MICROBIOLOGY

Taxonomic and carbon metabolic diversification of Bathyarchaeia during its coevolution history with early Earth surface environment

Jialin Hou[1], Yinzhao Wang[1]*, Pengfei Zhu[2], Na Yang[2], Lewen Liang[1], Tiantian Yu[1], Mingyang Niu[1], Kurt Konhauser[3], Ben J. Woodcroft[4], Fengping Wang[1,2,5]*

Bathyarchaeia, as one of the most abundant microorganisms on Earth, play vital roles in the global carbon cycle. However, our understanding of their origin, evolution, and ecological functions remains poorly constrained. Here, we present the largest dataset of Bathyarchaeia metagenome assembled genome to date and reclassify Bathyarchaeia into eight order-level units corresponding to the former subgroup system. Highly diversified and versatile carbon metabolisms were found among different orders, particularly atypical C1 metabolic pathways, indicating that Bathyarchaeia represent overlooked important methylotrophs. Molecular dating results indicate that Bathyarchaeia diverged at ~3.3 billion years, followed by three major diversifications at ~3.0, ~2.5, and ~1.8 to 1.7 billion years, likely driven by continental emergence, growth, and intensive submarine volcanism, respectively. The lignin-degrading Bathyarchaeia clade emerged at ~300 million years perhaps contributed to the sharply decreased carbon sequestration rate during the Late Carboniferous period. The evolutionary history of Bathyarchaeia potentially has been shaped by geological forces, which, in turn, affected Earth's surface environment.

INTRODUCTION

Bathyarchaeia, previously known as the miscellaneous Crenarchaeotal group (MCG) (*1*) or Bathyarchaeota (*2*), are ubiquitously distributed dominating archaeal communities in anoxic subsurface environments, including oceanic (*3*), freshwater (*4*), hydrothermal (*5*), hot spring sediment (*6*), and soil (*7*), particularly coastal settings (*8*). Globally, it is estimated that Bathyarchaeia account for ~2.0 to 3.9 × 10^{28} cells (*5*), ranking as one of the most abundant microbial groups on Earth. However, because there is still no laboratory cultivation of Bathyarchaeia thus far, their taxonomy was defined by the identities of 16*S* ribosomal RNA (rRNA) genes retrieved directly from environmental samples and metagenome-assembled genomes (MAGs). This has resulted in notable phylogenetic inconsistencies and instability. For example, Bathyarchaeia were initially classified into 17 subgroups (*8*) and further extended to 25 subgroups with accumulating sequences (*4*, *9*). Recently, the Genome Taxonomy Database (GTDB) archaeal taxonomy reclassified the phylum Bathyarchaeota to a class-level clade and renamed it as the class Bathyarchaeia within the phylum Thermoproteota, corresponding to the previous Thaumarchaeota-Aigarchaeota-Crenarchaeota-Korarchaeota (TACK) superphylum (*10*). However, the correspondence between the former subgroup system and the detailed taxonomy remains unknown, thus impeding our understanding of the evolution and ecological roles of different Bathyarchaeia lineages.

Bathyarchaeia have been suggested to have diversified capabilities in metabolizing a large spectrum of organic compounds to fulfill their carbon or energy demands, including peptides/proteins (*3*), polymeric carbohydrates (*11*), fatty acids (*12*), lignin (*13*), aromatic (*2*), and aldehyde compounds (*14*). At the same time, isotopic and enzymatic experiments have demonstrated their ability to assimilate inorganic carbon (*5*, *13*). Furthermore, recent findings imply that some Bathyarchaeia species might also undergo anaerobic methane or multicarbon alkane metabolism (*15*, *16*). Overall, given their numerical dominance and metabolic versatility in the anoxic deep biosphere, different Bathyarchaeia lineages are believed to be critical players in modern carbon cycling. Nevertheless, when and how Bathyarchaeia originated on early Earth and subsequently evolved into such a diversified and now flourishing branch of life remains unclear.

Recently, the evolution of several archaea with special ecological functions—such as ammonia-oxidizing Thaumarchaeota (*17*), methanogens (*18*), and methanotrophic archaea (*19*)—has been elucidated by molecular dating analysis. To demonstrate the metabolic characteristics and evolutionary history of Bathyarchaeia, we propose here an improved taxonomy derived from an exhaustive collection of 304 nonredundant representative MAGs refined from available public databases and our own datasets based on the GTDB criterion (*10*). This approach enabled a comprehensive investigation of the distinctive environmental distribution and taxonomy-specific carbon metabolic traits of different Bathyarchaeia lineages. Their origin and divergence time were inferred from phylogenomic and molecular dating analysis, reflecting the multiple independent correlations between the evolution of Bathyarchaeia and key geological processes that occurred in early Earth's history, including subaerial continental emergence, growth, and episodes of intensive submarine volcanism. In addition, we propose that the

[1]State Key Laboratory of Microbial Metabolism, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China. [2]School of Oceanography, Shanghai Jiao Tong University, Shanghai, China. [3]Department of Earth and Atmospheric Sciences, University of Alberta, Edmonton, Alberta, Canada. [4]Centre for Microbiome Research, School of Biomedical Sciences, Queensland University of Technology (QUT), Translational Research Institute, Woolloongabba, Australia. [5]Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai, China.
*Corresponding author. Email: fengpingw@sjtu.edu.cn (F.W.); wyz@sjtu.edu.cn (Y. W.)

origin of lignin-degrading Bathyarchaeia clade probably contributed to the contemporaneous sharp decrease in coal deposition rate during the Late Carboniferous period.

RESULTS AND DISCUSSION

Improved Bathyarchaeia taxonomy and subgroup assignment

A total of 304 nonredundant representative Bathyarchaeia MAGs identified from available public genome databases, including the National Center for Biotechnology Information (NCBI), the genomic catalog of Earth′s microbiome (GEM) (*20*), and GTDB (*10*), as well as our sequenced datasets (table S1), were used for further phylogenomic analysis (table S2). A total of 78 (25.7%) and 187 (61.5%) MAGs had high (completeness of ≥90% and contamination of <5%) and medium (completeness of ≥50% and contamination of <10%) genomic quality, respectively (table S3). On the basis of the phylogenomic inference, we propose a systematic taxonomy for each Bathyarchaeia representative MAG, whose taxonomic rank was hierarchically assigned and normalized by the approach implemented in GTDB (*21*). First, the calculated relative evolutionary divergence (RED) value (0.339) for the whole Bathyarchaeia lineage indicates it as an independent class-level taxonomic unit within the phylum Thermoproteota (previously known as TACK superphylum), in accordance with the archaeal GTDB taxonomy (06-RS202) (fig. S1 and table S4). Furthermore, the 304 representative MAGs were classified into eight order-level lineages and corresponding subsequent family and genus ranks (tables S3 and S4).

The eight Bathyarchaeia orders were named after different Chinese traditional mythological figures, denoting their distinct environmental features. These include Bifangarchaeales (Bifang, a single-legged bird as the symbol of wildfire; this lineage largely comes from hot springs), Jinwuousiales (Jinwu, a three-legged bird living in the Sun; their members are mostly from hydrothermal vents), Zhuquarculales (Zhuque, a red bird as the god of fire; all the species are from hydrothermal vents), Wuzhiqibiales (Wuzhiqi, the god of Huai River; most comes from marine sediments, particularly estuaries), Houtuarculales (Houtu, the god of Earth; the large parts are found from soils), Mazuousiales (the sea goddess, mostly comes from East China Sea sediments), Baizomonadales (Baize, a propitious creature living in cold Kunlun Mountains; this lineage has miscellaneous distribution but largely from cold marine sediments), and Xuanwuarculales (Xuanwu, a black turtle as the god of iciness and water; only has one MAG from underground water sediment) (Fig. 1 and table S5). Despite using the same RED intervals for taxonomic rank classification, the newly improved taxonomy of Bathyarchaeia has more designated families and genera for most orders compared with the reference GTDB taxonomy (06-RS202) (Table 1). Notably, the order Mazuousiales is a newly found Bathyarchaeia order proposed in this study, in which five of their seven MAGs are additionally retrieved from the GEM and our datasets (table S3). Consequently, these taxonomic differences can be attributed to the largely increased Bathyarchaeia MAGs analyzed here (304 MAGs) in comparison to the reference GTDB database (173 MAGs, 06-RS202) (table S2).

Previously, Bathyarchaeia were classified into 25 subgroups based on the phylogeny of 16*S* rRNA gene. However, subgroup is not a legitimate rank name and normally subjectively defined without a consistent phylogenetic threshold (*4*, *8*, *9*), e.g., the minimum intragroup identities of different subgroups range from 84 to 94% in the initial designation (*8*) and subsequently were updated to >90% (*9*), which generally fall into the universal taxonomic threshold range between the order and family level (89.2 to 92.25%) (*22*). As a consequence, the specific taxonomic rank for each subgroup and whether they are equivalent remain unresolved. To address this, a phylogenetic tree was constructed using the 16*S* rRNA genes retrieved from the representative Bathyarchaeia MAGs and subgroup-classified reference sequences (*9*) to establish links between each Bathyarchaeia subgroup and their corresponding taxonomic ranks (fig. S2). The result shows that 16*S* rRNA genes from 16 subgroups were identified in at least one MAG of seven Bathyarchaeia orders (except for the order Zhuquarculales as no 16*S* rRNA gene was retrieved from their MAGs), and most subgroups are assigned to remarkably distinct taxonomic levels (Fig. 1B). For example, the largest order Baizomonadales taxonomically corresponds to at least 10 subgroups, including subgroups 8 and 6, which dominate Bathyarchaeia lineages in global marine and freshwater sediments (*4*, *23*) and are assigned to the genus *Baizosediminiarchaeum* and *Baizomonas*, respectively (table S3). In contrast, subgroup 17 contains the MAGs from both orders Xuanwuarculales and Mazuousiales. Together, our results demonstrate that the previously widely used Bathyarchaeia subgroups are far from taxonomic equivalents. Instead, they correspond to remarkably distinct ranks in the improved taxonomic framework, ranging from the above order level to the below genus level. Accordingly, cautions should be applied when extrapolating the significance of different Bathyarchaeia subgroups as they may not be suitable for quantitative comparisons either among themselves or with other microbial lineages.

