB would be similar to the anaerobic methane-oxidizing archaea (ANME)/SRB consortium, and acetate would be maintained at a low level as a transient intermediate (Boetius *et al.* 2000; Hinrichs and Boetius 2002).

The energy landscape of a local environment, i.e. the census of energy availability for redox reactions, is used, to some extent, to constrain and predict the distribution of functional groups of chemotrophic microorganisms (Amend *et al.* 2011; LaRowe and Amend 2014). A model based on the thermodynamic considerations of chemicals and temperatures may be used to offer a framework linking the distribution of microbial groups and energy landscapes (Amend *et al.* 2011; LaRowe and Amend 2014; Dahle *et al.* 2015). In a recent study, Bathyarchaeota and ANME were shown to predominate on the flange of a hydrothermal chimney wall in the Soria Moria Vent field, where the local energy condition favors anaerobic methane oxidizers (Dahle *et al.* 2015). In some flange subsamples, Bathyarchaeota were even more dominant than ANME; however, compared with the well-studied metabolism of ANME, the exact function of Bathyarchaeota in that ecological setting remains unknown. Reconsideration of the potential methane-oxidizing contribution of Bathyarchaeota would refine the congruency between the predicted and observed microbial communities, i.e. the potential AOM metabolism of Bathyarchaeota in the flange of the hydrothermal vent would be consistent with the aforementioned genomic inferences (Evans *et al.* 2015). The possibility of the replacement of the AOM function of ANME by Bathyarchaeota was also suggested by a microbial community composition in a study of the microbial colonization within an artificial micro-niche, basaltic glass imposed by hydrothermal conditions (Callac *et al.* 2013).

The active microbial community in four SMTZ layers of the ODP Leg 201 subsurface sediment cores off Peru was dominated by MBG-B and Bathyarchaeota (Biddle *et al.* 2006). Energy flux analysis revealed that AOM and slow degradation of refractory sedimentary organic matter were the two principal energy generation pathways in the local community. The potential AOM metabolic capacity of Bathyarchaeota could help to fully address the isotopic relationship between the archaeal biomass and the ambient environmental carbon pools, as follows. (i) The $\delta^{13}C$ signature of the archaeal biomass suggests that only a small fraction of local archaea in SMTZ utilize methane, which might be explained by the contribution of Bathyarchaeota in the biomass; until now, only one line of evidence points to the acquisition of methane metabolism by Bathyarchaeota (Lloyd *et al.* 2013; Evans *et al.* 2015; Lazar *et al.* 2015; He *et al.* 2016). (ii) Similar $\delta^{13}C$ signatures of the archaeal biomass and total organic carbon suggest that the organic matter assimilation contributes to the bulk of the archaeal biomass; the relatively small $\delta^{13}C$ signature of the archaeal biomass in comparison with the dissolved inorganic carbon suggests that only a small amount of archaeal biomass

is derived from autotrophic CO_2 fixation (Biddle *et al.* 2006). This could be explained by the versatile pathways of organic matter assimilation present in the majority of Bathyarchaeota, reflected by inferences from genomic data. The wide availability of buried organic matter in the marine subsurface would favor the heterotrophic feeding of Bathyarchaeota. (iii) The relatively small $\delta^{13}C$ signature of the archaeal intact polar lipids in comparison with the archaeal biomass suggests that the C isotopic fractionation during lipid biosynthesis is different from that of typical methylotrophic methanogens (Summons, Franzmann and Nichols 1998). Based on the above, it is proposed that Bathyarchaeota might mediate the AOM without assimilating the carbon in methane. This would be supported by a coupled AOM and syntrophic SRB metabolism, with methane consumed by Bathyarchaeota through reverse acetoclastic methanogenesis with the production of acetate, which is readily oxidized by sulfate in SRB. In this process, methane is not assimilated by Bathyarchaeota but serves as an energy source. Metagenomic evidence of sulfate reductase-encoding genes in the upper region of SMTZ of the OPD site 1229 provides more hints to the potential synergistic metabolism of AOM coupled with sulfate reduction (Biddle *et al.* 2008).

More recently, the proposed genus 'Candidatus Syntrophoarchaeum' was shown to be able to anaerobically oxidize butane in a manner similar to ANME oxidation of methane, by reverse methanogenesis, a process that is initially mediated by MCR (Laso-Pérez *et al.* 2016). Two highly abundant MCR variants were detected in *Ca*. S. butanivorans protein extracts; they are probably responsible for the initial step of butane activation to generate butyl-CoM. Interestingly, one of the highly abundant McrA subunits of *Ca*. S. butanivorans forms a distinct cluster with those of Bathyarchaeota origin, separately from other methanogens and methanotrophs (Laso-Pérez *et al.* 2016). Given the substrate specificity of this MCR type in utilizing butane instead of methane, and amino acid divergence of this MCR type from its methane metabolizing related counterparts, it is possible that the MCR clusters in some members of Bathyarchaeota are responsible for butane oxidation instead of methane metabolism (Laso-Pérez *et al.* 2016). Future experiments investigating substrate specificity of these proteins and analyses of the intermediate metabolites will help establish their actual functions.