Taxonomically specific distribution in terrestrial and geothermal environments

For the 304 representative Bathyarchaeia MAGs, more than 75% of them (228) were retrieved from diverse oceanic and terrestrial sedimentary environments (Fig. 2A). This is consistent with an investigation of Bathyarchaeia abundance across global environmental metagenomes undertaken here (Fig. 2C and table S6). Nearly 50% of the MAGs (150) were isolated from oceanic environments, of which 91% were recovered from marine sediments that are associated with the seafloor (70) and hydrothermal systems (67), while the other 50% of MAGs are from diverse terrestrial ecosystems, mainly including soils (66), hot springs (66), and freshwater sediments (17). Notably, some Bathyarchaeia orders show taxonomically specific distribution in different terrestrial environments. For example, ~61% (30 of 49) of the MAGs assigned to the order Houtuarculales were recovered from soils, especially the genus *Houtuousia* (all of 23) (Fig. 2B and table S3). The metagenome-based global survey further demonstrates that Houtuarculales is the dominating Bathyarchaeia order in soil environments (Fig. 2C). Therefore, this lineage may already have metabolically adapted to the specific soil ecosystem, which expands the range of ecological impacts and significance of the class Bathyarchaeia from oceanic to terrestrial environments. The sole MAG assigned to the order Xuanwuarculales was recovered from groundwater sediment (Fig. 2B) (*24*), indicating that the deep terrestrial biosphere likely harbors other undiscovered Bathyarchaeia lineages.

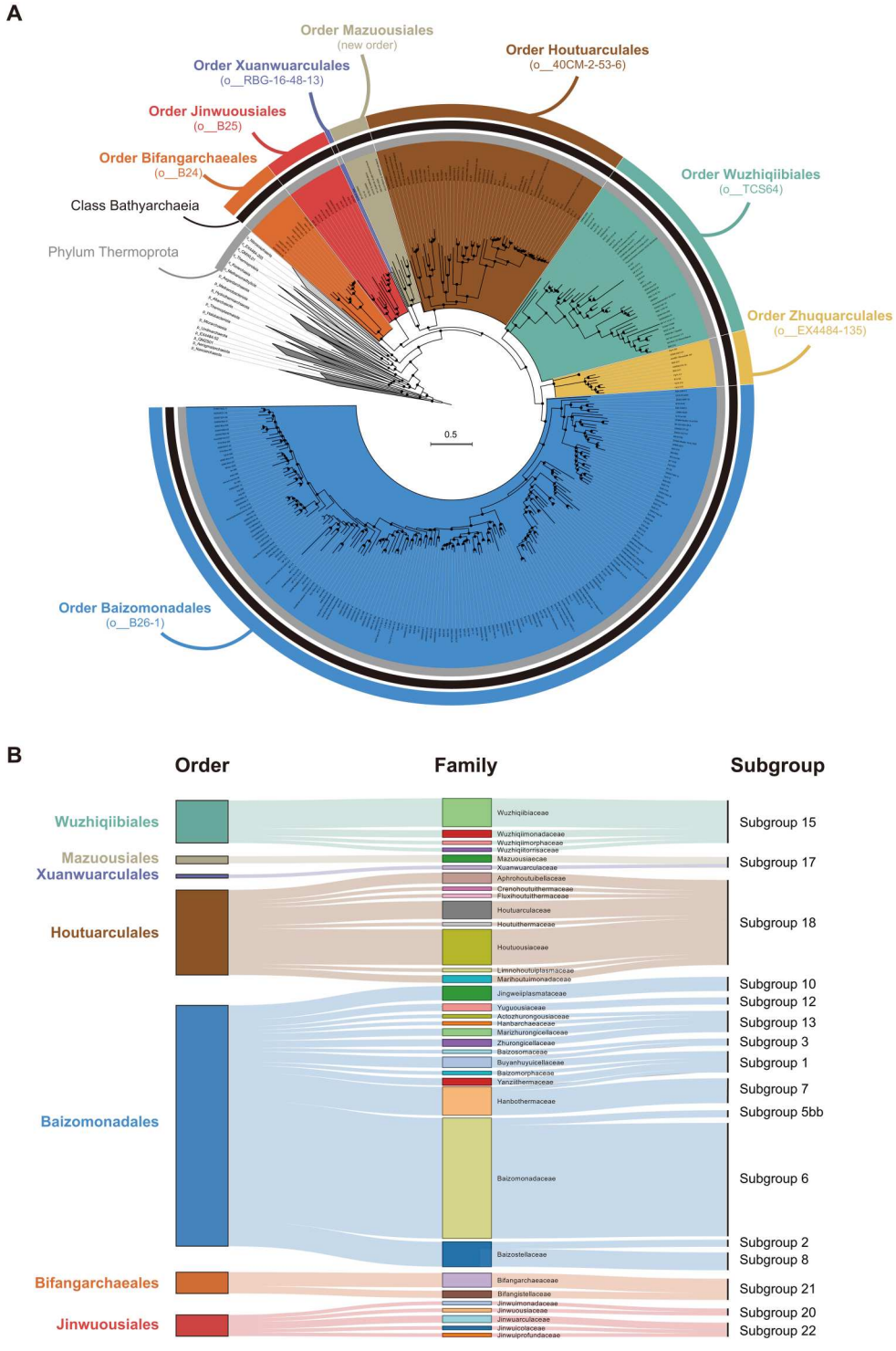

Fig. 1. The newly improved taxonomy and subgroup assignment of the class Bathyarchaeia. (**A**) Phylogenomic affiliation of 304 representative Bathyarchaeia MAGs refined from publicly available databases and our laboratory datasets based on a concatenated alignment of 122-archaeal marker proteins implemented in GTDB-Tk by using IQ-TREE2 with LG+F+R10+C60 model and Shimodaira Hasegawa–like approximate likelihood ratio test with 1000 bootstrap replicates (bootstrap higher than 0.9 are shown with black dots). The colored background and outer rings denote the eight proposed orders of class Bathyarchaeia, and the inner black and gray rings represent the whole class Bathyarchaeia and phylum Thermoproteota, respectively. The number of MAGs used in the phylogenetic analysis is listed under the name for each taxonomic lineage. (**B**) Taxonomic assignment of subgroups for the representative Bathyarchaeia MAGs with 16*S* rRNA genes. The phylogenetic tree was constructed with subgroup-classified sequences from Zhou *et al.* (*9*) as a reference by using RaxML 8.2.12 (*84*) with -m GTRGAMMA -N autoMRE and 1000 bootstrap replicates (fig. S2).

Table 1. The newly improved taxonomy of the class Bathyarchaeia in this study and the difference from the reference GTDB taxonomy (Release 06-RS202). Numbers in brackets represent the number of taxonomic units in GTDB Release 06-RS202.

Proposed orders	GTDB orders	No. of family	No. of genus	No. of MAG	Type genus	Putative nomenclatural type
Baizomonadales	o__B26-1	19(6)	24(33)	175(98)	*Baizomonas*	*Baizomonas nivis*
Bifangarchaeales	o__B24	3(2)	3(2)	11(10)	*Bifangarchaeales*	*Bifangarchaeales caldum*
Jinwuousiales	o__B25	4(1)	5(3)	12(10)	*Jinwuousia*	*Jinwuousisles tripozesto*
Zhuquarculales	o__EX4484-135	2(2)	2(2)	10(3)	*Zhuquarcula*	*Zhuquarcula teneretur*
Wuzhiqibiales	o__TCS64	5(1)	5(6)	39(10)	*Wuzhiqiibium*	*Wuzhiqiibium servoceani*
Houtuarculales	o__40CM-2-53-6	9(4)	10(9)	49(41)	*Houtuarcula*	*Houtuarcula pratum*
Mazuousiales	No	3	5	7	*Mazuarcula*	*Mazuousia monanthrakas*
Xuanwuarculales	o__RBG_16_48_13	1(1)	1(1)	1(1)	*Xuanwuarcula*	*Xuanwuarcula inferioraqua*

Except for the predominant Bathyarchaeia groups found in the temperate coastal zone or cold deep subsurface (*3*, *8*, *9*), diverse Bathyarchaeia species have also been found in different geothermal environments (*5*, *6*). Here, a total of 143 (~47%) representative Bathyarchaeia MAGs were retrieved from hot springs (66), hydrothermal sediments (67), crustal fluids (5), and deep formation waters (5) (Fig. 2A), along with notable taxonomic specificity at the order level. For instance, all MAGs of the orders Bifangarchaeales, Zhuquarculales, and Jinwuousiales (except one from saline water) were specifically recovered from geothermal environments (Fig. 2B and table S3), consistent with the previous subgroup-based studies (*5*, *6*, *9*, *25*). In particular, 8 of 11 MAGs from the order Bifangarchaeales were exclusively recovered from terrestrial hot springs, whereas the other two orders only contain the MAGs from hydrothermal sediments (Fig. 2B), which is also supported by the global survey results (Fig. 2C). Collectively, these results indicate that diversified Bathyarchaeia lineages thrive in high-temperature environments and play important roles in global geothermal ecosystems. All members from the orders Bifangarchaeales, Jinwuousiales, and Zhuquarculales are very likely obligate thermophilic archaea that evolved in high-temperature environments, in accordance with previous studies about the thermal origin and adaptation of Bathyarchaeia (*6*, *25*).

Diversified metabolic potentials for complex organic carbon utilization

Metabolic reconstruction of 86 high-quality Bathyarchaeia MAGs indicated that almost all eight proposed orders have relatively similar and complete central carbon metabolisms, including glycolysis, gluconeogenesis, reverse ribulose monophosphate, phosphoribosyl pyrophosphate biosynthesis, and pyruvate-ferredoxin oxidoreductase (PFOR complex), in addition to complete gene sets coding for the tetrahydromethanopterin (H_4-MPT) Wood-Ljungdahl pathway (WLP) (Fig. 3A and table S7). Furthermore, complete or near complete PFOR complex and H_4-MPT WLP were inferred to present in the common ancestors for each of the seven Bathyarchaeia orders (fig. S3 and table S8). In addition, the genes encoding carbohydrate-active enzymes (CAZymes) and peptidases were extensively predicted in the 86 MAGs, along with significantly different average numbers (per 100 genes per genome) among the eight orders (Fig. 3, B and C). All these core metabolic repertoires reflect the versatile capabilities of Bathyarchaeia in using different organic compounds. Consistent with this inference, previous studies have demonstrated that different Bathyarchaeia species can assimilate a variety of organic carbon compounds, such as acetate (*12*), amino acids (*26*), fatty acids, and complex hydrocarbons (*11*). Some Bathyarchaeia members are also proposed to be capable of using lignin as an energy source for acetogenesis through the reductive acetyl–coenzyme A (CoA) pathway (also reverse H_4-MPT WLP) (*5*, *13*).