Acetogenesis pathway as a hybrid of bacterial and archaeal features

In the two recent metagenomic bathyarchaeotal binning studies, nearly all the identified bins placed H₄MPT as a C₁-carrier in the Wood–Ljungdahl pathway, which is often used by the methanogenic archaea for carbon fixation (He *et al.* 2016; Lazar *et al.* 2016). However, in the above binning studies, none of the genomes encoded enzymes involved in the final methane production step (McrABG), suggesting that the Wood–Ljungdahl pathway is not used for methane production but for acetyl-CoA generation and further acetogenesis.

He *et al.* (2016) demonstrated that half of the bathyarchaeotal genomes encode a set of phosphate acetyltransferase (Pta) and acetate kinase (Ack) for acetate production or assimilation, usually observed in bacteria. However, according to the genomic information on most archaeal acetogens and bathyarchaeotal genomic bins obtained by Lazar *et al.* (2016), it appears that these microbes rely on the acetyl-CoA synthetase (Acd) to generate acetate (He *et al.* 2016). Although the Pta-Ack pathway has been

previously identified in the methanogenic genus *Methanosarcina*, it was shown that the encoding *pta-ack* gene pair might be derived from a horizontal transfer of genes of bacterial origin (Fournier and Gogarten 2008). Phylogenetic analysis of the Pta and Ack coding sequences in He *et al.*'s study revealed that these genes form a monophyletic clade and are different from all other know sequences, indicating that they evolved independently of the currently known bacterial counterparts (He *et al.* 2016). This is the first ever genomic evidence for homoacetogenesis, the ability to solely utilize CO_2 and H_2 to generate acetate, in an archaeal genome and of distinct archaeal phylogenetic origin other than that of Bacteria (He *et al.* 2016). Subsequent heterologous expression of bathyarchaeotal Ack revealed that the enzyme can catalyze the biochemical reaction in the direction from acetyl phosphate to acetate, with a higher affinity for the substrates than the products (He *et al.* 2016). The exclusive archaeal origin of the Ack-Pta homoacetogenesis pathway is different from other archaeal acetogenesis systems but shares functional similarity with its bacterial origin counterparts, although it is phylogenetically divergent (He *et al.* 2016). Collectively, these findings indicate a hybrid of archaeal and bacterial features for acetogenesis of Bathyarchaeota.