In addition to the common central carbon metabolism, different Bathyarchaeia orders show distinct metabolic preferences in metabolizing complex organic compounds. Among all Bathyarchaeia orders, the hot spring order Bifangarchaeales has the highest average number of CAZyme genes per genome in their MAGs (Fig. 3B), particularly those associated with decomposing biomass [e.g., GH109 (*27*)] and plant cell wall–derived polysaccharides [e.g., AA4 and AA6 (*28*)] (Fig. 3A). Moreover, the key genes for fructose and xylose degradation were also predicted in seven of their eight MAGs but rarely detected in others (Fig. 3A), which were also inferred to be encoded in the common ancestor of this order with high possibility (fig. S3 and table S8). Therefore, these results indicate that the members of the order Bifangarchaeales likely had the capacity to use a large spectrum of carbohydrates, especially those derived from vascular plants, for specifically adapting to hot spring environments, which are usually enriched with plant and microbial biomass (*29*). Moreover, we demonstrate that this ability is specific to the order Bifangarchaeales rather than a common metabolic trait for all thermophilic Bathyarchaeia species at high temperatures as proposed before (*6*).

For the orders Jinwuousiales and Zhuquarculales, these two hydrothermal lineages have significantly smaller average genome sizes, lower coding densities, and fewer gene numbers than the other orders, including the hot spring order Bifangarchaeales (fig. S4). It means that remarkable genome streamlining might have occurred in these hydrothermal Bathyarchaeia lineages, possibly as an adaptative consequence of the extreme conditions associated with hydrothermal activity and scarce external organic inputs in deep-sea environments. In terms of metabolism, four Jinwuousiales MAGs encode the phosphate acetyltransferase (Pta), acetate kinase (Ack), and/or acetyl-CoA synthetase (ACSS) coding genes for acetogenesis/acetate utilization (Fig. 3A), in which ACSS was also inferred to be present in the common ancestor of order Jinwuousiales (fig. S3 and table S8), thus implying that acetate is an important

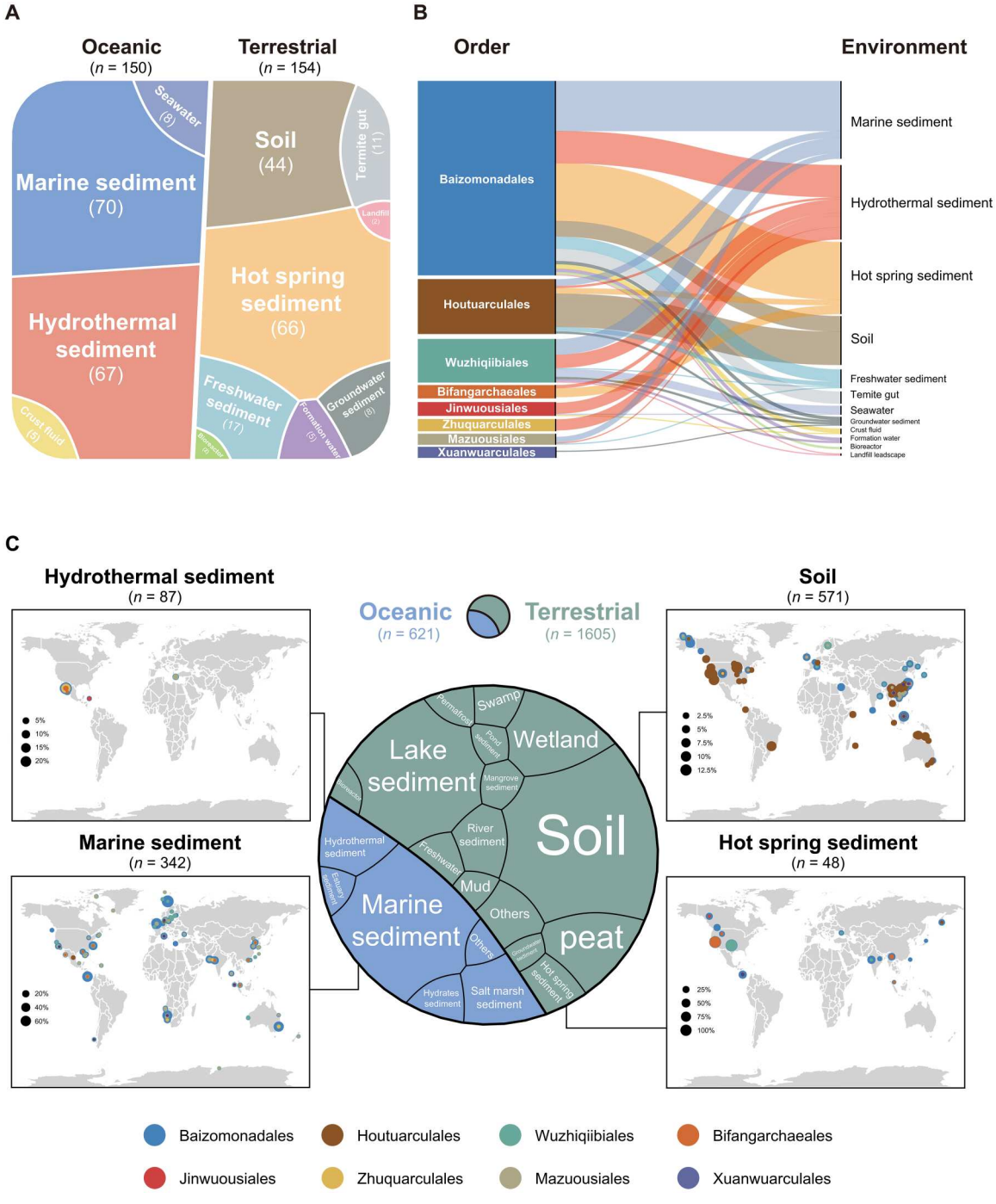

Fig. 2. Environmental and geographic distribution of the class Bathyarchaeia. (**A**) Distribution of 304 representative Bathyarchaeia MAGs across different oceanic and terrestrial environments. The number of MAGs associated with each environment is indicated below the name. (**B**) Environmental distribution of Bathyarchaeia MAGs from the eight proposed orders. (**C**) Environmental distribution of a total of 2226 metagenomic samples with Bathyarchaeia (>1%) from diverse marine and terrestrial ecosystems based on Sandpiper v0.0.23 (https://sandpiper.qut.edu.au/). Geographic distribution of the metagenomes from hydrothermal sediment, soil, marine sediment, and hot spring sediment. Relative abundance of the eight Bathyarchaeia orders within each sample is indicated with the circle size.

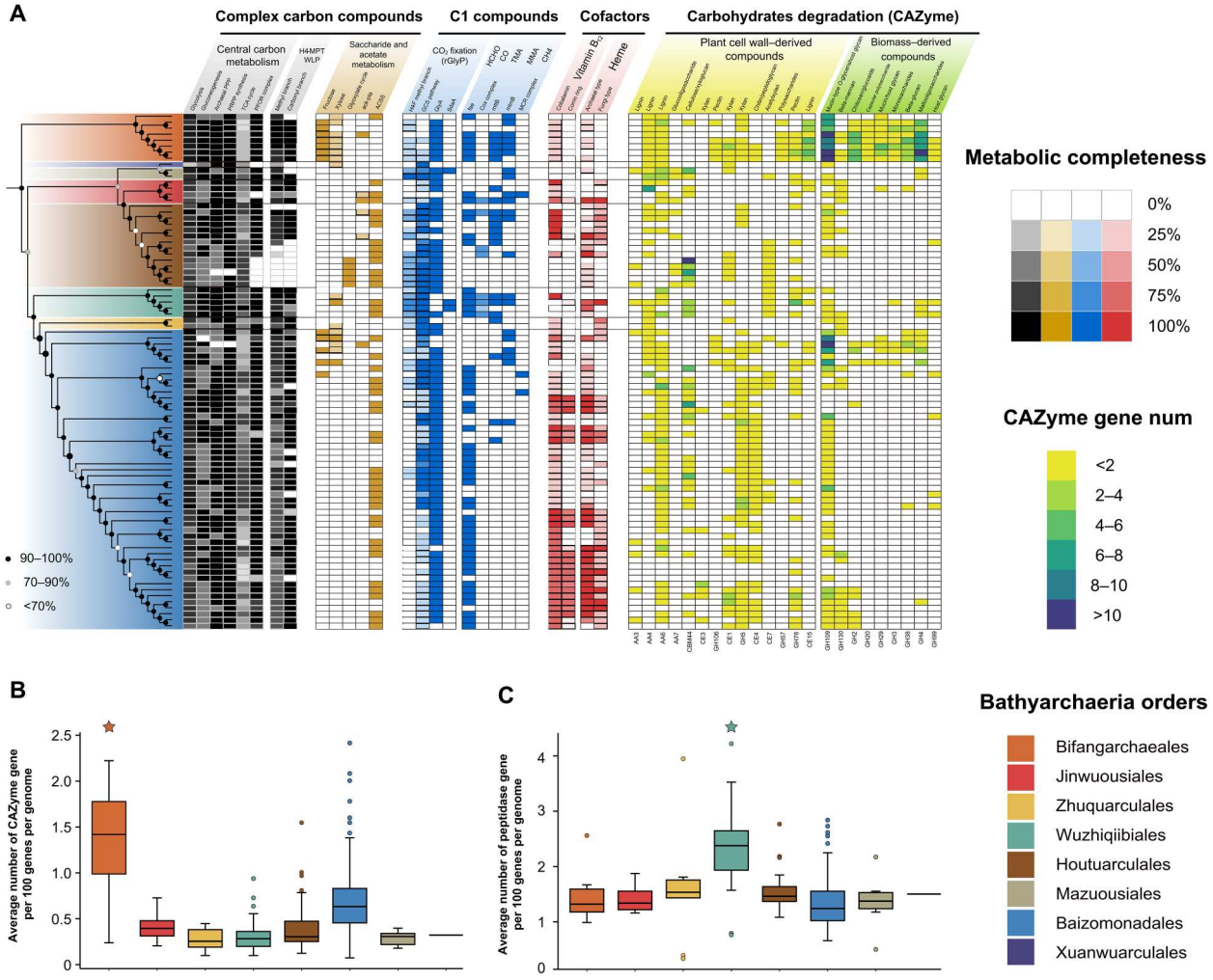

Fig. 3. Distinctive metabolic traits of the class Bathyarchaeia. (**A**) Phylogenomic tree includes 86 high-quality representative Bathyarchaeia MAGs and was inferred by iQ-TREE2 with LG+F+R7+C20 model from 122 archaeal marker proteins implemented in GTDB-Tk (see Materials and Methods). Colors on branches indicate eight proposed Bathyarchaeia orders. The completeness of each metabolic pathway is defined by the ratio of marker genes identified in the complete gene repertoire for each MAG, while for many key genes or complexes, such as those involved in C1 compound metabolism, their presence and absence in each MAG are directly indicated by the percentage of completeness. Metabolic capabilities of carbohydrate degradation are evaluated by the gene numbers of different CAZymes identified in each MAG. The following metabolic pathways and key enzymes were used in the figure (table S7): archaeal PPP, archaeal pentose phosphate pathway; PRPP synthesis, phosphoribosyl pyrophosphate synthesis; TCA cycle, tricarboxylic acid cycle; PFOR complex, pyruvate:ferredoxin oxidoreductase complex; ack-pta, acetate kinase and phosphate acetyltransferase; ACSS, acetyl-CoA synthetase; rGlyP, reductive glycine pathway; GCS, glycine cleavage system; GlyA, glycine hydroxymethyltransferase; SdaA, L-serine dehydratase; fae, formaldehyde-activating enzyme; Cox complex, aerobic carbon-monoxide dehydrogenase; MttB, trimethylamine-corrinoid protein co-methyltransferase; MtmB, methylamine-corrinoid protein co-methyltransferase; MCR complex, methyl-coenzyme M reductase complex; AA, auxiliary activity; GH, glycoside hydrolase; CE, carbohydrate esterase. (**B** and **C**) Box plots show the average number of CAZyme and peptidase genes per 100 genes in one MAG for each Bathyarchaeia order, respectively. The stars indicate that the order has a significantly higher average genomic gene number in comparison with all other Bathyarchaeia orders (Wilcoxon test with Benjamini-Hochberg correction, $P < 0.05$).

intermediate substrate for this Bathyarchaeia lineage inhabiting hydrothermal sediments. Compared with the order Jinwuousiales, members of the order Zhuquarculales have significantly higher guanine and cytosine (GC) content, lower nitrogen per amino acid residue side chain (N-ARSC) (fig. S4) (*30*), as well as fewer genes for amino acid synthesis and CAZymes (Fig. 3A and fig. S5). Furthermore, remarkable composition bias of amino acid was observed among different Bathyarchaeia orders, in which the MAGs of order Zhuquarculales have significantly less lysine, threonine, phenylalanine, and valine, which correlates well with the general absence of these synthesis pathways in their representative MAGs (figs. S5 and S6). Accordingly, their atypical inabilities in terms of both anabolism and catabolism indicate that the order Zhuquarculales might physiologically depend on other organisms for acquiring bioavailable nitrogen sources and other essential nutrients to survive in extreme hydrothermal environments. In addition, the positive correlation between genomic N-ARSC and GC content in the dominating phototrophic bacteria in the surface seawater (*31*, *32*) no longer holds among different Bathyarchaeia orders. It implies that the variation of GC content has discrepant implications

to the nitrogen use efficiency in this archaeal lineage probably due to the remarkably distinct physiological characteristics and environmental distribution between phototrophic bacteria and Bathyarchaeia.

More than 72% of the MAGs in the order Houtuarculales are from nongeothermal terrestrial environments (Fig. 2B). Intriguingly, all five MAGs of the genus *Houtuousia*, which were exclusively recovered from grassland soils (table S3), uniquely encode the two key genes in the glyoxylate cycle, isocitrate lyase, and malate synthase, in addition to the relatively complete tricarboxylic acid cycle and ACSS-coding gene (Fig. 3A), which allows them to directly assimilate acetate as a potential carbon source (*33*). Although the glyoxylate cycle has been identified in many bacterial and fungal species, only a few Haloarchaea have been demonstrated to operate this pathway to assimilate acetate (*34*), suggesting that it might facilitate their adaptation to high-salinity environments. Here, the glyoxylate cycle found in these soil Bathyarchaeia MAGs indicates that the unusual acetate assimilation pathway is not restricted to halophilic archaea, thus expanding its ecological significance from the extreme hypersaline environment to the broader soil environment. Another interesting finding for the genus *Houtuousia* is the absence of the H_4-MPT WLP and PFOR complex in their MAGs, although these two highly oxygen-sensitive pathways were found in all other Bathyarchaeia MAGs as core metabolic traits (Fig. 3A). Consistent with this finding, ancestral reconstruction results suggest that the common ancestor of genus *Houtuousia* also lacks the H_4-MPT WLP and PFOR complex (fig. S3 and table S8). In addition, their MAGs exclusively encode the cytochrome C oxidases putatively for aerobic respiration (Fig. 3A), resembling those in aerobic *Sulfolobus acidocaldarius* (*35*). Therefore, the members of the genus *Houtuousia* probably made physiological adaptations to the (micro)oxic soil environment with a (micro)aerobic lifestyle and atypical acetate assimilation capability.

The order Wuzhiqiibiales solely corresponds to the previous subgroup 15 (Fig. 1B), which frequently predominates the archaeal communities in anoxic coastal sediments with positive correlations with salinity and sedimentary depth (*23*). Here, among all eight Bathyarchaeia orders, the MAGs of Wuzhiqiibiales, on average, encode the highest number of genes for peptidases (~2.3 genes per 100 genes) (Fig. 3C), especially extracellular peptidases (fig. S7 and table S9). This indicates their remarkable potential in metabolizing amino acids, peptides, and detrital proteins. In accordance with our findings, the first single cell–sequenced Bathyarchaeia genome (SCGC AB-539-E09), assigned to the order Wuzhiqiibiales herein, was demonstrated to play active roles in degrading detrital proteins in anoxic sediments (*3*). Recently, ^{13}C-labeled incubation also indicated that some Wuzhiqiibiales species could use extracellular proteins through catabolic processes (*26*). Therefore, using detrital proteins seems to be a metabolic characteristic that is specific to all members of the order Wuzhiqiibiales, reflecting their important roles in protein remineralization and carbon/nitrogen cycling in global marine sediments.

As the most numerous and diversified Bathyarchaeia lineage, 50 MAGs from the order Baizomonadales encode most carbon metabolic pathways identified in other orders (Fig. 3A), and they were retrieved not only from common natural depositional systems but also from a few unique artificial environments, such as bioreactors and landfills (Fig. 2B). This suggests that the order Baizomonadales is now the most representative and successfully evolved Bathyarchaeia lineage. Intriguingly, the anaerobic cobalamin (vitamin B_{12}) synthesis pathway was specifically identified in most members of the genera *Baizomonas* and *Baizosediminiarchaeum* (Fig. 3A and fig. S5). Despite cobalamin being an essential cobalt-containing cofactor for all three domains of life, in the domain Archaea, only a few Thaumarchaeota and euryarchaeotal methanogens could de novo synthesize it, which might facilitate their key roles in seawater and the methane zone in deep sediments, respectively (*36*, *37*). Accordingly, the presence of the cobalamin synthesis pathway in these two genera makes Bathyarchaeia the third archaeal lineage with this unusual capability. This suggests that they could serve as a keystone species in the subseafloor microbial community by providing this essential cofactor for other members. In addition, some *Baizosediminiarchaeum* species could grow with lignin as the energy source by using a special *O*-demethylation methyltransferase (MT) system that require cobalamin-binding corrinoid protein (*13*, *38*), while the growth of certain *Baizomonas* species (assigned to subgroup 6) was substantially stimulated by different lignin-derived aldehydes with methoxy groups in enrichments (*14*). Overall, the cobalamin synthesis capability might be one of the key factors in explaining the ubiquitous distribution and predominance of genera *Baizomonas* and *Baizosediminiarchaeum* (previously subgroups 6/5bb and 8), even the whole class Bathyarchaeia, in global freshwater and marine sediments.

Atypical and flexible pathway for C1 compound metabolism

In addition to complex organics, Bathyarchaeia exhibit notable potential in using multiple C1 compounds, including CO_2, CO, formate, formaldehyde, methanol, methylamine, and methane. For the orders Wuzhiqiibiales and Mazuousiales, the key genes of reductive glycine pathway (rGlyP), which is the seventh natural carbon fixation pathway recently reported (*39*), were identified in their MAGs, including the genes coding for glycine cleavage system (GCS), serine hydroxymethyltransferase (GlyA), and serine deaminase (Sda), as well as the tetrahydrofolate methyl branch of WLP (Figs. 3A and 4). Consistent with this finding, there was strong support for the presence of these genes in the ancestral Wuzhiqiibiales and Mazuousiales genomes (fig. S3 and table S8), suggesting that the rGlyP is an ancient metabolic trait for these two Bathyarchaeia orders. Moreover, both analyzed Mazuousiales MAGs uniquely encode formate dehydrogenase (Fdh), while the key genes (*coxM* and *coxS*) for aerobic CO dehydrogenase are largely found in the order Wuzhiqibiales (Fig. 4). These results indicate that some species from the orders Wuzhiqiibiales and Mazuousiales are likely able to assimilate CO_2, CO, and/or formate via the rGlyP. Distinct from the archaeal phylum Brockarchaeota that also has rGlyP (*40*), additional complete H_4-MPT WLP and CO/formate metabolic potentials are identified in the orders Wuzhiqiibiales and Mazuousiales (Figs. 3A and 4), implying that these Bathyarchaeia lineages have more diversified and flexible capabilities in metabolizing different C1 compounds. Previous studies demonstrated that some Bathyarchaeia species could assimilate ^{13}C-labeled carbonates, presumably through the H_4-MPT WLP (*13*). Here, we provide another possibility: A considerable number of Bathyarchaeia members are capable of autotrophic growth through the rGlyP. Given the prevalence of the orders Wuzhiqiibiales and Mazuousiales in modern marine sediments (*23*, *41*),