Heterotrophism centralized on the generation of acetyl-CoA

Physiological incubation experiments with stable isotopic probing demonstrated that members of Bathyarchaeota are able to assimilate a wide variety of the tested ^{13}C-organic compounds, including acetate, glycine, urea, simple biopolymers (extracted algal lipids) and complex biopolymers (ISOGRO) (Webster *et al.* 2010; Seyler, McGuinness and Kerkhof 2014). Genomic inferences from SAGs and genome-resolved metagenomic bins provide further genomic support for the heterotrophic lifestyle of Bathyarchaeota, rendering them capable of adapting to various environments and becoming one of the most successful lineages globally (Fig. 3). Four major heterotrophic pathways centralized on the acetyl-CoA generation are summarized below, reflecting the core metabolism of fermentation and acetogenesis (Fig. 4), although these might not necessarily exist in all bathyarchaeotal subgroups (Fig. 3) (Lloyd *et al.* 2013; Evans *et al.* 2015; Lazar *et al.* 2015; He *et al.* 2016; Lazar *et al.* 2016; Lever 2016). Proteins or polypeptides are first degraded by extracellular peptidases, with the resultant amino acids and oligopeptides imported into the cell, where they would be finally metabolized into acetyl-CoA via the peptide-degradation pathway. The uptake and breakdown of polymeric hydrocarbons is facilitated by extracellular hydrolases; Bathyarchaeota also acquired the Embden–Meyerhof –Parnas/Entner–Doudoroff glycolysis and gluconeogenesis pathway for the core hydrocarbon utilization metabolism. Furthermore, evidence of fatty acid and aromatic compound utilization by Bathyarchaeota has been presented (Meng *et al.* 2014; Evans *et al.* 2015; He *et al.* 2016); these transformations would be supported by the beta-oxidation pathway and a potential anaerobic aromatic compound degradation pathway. The available genomic evidence of various known and unknown methyltransferases harbored by BA1 and BA2 suggests the existence of a methylated compound utilization pathway, with the methyl group being ultimately reduced to CH_3-H₄MPT and integrated into the methyl-branch of the Wood–Ljungdahl pathway (Evans *et al.* 2015). Moreover, the carbonyl branch of the Wood–Ljungdahl pathway might reduce CO_2 into acetyl-CoA. The capability to utilize a wide variety of substrates might comprise
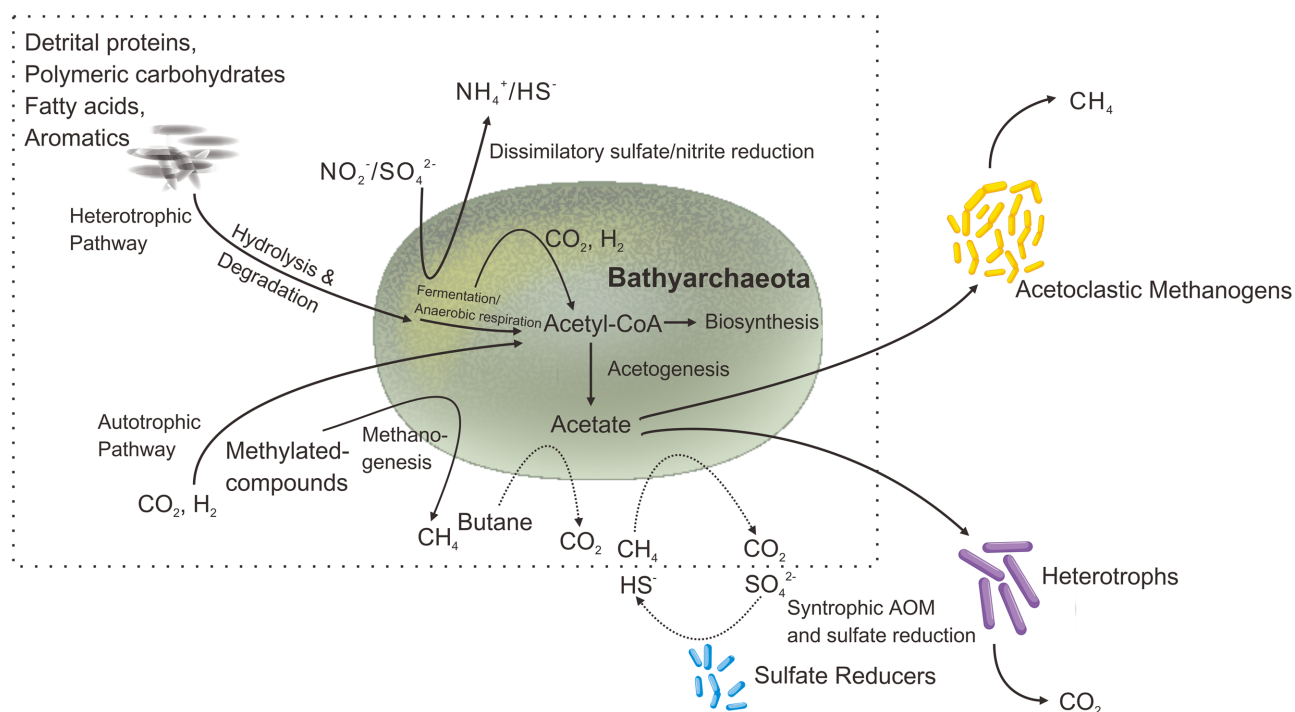
Figure 4. Metabolic potential of Bathyarchaeota and their interactive relationships with other microorganisms.

an effective strategy for competing with substrate specialists for energy sources in various environments (Li *et al.* 2015), such as detrital protein-rich deep seafloor sediments and estuarine sediments containing various carbohydrates. The central product, acetyl-CoA, would either (i) be involved in substrate-level phosphorylation to generate acetate and ATP, catalyzed by an ADP-forming acetyl-CoA synthase as in other peptide-degrading archaea; (ii) be metabolized to generate acetate through the Pta-Ack pathway, similarly to *bona fide* bacterial homoacetogens; or (iii) be utilized for biosynthesis, e.g. lipid and amino acid synthesis (Fig. 4) (Evans *et al.* 2015; He *et al.* 2016; Lazar *et al.* 2016).