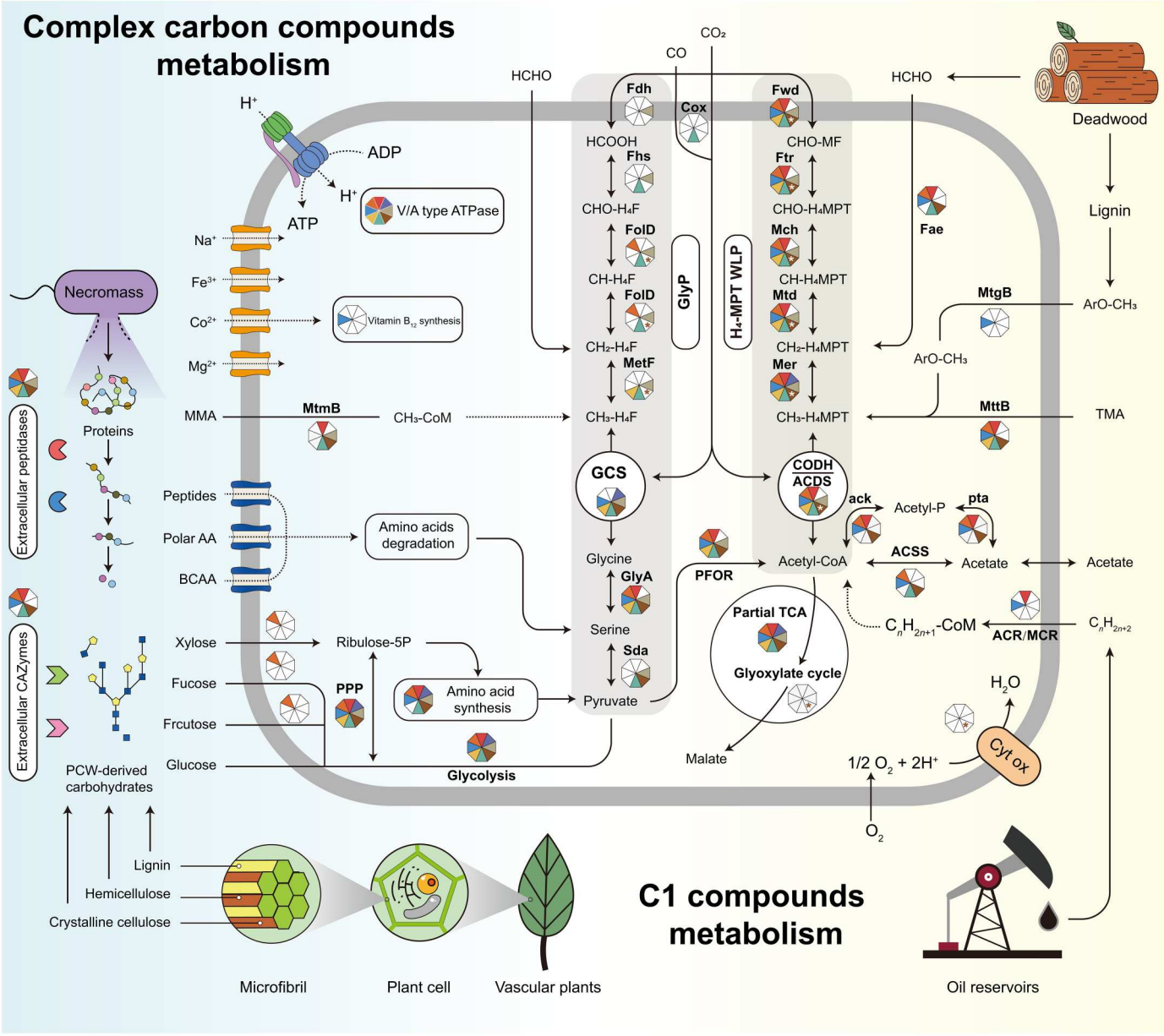

Fig. 4. Overall metabolic reconstruction of the class Bathyarchaeia. The presence of specific pathways or enzymes in at least two MAGs from different samples for each Bathyarchaeia order is indicated by different colored parts in octagon. White color denotes that all enzymes of a specific pathway are absent in any MAGs of the order. The enzymes of the H_4-MPT pathway marked with white stars in brown background, including Fwd, Ftr, Mch, Mtd, Mer, and CODH/ACDS, indicate that they are present in all other MAGs in the order Houtuarculales, except for the special genus *Houtuousia* recovered from soil. In contrast, those marked with brown stars in white background represent that they are exclusively identified in the members of genus *Houtuousia*, but not in other Houtuarculales MAGs. The following metabolic pathways, key enzymes, and compounds were used in the figure (table S7): PPP, archaeal pentose phosphate pathway; TCA cycle, tricarboxylic acid cycle; PFOR, pyruvate:ferredoxin oxidoreductase complex; ack-pta, acetate kinase and phosphate acetyltransferase; ACSS, acetyl-CoA synthetase; rGlyP, reductive glycine pathway; GCS, glycine cleavage system; GlyA, glycine hydroxymethyltransferase; Sda, L-serine dehydratase; Fae, formaldehyde-activating enzyme; Cox, aerobic carbon-monoxide dehydrogenase; Fwd, formylmethanofuran dehydrogenase complex; Ftr, formylmethanofuran–tetrahydromethanopterin *N*-formyltransferase; Mch, methenyltetrahydromethanopterin cyclohydrolase; Mtd, methylenetetrahydromethanopterin dehydrogenase; Mer, coenzyme F_{420}-dependent 5,10-methenyltetrahydromethanopterin reductase; Fdh, formate dehydrogenase; Fhs, formate-tetrahydrofolate ligase; FolD, methylenetetrahydrofolate dehydrogenase ($NADP^+$)/methenyltetrahydrofolate cyclohydrolase; MetF, methylenetetrahydrofolate reductase (NADPH); Cyt ox, cytochrome c oxidase; MttB; MtmB; MCR complex; ACR, alkyl-coenzyme M reductase complex; TMA, trimethylamine; MMA, monomethylamine; BCAA, branched-chain amino acids; $ArOCH_3$, methoxylated aromatic compounds; ATPase, adenosine triphosphatase; ADP, adenosine 5′-diphosphate; CODH/ACDS, anaerobic carbon-monoxide dehydrogenase/acetyl-CoA decarbonylase/synthase; PCW, plant cell wall.

their overlooked metabolic capability and contribution to dark carbon fixation warrant further study.

The genes coding for mono- and trimethylamine-cobalamin MT systems (*mtmB* and *mttB*, respectively) are predicted in seven Bathyarchaeia orders (except for the sole MAG of the order Xuanwuarculales), most of which are distributed in the orders Wuzhiqibiales and Bifangarchaeales (Fig. 3A). Two Jinwuousiales MAGs also contain the encoding genes of methanol and dimethylamine-cobalamin MTs (*mtaB* and *mtbB*, respectively). Ancestral reconstruction results also suggest that the common ancestors of order Jinwuousiales and Wuzhiqibiales likely had the *mtmB* and *mttB* genes (fig. S3). Under anoxic conditions, methylated C1 compounds are mostly used by methanogens for methane production. However, it is unlikely for Bathyarchaeia, as the key gene coding for methyl-coenzyme M reductase (MCR) in methanogenesis is absent in most Bathyarchaeia lineages (see the next part for details). In addition, anaerobic respiration and acetogenesis are alternative pathways for metabolizing methylated compounds (*42*). Accordingly, given the generally missing sulfate reduction or denitrification pathways but the prevalent acetate metabolism-involved enzymes (e.g., ack, pta, and ACSS) in most Bathyarchaeia MAGs (Fig. 3A), it is plausible that their MT system–encoded species are nonmethanogenic methylotrophic archaea that could metabolize multiple methylated compounds via acetogenesis. Recently, ^{13}C-labeled incubation experiments demonstrated active methylotrophy of methylamine and methanol by microbial communities from coalbeds 2 km below the seafloor with minor methanogenesis (*43*). Therefore, the nonmethanogenic methylotrophic lifestyle might be an important adaptive strategy for microbial life in the extremely energy-limited deep biosphere, and so do these Bathyarchaeia that are widely distributed in global deep marine sediments (*9*, *26*).

Moreover, our other work has also demonstrated that one *Baizosediminiarchaeum* species of the lignin-degrading clade used a special MT system to transfer the methyl groups from lignin-derived $ArOCH_3$ to H_4MPT, which are either further oxidized to CO_2 or converted to acetate for energy production (*38*). This suggests that this Bathyarchaeia-specific MT system plays a key role in anaerobic lignin degradation, thus likely explaining their predominance in the lignin-enriched estuarine and nearshore sediments. In addition, formaldehyde is a common toxic C1 by-product released during lignin-derived aromatic degradation (*44*). Here, the gene encoding formaldehyde-activating enzyme (*fae*) is widely identified in the analyzed MAGs, including the lignin-degrading clade, implying that formaldehyde could be detoxified or used via the H_4-MPT WLP for most Bathyarchaeia lineages.

Together, our results reveal the currently unrecognized and diverse roles of Bathyarchaeia in C1 compound transformation: Most of them are inferred to be putative nonmethanogenic methylotrophic archaea with the potential to metabolize different methylated compounds via diverse MT systems, including methanol, methylamines, and $ArOCH_3$. In comparison with Brockarchaeota (*40*), most of these Bathyarchaeia lineages have an extra complete H_4-MPT WLP and more methylamine-associated MT systems (Fig. 4), reflecting their more flexible methylotrophic strategies and substrate preferences, the latter of which might facilitate their survival in the nutrient-limited deep subseafloor as the amino residues released from methylamines are bioavailable nitrogen sources (*45*). Multiple C1 compounds have also been shown to be metabolized through methylotrophic methanogenesis and nonmethanogenic methylotrophy in anoxic sulfate–reducing sediments, where Bathyarchaeia are enriched (*46*). In conclusion, these results provide insights into the diverse and versatile C1 carbon metabolism of Bathyarchaeia by demonstrating their prevalent but overlooked methylotrophic capabilities and thus highlighting their unparalleled significance and contribution to the global carbon cycle.

Two MAGs (BA1 and BA2) from coalbed well formation water contain the MCR-encoding genes (*mcr*) (*15*), which were assigned to different families of the order Baizomonadales herein (table S7). These *mcr*-like genes are inferred to be *acr* genes encoding the alkyl-coenzyme M reductases (ACRs) involved in the anaerobic oxidation of multicarbon alkanes (*16*). After eliminating the possibility of contamination (table S10), we found a complete *mcr* gene cluster in another MAG (3300028193_21) from the order Jinwuousiales. These are the only three MAGs containing the *mcr* or *acr* genes in the 304 MAGs checked in this study and, as far as we are aware, also for all available Bathyarchaeia MAGs from public databases. Further phylogenetic analysis placed the two Baizomonadales-ACRs in the MCR group III, while the only MCR from the order Jinwuousiales was assigned to MCR group II (with mainly putative methanogens from the previous TACK superphylum and a few Archaeoglobi species) (fig. S8). Previous studies indicate that the two *acr* genes of Baizomonadales might have been acquired via horizontal gene transfer (HGT) events from other ACR-encoded euryarchaeotal hosts (*16*). However, the Jinwuousiales *mcr* gene found here more likely evolved vertically from the common ancestor with other MCR-encoding archaea within MCR group II because Bathyarchaeia and the most previous TACK members in this group—such as *Ca*. Methanomethylicia, *Ca*. Korarchaeia, and *Ca*. Nezhaarchaeales—are reclassified to different class- or order-level units of the same phylum Thermoproteota in the GTDB taxonomy (*10*). Overall, these remarkable phylogenetic divergences among the three Bathyarchaeia *acr* genes not only reflect their complex evolutionary history of methane metabolism but also expand their putative ecological roles in methane/alkane cycling (methanogenesis or alkane oxidization).