Implications for the ecology and evolution of Bathyarchaeota

Both Bathyarchaeota and the recently identified more basally branched Lokiarchaeota acquired the H_4MPT-dependent Wood–Ljungdahl pathway and the hydrogen-dependent electron bifurcating system MvhADG-HdrABC, viewed as typical for the anaerobic and hydrogen-dependent archaeal lifestyle (Lazar *et al.* 2016; Sousa *et al.* 2016). They also acquired some subunits of coenzyme F420 hydrogenase; this enzyme generates reduced ferredoxin, with hydrogen as the electron donor, as an alternative to MvhADG in many Methanomicrobiales (Thauer *et al.* 2008; Lazar *et al.* 2016; Sousa *et al.* 2016). However, Lokiarchaeota and most members of the Bathyarchaeota phylum lack the essential methane metabolizing elements, such as CoB or CoM synthase and methyl-CoM reductase, etc., though they use H_4MPT as the C_1-carrier, which is common in methanogens. These archaeal groups are the phylogenetically closest ones to the protoeukaryote that served as the mitochondrion-acquiring host; this gave rise to a hydrogen hypothesis that explains their hydrogen-dependent metabolism to address the mitochondrion acquisition and subsequent endosymbiont processes. According to that hypothesis, the proto-mitochondrion bacterium was capable of both respiration and anaerobic H_2-producing fermentation; anaerobic syntrophy with respect to H_2 brought about a physical association with an H_2-dependent host and initiated a symbiotic association with the host; this led to endosymbiosis, after engulfment by the host cell (Martin and Muller 1998; Martin *et al.* 2016).

Methanogenesis and acetogenesis are considered to be the two most fundamental and ancient microbial biochemical energy conservation processes, and they both employ the Wood–Ljungdahl pathway for CO_2 reduction and ATP generation (Weiss *et al.* 2016). Methanogens and acetogenic Clostridia are the most frequent basal-branching archaea and bacteria, respectively, in phylogenetic reconstructions reflecting the descendants of the last universal common ancestor; gene categories proposed for the last universal common ancestor also point to the acetogenic and methanogenic roots, reflecting its autotrophic lifestyle as H_2-dependent and N_2-fixing, utilizing the Wood–Ljungdahl pathway and originating from a hydrothermal environmental setting (Weiss *et al.* 2016). Bathyarchaeota possesss a *bona fide* homoacetogenesis pathway of archaeal phylogenetic origin, as confirmed by functional studies, indicating a distinct evolutionary pathway of acetogenesis in archaea, different from horizontal transfer from bacteria (He *et al.* 2016). Methane metabolism pathways have been identified in members of phylum Bathyarchaeota and in the recently discovered phylum Verstraetearchaeota, placing the origin of methanogenesis before the divergence of Euryarchaeota (Evans *et al.* 2015; Vanwonterghem *et al.* 2016). The discovery of BchG of archaeal origin in the genomic content of Bathyarchaeota also suggests that an archaeon-based photosynthetic pathway might exist in nature, and that photosynthesis might have evolved before the divergence of bacteria and archaea (Meng *et al.* 2009). The versatile metabolic properties of Bathyarchaeota, including acetogenesis, methane cycling, potential photosynthesis, and dissimilatory nitrite and sulfate reduction, etc., indicate that their

ecological and phylogenetic characteristics are quite diverse, and given their basal phylogenetic position at the root of archaea, the evolutionary paths of those capabilities are also of great meaning for understanding the evolution of early life (Evans *et al.* 2015; He *et al.* 2016; Lazar *et al.* 2016; Zhang *et al.* 2016).

Members of Bathyarchaeota are able to use CO_2 and H_2 from natural sources and fermentation products to fuel acetogenesis (He *et al.* 2016; Martin *et al.* 2016). The ability to use a wide range of substrates for energy conservation and biosynthesis, rather than a single reductive acetyl-CoA pathway, enhances the survival of Bathyarchaeota in energy-limited environments (Lazar *et al.* 2016). Meanwhile, the ability to utilize a wide variety of substrates could have allowed Bathyarchaeota to avoid a direct competition with other substrate specialists, such as methanogens and sulfate reducers; in contrast, organic matter degradation to generate acetate might be more energetically favorable for Bathyarchaeota than for other bacterial acetogens, as the former do not need to invest in ATP to activate formate; subsequently, Bathyarchaeota plays the role of active carbon transformers, especially in the subsurface sediments, to fuel the heterotrophy and acetoclastic methanogenesis processes and facilitate coupled carbon cycling (Fig. 1) (He *et al.* 2016; Lazar *et al.* 2016).

SUMMARY AND OUTLOOK

In summary, the most recent research advances have considerably expanded our knowledge of Bathyarchaeota, their distribution, ecology and physiological and genomic properties since their first discovery and definition about two decades ago. Recent data point to the global occurrence of Bathyarchaeota and their potential impact on global carbon transformation, highlighting their important role as a group of global generalists participating in carbon cycling, similar to euryarchaeotal methanogens and Thaumarchaeota. Among the presently recognized 25 bathyarchaeotal subgroups, eight are delineated as significantly niche-specific based on their marine/freshwater segregation. The deduced last common ancestor of Bathyarchaeota might be a saline-adapted organism, which evolved from saline to freshwater habitats during the diversification process, with the occurrence of few environmental transitional events. Multiple genomic and physiological traits of these microorganisms have been coming to light in recent decades with the advent of stable isotope labeling and metagenomic profiling methods. It is evident that the phylogenetically diverse subgroups are heterotrophs with metabolism centralized around acetyl-CoA generation. They are able to use a variety of substrates, including (i) detrital proteins, (ii) polymeric carbohydrates, (iii) fatty acids/aromatic compound, (iv) methane (or short alkane) and methylated compounds, and/or (v) potentially other organic matter to generate acetyl-CoA, subsequently using it to obtain energy or assimilate it in biosynthetic processes. Because of the universal distribution and predominance of Bathyarchaeota, not only in the marine sediments but also in terrestrial sediments and various other eco-niches, and because of their versatile metabolism (including acetogenesis, methane metabolism, and dissimilatory nitrate and sulfate reduction) and potential interactions with ANME archaea, acetoclastic methanogens and heterotrophic bacteria, the ecological importance of this group of generalists has entered the limelight and needs further exploration.

Given that they are abundant, globally distributed and phylogenetically diverse, continued exploration of new potential bathyarchaeotal subgroups is encouraged. The knowledge of their physiological and genomic properties, as well as their adaptive strategies in various eco-niches, is nonetheless still rudimentary. First, successful enrichment methods that would allow harvesting sufficient bathyarchaeotal biomass to explore their physiological and genomic characteristics have not yet been established. A successful enrichment, with nearly pure biomass of certain subgroups of Bathyarchaeota, would enable a more efficient investigation of their metabolic capacities using stable isotope-labeled substrates, and establishing a direct link between the genotype and phenotype. Second, determining whether the methane cycling capacity is confined to certain subgroups or whether numerous subgroups or lineages are capable of methane cycling, and if so, the nature of their shared evolutionary or genomic characteristics, is of utmost importance. Obtaining direct physiological evidence for the generation or oxidization of methane by Bathyarchaeota in the future is also important. Third, only limited reports on the distribution patterns of bathyarchaeotal subgroups and the associated environmental factors are available. For us, phenotypical and genotypical information on subgroups whose existing patterns have only been sporadically reported still remains elusive and more explicit investigations are lacking. Future efforts should be encouraged to address the fundamental issues of the diversity and distribution patterns of Bathyarchaeota, and their vital roles in global carbon cycling.

SUPPLEMENTARY DATA

Supplementary data are available at *FEMSRE* online.

REFERENCES

Amend JP, McCollom TM, Hentscher M *et al.* Catabolic and anabolic energy for chemolithoautotrophs in deep-sea hydrothermal systems hosted in different rock types. *Geochim Cosmochim Acta* 2011;**75**:5736–48.

Anantharaman K, Brown CT, Hug LA *et al.* Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. *Nat Commun* 2016;**7**: 13219.

Auguet J-C, Barberan A, Casamayor EO. Global ecological patterns in uncultured Archaea. *ISME J* 2010;**4**:182–90.

Barns SM, Delwiche CF, Palmer JD *et al.* Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *Proc Natl Acad Sci USA* 1996;**93**:9188–93.

Battistuzzi FU, Feijao A, Hedges SB. A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land. *BMC Evol Biol* 2004;**4**:44.

Biddle JF, Lipp JS, Lever MA *et al.* Heterotrophic Archaea dominate sedimentary subsurface ecosystems off Peru. *Proc Natl Acad Sci USA* 2006;**103**:3846–51.

Biddle JF, Fitz-Gibbon S, Schuster SC *et al.* Metagenomic signatures of the Peru Margin subseafloor biosphere show a genetically distinct environment. *Proc Natl Acad Sci USA* 2008;**105**:10583–8.

Boetius A, Ravenschlag K, Schubert CJ *et al.* A marine microbial consortium apparently mediating anaerobic oxidation of methane. *Nature* 2000;**407**:623–6.

Buckles LK, Villanueva L, Weijers JWH *et al.* Linking isoprenoidal GDGT membrane lipid distributions with gene abundances of ammonia-oxidizing Thaumarchaeota and uncultured crenarchaeotal groups in the water column of a tropical lake (Lake Challa, East Africa). *Environ Microbiol* 2013;**15**:2445–62.

Callac N, Rommevaux-Jestin C, Rouxel O *et al.* Microbial colonization of basaltic glasses in hydrothermal organic-rich sediments at Guaymas Basin. *Front Microbiol* 2013;**4**:250.

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 2009;**25**:1972–3.

Coolen MJL, Cypionka H, Sass AM *et al.* Ongoing modification of Mediterranean Pleistocene sapropels mediated by prokaryotes. *Science* 2002;**296**:2407–10.

Dahle H, Okland I, Thorseth IH *et al.* Energy landscapes shape microbial communities in hydrothermal systems on the Arctic Mid-Ocean Ridge. *ISME J* 2015;**9**:1593–606.

Devereux R, Mosher JJ, Vishnivetskaya TA *et al.* Changes in northern Gulf of Mexico sediment bacterial and archaeal communities exposed to hypoxia. *Geobiology* 2015;**13**:478–93.