Coevolution of Bathyarchaeia and early Earth surface environment

By using the temporal constraints of an HGT event of the chromosome segregation (SMC) gene (*47*), the common ancestor of Bathyarchaeia is predicted at ~3.37 billion years (Ga) ago [posterior 95% confidence interval (CI), 3.62 to 3.13 Ga ago], coinciding with the subaerial emergence of the first continent on Earth ~3.46 to 3.2 Ga ago (Fig. 5 and fig. S9) (*48*). In the absence of plant root systems, Paleoarchean surface terrains would have been subject to rapid migration of riverbeds, and given that the tidal range was larger in the past, it is likely that emergent continental crust would have experienced enhanced wetting (*49*). Therefore, the ancient terrestrial phototrophic microbial communities, quite possibly cyanobacteria (*50*), would have colonized the exposed land surface and grew as biological soil crusts and microbial mats. These mats certainly could have been sloughed off and transported by rivers to coastal settings where their detritus accumulated in the sediment pile (*49*). Concomitant with increased continental weathering, the enhanced transport of suspended and dissolved loads to the oceans would ultimately cause the accumulation of sediments to form shallow continental

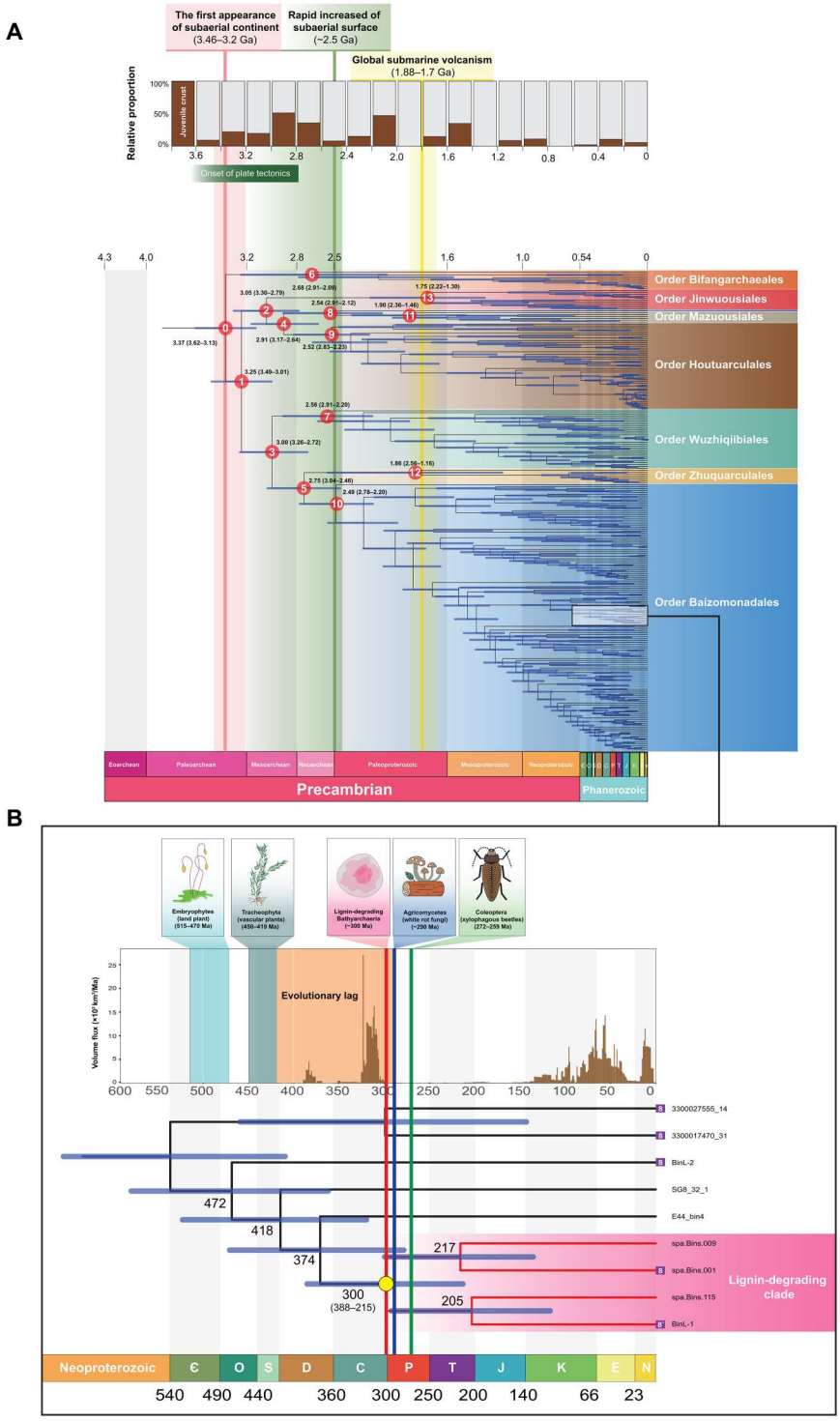

Fig. 5. Evolutionary history of the class Bathyarchaeia and timing correlations with major geological activities. (**A**) Phylogenomic tree and estimated divergence times of the major Bathyarchaeia lineages. The whole tree was constructed on the basis of 259 Bathyarchaeia and 190 reference MAGs by the concatenated alignment of their SMC and 16 conserved proteins (fig. S9). Different color schemes represent the descendent lineages for eight Bathyarchaeia orders (the label of order Xuanwuarculales is removed as the only one MAG). The divergence ages of major nodes are numbered 0 to 13 and labeled with the posterior 95% CIs (flanking horizontal blue bar). The relative proportions of juvenile crust through time are represented as a brown bar chart in the top panel (*60*). The vertical pink, green, and yellow bands indicate the first subaerial continent exposed above the sea at 3.46 to 3.2 Ga ago (*48*), rapidly increasing subaerial landmass at ~2.5 Ga ago (*62*), and the global submarine volcanism intensively occurred at 1.88 to 1.7 Ga ago (*64*), respectively. (**B**) Evolutionary timeline of the Bathyarchaeia lignin-degrading lineage and contemporaneous Permo-Carboniferous coal peak. The tree is part of the whole phylogenomic tree of Bathyarchaeia, comprising nine MAGs from the genus *Baizosediminiarchaeum*. The four MAGs with pink background, which were specifically recovered from the lignin enrichment, are designated as the lignin-degrading Bathyarchaeia clade in this study. The MAGs labeled with the number "8" within the purple box are assigned to the previous subgroup 8. The top panel shows the burial flux of terrestrial organic sediments accumulated in North America through time (*69*) and contemporaneous key evolutionary time points for terrestrial plants (*105*), white rot fungi (*68*), and xylophagous beetles (*70*, *71*).

shelves with increased nutrient supply [e.g., phosphorus (*51*)]. As discussed above, modern Bathyarchaeia species have versatile heterotrophic capabilities and mostly inhabit anoxic coastal sediments, which are characterized by rapid terrestrial sedimentation rates. Accordingly, it is reasonable to infer that the ancestral Bathyarchaeia originated in these newly formed coastal settings associated with continental emergence on Archean Earth. The presence of heterotrophic Bathyarchaeia so early in Earth's history not only implies abundant organic carbon availability but also points to an early terrestrial biosphere that directly arose through the creation of habitable space on land (*52*).

The Paleoarchean rock record certainly points to the presence of stromatolites that grew in the shallow waters fringing the newly exposed continental landmass. For example, the ~3.42-Ga Buck

Reef Chert contains fine anastomizing carbonaceous laminations that have been interpreted as fossil microbial mats (*53*). Similarly, one of the oldest known stromatolites, the ~3.43-Ga Strelley Pool Formation in northwestern Australia, contains complex morphologies that vary along a carbonate platform that is several kilometers long and has been interpreted as a microbial reef (*54*). More recently, the earliest biosignatures were found from the ~3.5-Ga hot spring deposits (*55*). What is consistent among these deposits is the influence of hydrothermal activity in association with frequent volcanic eruptions.

We found that the orders Bifangarchaeales and Jinwuousiales are specific thermophilic lineages inhabiting hot spring and hydrothermal sediments, respectively (Fig. 2B). As the two most ancient lineages, phylogenetic results reveal that the common ancestor of order Bifangarchaeales diverged from the root directly, followed by the emergence of order Jinwuousiales at ~3.05 Ga ago (CI, 3.30 to 2.79 Ga ago) (Fig. 5 and table S11). This implies that Bathyarchaeia very likely evolved in an ancient hydrothermal setting, in accordance with the hot origin theory proposed before (*25*). In this context, we inferred that the common ancestor of Bathyarchaeia might emerge from the shallow submarine volcanic (or terrestrial hot spring) sediment enriched with organics from either continental weathering or in situ production by microbial mats. In addition, early Earth surface environment is likely to have accreted prebiotic methylated compounds from cometary delivery (*56*). Together with the enriched MT systems identified in the two ancient thermophilic orders (Fig. 4 and fig. S5), we suggest that the ancestral Bathyarchaeia might be nonmethanogenic methylotrophic archaea, and this view is consistent with a recent study suggesting the methylotrophic origin of methanogenesis at ~3.8 to 4.1 Ga ago (*18*).

The first extensive diversification of Bathyarchaeia occurred near ~3.0 Ga ago, including monophyletic divergences of the orders Jinwuousiales and Wuzhiqiibiales from other Bathyarchaeia lineages (Fig. 5A). Although the specific timing and mechanism remain highly controversial, ~3.0 Ga ago has been interpreted as a pivotal period during the evolution of Earth's tectonics (*57*, *58*). At that time, the majority (~65%) of juvenile continental crust had already formed (*58*), along with notably altered chemical composition (*59*), increased reworking/growth rate (*57*, *60*), and thickness (*58*), in turn marking the putative onset of plate tectonics (*59*). This timing is coincident with the Kaapvaal and Pilbara cratons (*61*). As the cratons were in isostatic uplift, the felsic rocks became exposed to erosion and weathering, leading to enhanced export of nutrients to the Archaean ocean. We suggest that increased weathering led to higher nutrient delivery, primary productivity, and organic carbon burial, which drove the evolution of certain Bathyarchaeia lineages, such as peptide-using Wuzhiqiibiales, to adapt to the great change in sediment composition.