Evans PN, Parks DH, Chadwick GL *et al.* Methane metabolism in the archaeal phylum Bathyarchaeota revealed by genome-centric metagenomics. *Science* 2015;**350**:434–8.

Ferry JG, House CH. The stepwise evolution of early life driven by energy conservation. *Mol Biol Evol* 2006;**23**:1286–92.

Fillol M, Sànchez-Melsió A, Gich F *et al.* Diversity of Miscellaneous Crenarchaeotic Group archaea in freshwater karstic lakes and their segregation between planktonic and sediment habitats. *FEMS Microbiol Ecol* 2015;**91**.

Fillol M, Auguet J-C, Casamayor EO *et al.* Insights in the ecology and evolutionary history of the Miscellaneous Crenarchaeotic Group lineage. *ISME J* 2016;**10**:665–77.

Fournier GP, Gogarten JP. Evolution of acetoclastic methanogenesis in *Methanosarcina* via horizontal gene transfer from cellulolytic *Clostridia*. *J Bacteriol* 2008;**190**:1124–7.

Fry JC, Parkes RJ, Cragg BA *et al.* Prokaryotic biodiversity and activity in the deep subseafloor biosphere. *FEMS Microbiol Ecol* 2008;**66**:181–96.

Gagen EJ, Huber H, Meador T *et al.* Novel cultivation-based approach to understanding the Miscellaneous Crenarchaeotic Group (MCG) archaea from sedimentary ecosystems. *Appl Environ Microbiol* 2013;**79**:6400–6.

Hallam SJ, Putnam N, Preston CM *et al.* Reverse methanogenesis: testing the hypothesis with environmental genomics. *Science* 2004;**305**:1457–62.

He Y, Li M, Perumal V *et al.* Genomic and enzymatic evidence for acetogenesis among multiple lineages of the archaeal phylum Bathyarchaeota widespread in marine sediments. *Nat Microbiol* 2016;**1**:16035.

Hélène A, Mylène H, Christine D *et al.* C3 group: a rare but active Thaumarchaeal group in intertidal muddy sediment. In: *Aquatic Sciences Meeting*, Granada, Spain. Wisconsin, USA: Association for the Sciences of Limnology & Oceanography, 2015.

Hinrichs K, Boetius A. The anaerobic oxidation of methane: new insights in microbial ecology and biogeochemistry. In: Wefer GBD, Hebbeln D, Jørgensen BB *et al.* (eds) *Ocean Margin Systems*. Verlag Berlin Heidelberg: Springer, 2002, 457–77.

Inagaki F, Suzuki M, Takai K *et al.* Microbial communities associated with geological horizons in coastal subseafloor sediments from the Sea of Okhotsk. *Appl Environ Microbiol* 2003;**69**:7224–35.

Inagaki F, Nunoura T, Nakagawa S *et al.* Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin. *Proc Natl Acad Sci USA* 2006;**103**:2815–20.

Jacquemet A, Barbeau J, Lemiegre L *et al.* Archaeal tetraether bipolar lipids: Structures, functions and applications. *Biochimie* 2009;**91**:711–7.

Jiang L, Zheng Y, Chen J *et al.* Stratification of Archaeal communities in shallow sediments of the Pearl River Estuary, Southern China. *Antonie Van Leeuwenhoek* 2011;**99**:739–51.

Kallmeyer J, Pockalny R, Adhikari RR *et al.* Global distribution of microbial abundance and biomass in subseafloor sediment. *Proc Natl Acad Sci USA* 2012;**109**:16213–6.

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

Kellermann MY, Wegener G, Elvert M *et al.* Autotrophy as a predominant mode of carbon fixation in anaerobic methane-oxidizing microbial communities. *Proc Natl Acad Sci USA* 2012;**109**:19321–6.

Knittel K, Boetius A. Anaerobic oxidation of methane: progress with an unknown process. *Annu Rev Microbiol* 2009;**63**:311–34.

Kubo K, Lloyd KG, Biddle JF *et al.* Archaea of the Miscellaneous Crenarchaeotal Group are abundant, diverse and widespread in marine sediments. *ISME J* 2012;**6**:1949–65.

Kubota K. CARD-FISH for environmental microorganisms: technical advancement and future applications. *Microb Environ* 2013;**28**:3–12.

LaRowe D, Amend J. Energetic constraints on life in marine deep sediments. In: Kallmeyer J, Wagner D (eds) *Microbial Life of the Deep Biosphere* Vol. **1**. 2014, 279–302.

Laso-Pérez R, Wegener G, Knittel K *et al.* Thermophilic archaea activate butane via alkyl-coenzyme M formation. *Nature* 2016;**539**:396–401.

Lazar CS, Biddle JF, Meador TB *et al.* 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* 2015;**17**:2228–38.

Lazar CS, Baker BJ, Seitz K *et al.* Genomic evidence for distinct carbon substrate preferences and ecological niches of Bathyarchaeota in estuarine sediments. *Environ Microbiol* 2016;**18**:1200–11.

Lever MA. A new era of methanogenesis research. *Trends Microbiol* 2016;**24**:84–6.

Li M, Baker BJ, Anantharaman K *et al.* Genomic and transcriptomic evidence for scavenging of diverse organic compounds by widespread deep-sea archaea. *Nat Commun* 2015;**6**:8933.

Li PY, Xie BB, Zhang XY *et al.* Genetic structure of three fosmid-fragments encoding 16S rRNA genes of the Miscellaneous Crenarchaeotic Group (MCG): implications for physiology and evolution of marine sedimentary archaea. *Environ Microbiol* 2012;**14**:467–79.

Lipp JS, Liu X, Hinrichs K-U. Structural diversity and fate of intact polar lipids in marine sediments. *Geochim Cosmochim Acta* 2009;**73**:6816–33.

Lipp JS, Morono Y, Inagaki F *et al*. Significant contribution of Archaea to extant biomass in marine subsurface sediments. *Nature* 2008;**454**:991–4.

Liu J, Yang H, Zhao M *et al*. Spatial distribution patterns of benthic microbial communities along the Pearl Estuary, China. *Syst Appl Microbiol* 2014;**37**:578–89.

Lloyd K. Beyond known methanogens. *Science* 2015;**350**:384.

Lloyd KG, Schreiber L, Petersen DG *et al*. Predominant archaea in marine sediments degrade detrital proteins. *Nature* 2013;**496**:215–8.

Logares R, Brate J, Bertilsson S *et al*. Infrequent marine-freshwater transitions in the microbial world. *Trends Microbiol* 2009;**17**:414–22.

Lomstein BA, Langerhuus AT, D'Hondt S *et al*. Endospore abundance, microbial growth and necromass turnover in deep sub-seafloor sediment. *Nature* 2012;**484**:101–4.

Lozupone CA, Knight R. Global patterns in bacterial diversity. *Proc Natl Acad Sci USA* 2007;**104**:11436–40.

Ludwig W, Strunk O, Westram R *et al*. ARB: a software environment for sequence data. *Nucleic Acids Res* 2004;**32**: 1363–71.

Martin W, Muller M. The hydrogen hypothesis for the first eukaryote. *Nature* 1998;**392**:37–41.

Martin WF, Neukirchen S, Zimorski V *et al*. Energy for two: New archaeal lineages and the origin of mitochondria. *Bioessays* 2016;**38**:850–6.

Meador TB, Bowles M, Lazar CS *et al*. The archaeal lipidome in estuarine sediment dominated by members of the Miscellaneous Crenarchaeotal Group. *Environ Microbiol* 2015;**17**:2441–58.

Meng J, Wang F, Wang F *et al*. An uncultivated Crenarchaeota contains functional bacteriochlorophyll a synthase. *ISME J* 2009;**3**:106–16.

Meng J, Xu J, Qin D *et al*. Genetic and functional properties of uncultivated MCG archaea assessed by metagenome and gene expression analyses. *ISME J* 2014;**8**:650–9.

Miller MA, Pfeiffer W, Schwartz T. *Creating the CIPRES Science Gateway for inference of large phylogenetic trees*. Gateway Computing Environments Workshop (GCE), 2010. New Orleans, Louisiana, USA: IEEE, 2010, 1–8.

Na H, Lever MA, Kjeldsen KU *et al*. Uncultured Desulfobacteraceae and Crenarchaeotal group C3 incorporate ^{13}C-acetate in coastal marine sediment. *Environ Microbiol Rep* 2015;**7**:614–22.

Newberry CJ, Webster G, Cragg BA *et al*. Diversity of prokaryotes and methanogenesis in deep subsurface sediments from the Nankai Trough, Ocean Drilling Program Leg 190. *Environ Microbiol* 2004;**6**:274–87.

Parkes RJ, Webster G, Cragg BA *et al*. Deep sub-seafloor prokaryotes stimulated at interfaces over geological time. *Nature* 2005;**436**:390–4.

Pruesse E, Peplies J, Gloeckner FO. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 2012;**28**:1823–9.

Reed DW, Fujita Y, Delwiche ME *et al*. Microbial communities from methane hydrate-bearing deep marine sediments in a forearc basin. *Appl Environ Microbiol* 2002;**68**:3759–70.

Rossel PE, Lipp JS, Fredricks HF *et al*. Intact polar lipids of anaerobic methanotrophic archaea and associated bacteria. *Org Geochem* 2008;**39**:992–9.

Schouten S, Hopmans EC, Damste JSS. The organic geochemistry of glycerol dialkyl glycerol tetraether lipids: A review. *Org Geochem* 2013;**54**:19–61.