The second extensive diversification of Bathyarchaeia occurred between ~2.75 and ~ 2.49 Ga ago, during which the taxonomic radiations of the orders Wuzhiqiibiales (~2.56 Ga ago), Houtuarculales (~2.52 Ga ago), and Baizomonadales (~2.49 Ga ago), as well as the common ancestor of Mazuousiales and Xuanwuarculales (~2.54 Ga ago) (Fig. 5A and table S11) emerged. This timing is associated with the rapid and expansive emergence of the subaerial landmass at the Archean-Proterozoic boundary (*62*), similar to the ~3.0 Ga ago, which remarkably increased subaerial weathering (*63*), the burial rate of organic carbon, and the extent of continental and sediment volume (*51*).

The phylogenetic radiation of different Bathyarchaeia lineages with diversified carbon metabolism might have been triggered by changes in emergent crustal lithology. For instance, between ~3.5 and ~3.2 Ga ago, the crust was primarily mafic, but by ~3.0 Ga ago, it became more felsic in composition (*59*). In addition, as a consequence of adapting to cooler Earth surface condition, the depositional environments for Bathyarchaeia changed from inhabiting early geothermal environments and then transitioning to moderate or cold modern marine sediments. Along those lines, the third major diversification event, leading to taxonomic radiation of two hydrothermal orders Jinwuousiales and Zhuquarculales, occurred at ~1.86 (CI, 2.56 to 1.16) and ~1.75 (CI, 2.22 to 1.30) Ga ago, respectively. This time window was concordant with a short period of intense submarine volcanic activity that took place globally at ~1.88 to 1.70 Ga ago (*64*). Therefore, the geological consequences of rapid continental growth and global magmatic eruption approximately 1.8 Ga ago might have profoundly affected the interior diversification of the ancestral hydrothermal Bathyarchaeia, probably by isolating them into distinct ecological niches.

Potential contribution of lignin-degrading Bathyarchaeia clade to the end of Permo-carboniferous coal production peak

An unparalleled interval of carbon sequestration occurred during the late Carboniferous to early Permian (~323 to 252 Ma ago), which corresponded to the greatest coal deposits in Earth's history (*65*), the so-called Permo-Carboniferous coal production peak. As the second most abundant biopolymer on Earth, lignin is sourced from dry woody (vascular) plant material. It is extraordinarily resistant to biological degradation due to its highly amorphous and polyphenolic structure (*66*), making it the main precursor of coal. Today, white rot fungi are the primary lignin degraders, except for a few bacterial species with very limited capabilities for delignification (*66*). Thus, Robinson *et al.* (*67*) hypothesized that a temporal gap of ~120 Ma between the origin of lignin-synthesizing vascular plants and the delayed evolution of lignin-degrading white rot fungi provided a window of opportunity for a substantial amount of recalcitrant vegetal polymers to accumulate and, in turn, become coal. The subsequent molecular dating evidence makes this paradigm more convincing by demonstrating that the common ancestor of white rot fungi did not evolve until ~290 Ma ago (*68*). Recently, certain Bathyarchaeia species from the genus *Baizosediminiarchaeum* (previous subgroup 8) have been reported to be capable of lignin utilization in the anoxic condition (*13*). Here, four MAGs of the genus *Baizosediminiarchaeum* that were specifically recovered from the lignin-enrichment culture in our laboratory were designated as the lignin-degrading Bathyarchaeia clade.

To increase the resolution of the phylogenetic signal of its adjacent lineages, an additional 21 single conserved genes were added to the phylogenetic marker set for the molecular dating of this lignin-degrading clade (see Materials and Methods for details). Phylogenetic analysis suggests that the lignin-degrading lineage is monophyletically evolved from one common ancestor, whose divergence time was estimated at ~300 Ma ago (CI, 388 to 215 Ma ago), coinciding with the emergence of ancestral white rot fungi and the peak of carbon sequestration at the end of Carboniferous period (Fig. 5B). Therefore, we infer that the lignin-degrading Bathyarchaeia ancestor also likely contributed to the sharp decline

in Permo-Carboniferous coal production, similar to what white rot fungi did according to the evolution lag theory.

Distinct from the aerobic white rot fungi, this special Bathyarchaeia clade can degrade lignin under strictly anoxic condition, which means that their ancestor could have continuously consumed buried woody material over hundreds of million years during the coal formation process (*65*). Thus, this Bathyarchaeia clade putatively plays a key role in determining organic carbon sequestration over geological time scales. Furthermore, the putative existence of anaerobic lignin-degrading archaea 300 Ma ago makes anoxic conditions no longer a vital factor determining the fate of recalcitrant woody material after burial in Carboniferous sediments. This finding directly challenges the alternative abiotic hypothesis that explains the unusual coal abundance as a confluence of the disappearance of widespread anoxic waterlogged environments driven by notable climate and tectonic changes (*69*). Moreover, the origin of xylophagous beetles (order Coleoptera) was also coeval with the two microbial lignin degraders (Fig. 5B) (*70*, *71*). Given that their descendants accounted for ~29% of the deadwood degradation in modern forest ecosystems (*72*), the contemporaneous ancestral xylophagous insects might also have contributed to the decline of the Permo-Carboniferous coal peak. Together, our results provide supporting evidence for the evolution lag theory in explaining the unusual coal abundance during the Permo-Carboniferous period. The improved evolutionary lag theory is an excellent case for illustrating the feedback and coevolution between life and Earth over geological time scales, in which Bathyarchaeia, as the most abundant and widespread microorganism on Earth today, might have played a previously unrecognized but substantial role in global carbon cycling at least ~300 Ma ago.

In summary, this study provides a comprehensive story on the phylogeny, metabolism, and evolution of Bathyarchaeia within an improved phylogenetically congruent taxonomic framework, which is established on the basis of the largest Bathyarchaeia MAG dataset and GTDB methodology to date. Comparative genomics results show the distinct environmental distribution and diversified carbon metabolism among different Bathyarchaeia lineages. In particular, multiple atypical C1 metabolic pathways—including different MT systems, the rGlyP and the newly found MCR—reflect their previously overlooked potential roles as anaerobic methylotrophs, dark carbon fixers, and methane/alkane oxidizers. Moreover, the evolutionary history of Bathyarchaeia may be linked to ancient geothermal sedimentary environments, while the subsequent continental expansion and intensive hydrothermal activity probably facilitated the notable diversifications of different Bathyarchaeia lineages. The temporal coincidence between the appearance of the lignin-degrading Bathyarchaeia clade and white rot fungi implies that certain ancient Bathyarchaeia groups likely contributed to the sharp decline in coal deposits during the Permo-Carboniferous period.

Our work provides systematic insights into the taxonomic and metabolic diversity of the whole class Bathyarchaeia, which not only leads to a better understanding of their overall ecological and biogeochemical significance but also lays the foundation for future ecological and functional studies of different Bathyarchaeia lineages. This study provides a good case for illustrating the feedbacks and coevolution between life and Earth, specifically the roles of Bathyarchaeia in carbon cycling on geological time scales. Nevertheless, future wider ecological investigations and continuous efforts in cultivating or isolating pure strains are needed to further decipher their specific roles in modern environments and potential implications in deep time.

MATERIALS AND METHODS

Sample collection and DNA extraction

In this study, two natural sediment samples and four sediment incubation enrichment cultures were sequenced by our laboratory and used to achieve more Bathyarchaeia MAGs. All necessary information involved in sampling and cultivation conditions is listed in table S1.

Deep-sea sediment from South China Sea

One 8.3-m piston core named QDN-14B was sampled from the Haima cold seep located in the South China Sea during the cruise"20150402" R/V *Haiyang IV* in April 2015, and the specific subsampling and preservation method was described in the previous study (*73*).

Coastal sediment from East China Sea

One 6-m gravity core A2-3 was collected from the Yangzi estuary (30.927933°N, 122.473233°E) in the East China Sea during a cruise in July 2017. The subsamples used for molecular analysis were cut into 10-cm intervals, kept in liquid nitrogen during transport to the laboratory, and stored at −80°C until further analysis.

Sediment enrichment with lignin

Four sets of sediments from different natural environments were aerobically incubated with lignin to enrich Bathyarchaeia. One intertidal sediment sample (DYS) was collected from Dayangshan Island (30.592817°N, 122.083493°E), Hangzhou Bay, China, as described previously (*13*). One freshwater sediment sample (PYH) was obtained from Poyang Lake, Jiangxi, China. The other two sediment samples (SQ and SRBZ) were collected from two hot spring pools with different temperatures in Tengchong, China. The detailed cultivation method for the DYS set was described in (*13*), and the cultivation conditions for the other three sets are listed in table S1.

For enrichment cultures, the DNA extraction method was the same as previously described (*13*). A modified SDS-based method was used to extract high-quality DNA from two natural sediment samples for metagenomic library construction and sequencing.

Metagenomic sequencing, assembly, and binning

For all sediment samples mentioned above, metagenomic libraries were constructed and sequenced on the HiSeq X Ten platform (2 × 150–base pair paired-end reads). First, raw metagenomic reads for each sample were trimmed by using Trimmomatic v0.39 (*74*) with default parameters and possible adaptors. After quantity control, clean reads were assembled into contigs by using MEGAHIT v1.2.9 (*75*) with the parameter -min-count 2 -k-min 41 -kmin-1pass -k-max 147 -k-step 10, followed by mapping clean reads to contigs by using bowtie2 v2.4.4 (*76*) with very-sensitive pattern for estimating the sequencing coverage of each contig.

Metagenomic binning was performed individually for each sample by a modified protocol from Hou *et al.* (*77*). First, MetaBAT2 v2.2.15 (*78*) and MaxBin v2.2.7 (*79*) were used to recover MAGs with contigs more than 1 Kb (1.5 Kb for MetaBAT2), 3Kb, and 5 Kb in each metagenomic sample, respectively. The resulting MAGs were refined by removing contigs with incongruent taxonomic information and divergent genomic

properties (e.g., GC content and coverage) based on the "outliers" pattern from RefineM v0.0.22 (https://github.com/donovan-h-parks/RefineM). All refined MAGs from the same sample were integrated into optimized, nonredundant bins by using DAS Tool v1.1.4 (*80*). The quality and initial taxonomic classification of integrated MAGs were estimated by checkM v1.1.3 (*81*) and GTDB-Tk v1.5.1 (*21*), respectively. High-quality MAGs (completeness of >90% and contamination of <10%) assigned to Bathyarchaeia were retained for downstream analysis.