Seyler LM, McGuinness LM, Kerkhof LJ. Crenarchaeal heterotrophy in salt marsh sediments. *ISME J* 2014;**8**:1534–43.

Sorensen KB, Teske A. Stratified communities of active archaea in deep marine subsurface sediments. *Appl Environ Microbiol* 2006;**72**:4596–603.

Sousa FL, Neukirchen S, Allen JF *et al*. *Lokiarchaeon* is hydrogen dependent. *Nat Microbiol* 2016;**1**:16034.

Spang A, Caceres EF, Ettema TJG. Genomic exploration of the diversity, ecology, and evolution of the archaeal domain of life. *Science* 2017;**357**:eaaf3883.

Stahl DA, Flesher B, Mansfield HR *et al*. Use of phylogenetically based hybridization probes for studies of ruminal microbial ecology. *Appl Environ Microbiol* 1988;**54**:1079–84.

Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014;**30**:1312–3.

Sturt HF, Summons RE, Smith K *et al*. Intact polar membrane lipids in prokaryotes and sediments deciphered by high-performance liquid chromatography/electrospray ionization multistage mass spectrometry—new biomarkers for biogeochemistry and microbial ecology. *Rapid Commun Mass Spectrom* 2004;**18**:617–28.

Summons RE, Franzmann PD, Nichols PD. Carbon isotopic fractionation associated with methylotrophic methanogenesis. *Org Geochem* 1998;**28**:465–75.

Takai K, Moser DP, DeFlaun M *et al*. Archaeal diversity in waters from deep South African gold mines. *Appl Environ Microbiol* 2001;**67**:5750–60.

Teske A, Sørensen KB. Uncultured archaea in deep marine subsurface sediments: have we caught them all? *ISME J* 2008;**2**:3–18.

Teske AP. Microbial communities of deep marine subsurface sediments: molecular and cultivation surveys. *Geomicrobiol J* 2006;**23**:357–68.

Thauer RK, Kaster A-K, Seedorf H *et al*. Methanogenic archaea: ecologically relevant differences in energy conservation. *Nat Rev Microbiol* 2008;**6**:579–91.

Vanwonterghem I, Evans PN, Parks DH *et al*. Methylotrophic methanogenesis discovered in the archaeal phylum *Verstraetearchaeota*. *Nat Microbiol* 2016;**1**:16170.

Wang F-P, Zhang Y, Chen Y *et al*. Methanotrophic archaea possessing diverging methane-oxidizing and electron-transporting pathways. *ISME J* 2014;**8**:1069–78.

Weber J, Kayser A, Rinas U. Metabolic flux analysis of *Escherichia coli* in glucose-limited continuous culture. II. Dynamic response to famine and feast, activation of the methylglyoxal pathway and oscillatory behaviour. *Microbiology* 2005;**151**:707–16.

Webster G, Parkes RJ, Cragg BA *et al*. Prokaryotic community composition and biogeochemical processes in deep sub-seafloor sediments from the Peru Margin. *FEMS Microbiol Ecol* 2006;**58**:65–85.

Webster G, Rinna J, Roussel EG *et al*. Prokaryotic functional diversity in different biogeochemical depth zones in tidal sediments of ?the Severn Estuary, UK, revealed by stable-isotope probing. *FEMS Microbiol Ecol* 2010;**72**:179–97.

Webster G, Sass H, Cragg BA *et al*. Enrichment and cultivation of prokaryotes associated with the sulphate-methane transition zone of diffusion-controlled sediments of Aarhus Bay, Denmark, under heterotrophic conditions. *FEMS Microbiol Ecol* 2011;**77**:248–63.

Weiss MC, Sousa FL, Mrnjavac N *et al*. The physiology and habitat of the last universal common ancestor. *Nat Microbiol* 2016;**1**:16116.

Xiang X, Wang R, Wang H *et al*. Distribution of Bathyarchaeota communities across different terrestrial settings and their potential ecological functions. *Sci Rep* 2017;**7**: 45028.

Yarza P, Yilmaz P, Pruesse E *et al*. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol* 2014;**12**:635–45.

Yoon SH, Ha SM, Lim J *et al*. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek* 2017;**110**:1281–6.

Yu T, Liang Q, Niu M *et al*. High occurrence of Bathyarchaeota (MCG) in the deep-sea sediments of South China Sea quantified using newly designed PCR primers. *Environ Microbiol Rep* 2017;**9**:374–82.

Yu T, Wu W, Liang W *et al*. Growth of sedimentary Bathyarchaeota on lignin as an energy source. *Proc Natl Acad Sci USA* 2018;**115**:6022–7.

Zhang W, Ding W, Yang B *et al*. Genomic and transcriptomic evidence for carbohydrate consumption among microorganisms in a cold seep brine pool. *Front Microbiol* 2016;**7**:1825.