Public data collection

Candidate Bathyarchaeia MAGs were downloaded from the genome reference databases of NCBI (https://ncbi.nlm.nih.gov/), GTDB (*10*), and GEM catalog (*20*) based on the key words "Bathyarchaeota," "MCG" (miscellaneous Crenarchaeota group), and "Bathyarchaeia" (final data collection in December 2020). The quality of these candidate Bathyarchaeia MAGs was estimated by checkM v1.1.3 (*81*). In addition, sampling and environmental information for each available candidate Bathyarchaeia were manually searched and curated in the corresponding literatures.

MAG refinement and subgroup affiliation

Because of the overlapping among databases and possible incorrect taxonomic assignment, all available candidate Bathyarchaeia MAGs from public databases and our own datasets were carefully checked for potential false positivity in taxonomic classification as Bathyarchaeia by using the classify_wf function in GTDB-Tk v1.5.1 (*21*). After removing the MAGs assigned to other microbial lineages, the average nuclear acid identities (ANIs) between pairwise sets of MAGs were calculated by using FastANI (*82*). These Bathyarchaeia MAGs were clustered into different species-level groups based on 95% genome-wide ANI threshold. For each group, the one with the best quality score (completeness − contamination × 5) was selected as the representative MAG and retained for downstream analysis. In general, the genome quality of selected representative Bathyarchaeia MAGs satisfies the median quality based on the MIMAG standard, although a few exceptions exist to retain some well-known or important Bathyarchaeia species, such as the first single cell–sequenced Bathyarchaeia genome (SCGC AB-539-E09) (*3*), which facilitate the further discussion about their metabolic characteristics and evolutionary history. The accession number and source for each representative Bathyarchaeia MAG are listed in table S2.

For each representative Bathyarchaeia MAG, their 16*S* rRNA genes were initially identified and extracted by using Barrnap 0.9 (https://github.com/tseemann/barrnap) and manually curated BLASTN check in SILVA 138.1 SSU Ref database (https://arb-silva.de/) for potential contamination. The reference dataset includes two parts: The first is the 16*S* rRNA genes with confirmed subgroup classification information from Zhou *et al.* (*9*), while the other is collected from SILVA 138.1 SSU Ref database. The reference sequences from SILVA were clustered by using CD-HIT v4.8.1 (https://github.com/weizhongli/cdhit) with 90% identity and compared with the first dataset for deredundancy. Subsequently, the nonredundant reference dataset, 16*S* rRNA genes identified from Bathyarchaeia MAGs, and additional archaeal outgroup sequences were aligned by using SILVAngs (https://ngs.arb-silva.de/silvangs) and trimmed by trimAI v1.4.rev15 with the -automated1 pattern (*83*). The maximum likelihood phylogenetic tree was constructed by using RaxML version 8.2.12 (*84*) with the -m GTRGAMMA -N autoMRE parameter. Note that some representative Bathyarchaeia MAGs without 16*S* rRNA genes were assigned to subgroups identical to their cospecies (ANI > 95%) from the same cluster if the latter had the subgroup-assigned 16*S* rRNA genes (table S3).

Global survey of geographic distribution

To investigate the environmental and geographic distribution of Bathyarchaeia in natural settings, we searched the taxonomic profiles of 244,459 public metagenomes based on Sandpiper database v0.0.23 (https://sandpiper.qut.edu.au/), which were generated using the SingleM v1.0.0beta5 (https://github.com/wwood/singlem). For those samples containing Bathyarchaeia (12,078 metagenomes), we further analyzed the proportion of the seven Bathyarchaeia orders with corresponding GTDB taxa. As the newly found order Mazuousiales is not included in the default GTDB taxonomy (07-RS207) used in Sandpiper v0.0.23, we specifically created a custom SingleM reference dataset (SingleM"metapackage"), which included the order Mazuousiales based on their MAGs for profiling its abundance in the public metagenome dataset. This Mazuousiales-inclusive metapackage was used to reassign taxonomy to each of the 12,078 Bathyarchaeia-containing community profiles using SingleM "renew," creating Mazuousiales-inclusive community profiles for each. The global survey information reported here are those profiles containing >1% Bathyarchaeia in these community profiles (table S6). Metadata (e.g., sampling environment, ecotype, and geographic location) associated with each sample was collected using Kingfisher "annotate" (https://github.com/wwood/kingfisher-download) and manually curated because of the possible incorrect, vague, or missing information provided by the NCBI SRA dataset.

Phylogenetic rank normalization and nomenclature

Phylogenomic inference

A total of 304 representative Bathyarchaeia MAGs were selected to construct the phylogenomic tree for taxonomic rank normalization, with an additional 55 archaeal genomes from other phyla (and other orders from the phylum Thermoproteota) as the reference dataset (table S2). Their 122 archaeal-specific conserved proteins were identified, aligned, trimmed individually, and concatenated together by using the identity function in GTDB-Tk v1.5.1 (*21*). The maximum likelihood phylogenomic tree for these 359 archaeal MAGs/genomes was inferred by IQ-TREE v2.1.2 (*85*) using the LG+F+R10+C60 model with 1000 ultrafast bootstrap replicates and 1000 bootstrap replicates for Shimodaira Hasegawa–like approximate likelihood ratio test (SH-aLRT).

Rank classification and normalization

The taxonomic rank for each Bathyarchaeia MAG was initially inferred by using the classify_wf function in GTDB-Tk v1.5.1 (*21*) and further curated carefully in consideration of the phylogenetic topological structure and environmental information of each lineage. The modified taxonomic ranks of Bathyarchaeia were normalized and evaluated by PhyloRank v0.1.10 (https://github.com/dparks1134/PhyloRank) based on the RED values of each taxonomic rank (table S4).

Automated nomenclature

Given the difficulty in manually creating hundreds of genus names for Bathyarchaeia MAGs, an automated approach was applied to solve this problem by GAN (*86*) based on three individual word

roots from traditional Chinese mythological figures, environmental terms, and universal Latin endings or diminutives provided by ICNP (International Code of Nomenclature of Prokaryotes) (table S12). For each proposed genus, their own type materials are designated as those with genomic quality that meets the defined standard (*87*) (or only the best one in all MAGs if no other MAGs satisfy the standard) and environmental representativity, which are used to define the names of the genus and corresponding higher-level taxonomic ranks by adding an appropriate suffix to the stem based on the ICNP rules (*88*). The nomenclature and etymology of each taxonomic rank are listed in table S5.

Genomic analysis and functional annotation

Genomic characteristics

General genomic traits—including predicted genome size [(original assembled MAG size) / (completeness + contamination)], GC content, predicted gene number, and coding density—were accessed from the checkM results directly. Carbon-ARSC and N-ARSC were calculated by using custom scripts from Feng *et al.* (*89*). The average percentage of amino acid composition for the 304 representative Bathyarchaeia MAGs were analyzed using CodonW (https://codonw.sourceforge.net/). The significant differences in each genomic characteristic among different orders were compared by using the Wilcoxon test with Bonferroni correction.

Functional gene annotation

To minimize the error possibility from incompleteness or contamination, 86 high-quality MAGs were selected as representative Bathyarchaeia MAGs to perform metabolic reconstruction in detail. Open reading frame was predicted from each MAG by Prodigal v2.6.3 (*90*) with single mode and further annotated in a series of databases in parallel, including Kyoto Encyclopedia of Genes and Genomes (*91*) and Pfam (*92*) by using DRAM (*93*) with default parameters. The resulting annotation for specific key genes was further verified across at least two databases or manually checked on the basis of the self-built custom database using Diamond blastp v0.9.14 (*94*). Genes encoding carbohydrate degradation enzymes described in the CAZymes database V10 (*95*) were searched using eCAMI with default 8 k-mer peptide library (*96*), and the annotated CAZymes subfamilies were further mapped to the dbCAN2 database for substrate annotation (*97*). The peptidase and proteinase encoding genes were annotated in the MEROPS database 12.4 (*98*) using Diamond blastp v0.9.14 (*94*) with a threshold of coverage >30% and *e* value < 1×10^{-10}. The possible signal peptides in each peptidase sequences were predicted by SignalP-5.0b (*99*) for identifying the extracellular peptidases in these Bathyarchaeia MAGs (table S9). The metabolic completeness is defined by the proportion of annotated genes in all genes for one specific pathway, while the carbohydrate and protein degrading potential for each MAG is evaluated by the total number of genes annotated in the CAZyme and MEROPS databases, respectively. The functional gene annotation for each Bathyarchaeia MAG and complete gene catalog of each pathway are listed in table S7. For specific key genes only found in the single MAG, including the *mcr* gene cluster in the MAG 3300028193_21 and other functional genes in the only MAG of order Xuanwuarculales (_RBG_16_48_13-), the taxonomy of their adjacent genes from the same contig was carefully checked on the basis of the blastp results in the NCBI database to eliminate the possibility that these key genes are from the potential contamination (table S10).

Phylogenetic analysis of McrA

The phylogenies of McrA protein sequences retrieved from the representative Bathyarchaeia MAGs were further analyzed in this study. The phylogenetic tree of McrA was constructed together with the reference sequences from a previous study on anaerobic alkane–oxidizing archaea (*16*) by IQ-TREE v2.1.2 with LG+F+G+C60 model.

Ancestral metabolic reconstruction

For inferring the ancestral metabolic repertoire of each Bathyarchaeia order, amalgamated likelihood estimation (ALE) analyses were performed using the maximum likelihood implementation of the undated ALE algorithm using a sample of 1000 ultrafast bootstrap trees for the gene families of important carbon metabolisms and the species tree of 86 selected representative Bathyarchaeia MAGs. The species tree was constructed using IQ-TREE v2.1.2 on the basis of the concatenated 122 archaeal marker proteins implemented in GTDB-Tk v1.5.1 with LG+F+R7+C20 model, while the gene trees were constructed on the basis of their own predicted model individually (table S8). The gene copy number for each gene family that was present in the root node of each Bathyarchaeia order was estimated using the ALE package (*100*).

Phylogenetic divergence time estimation

Phylogenomic analysis

The molecular dating analysis of Bathyarchaeia used the same methodology as previous studies on the evolution of methanogens (*47*) and methanogenesis (*18*), which are based on a temporal constraint of the HGT event of SMC-encoding gene from certain euryarchaeotal methanogen lineages to the ancestor of Cyanobacteria, for which no unambiguous fossil records are available. For this purpose, a total of 259 high- and median-quality Bathyarchaeia MAGs were selected for phylogenetic analysis, together with an additional 190 reference genomes/MAGs from representative Cyanobacteria and other archaeal phyla (table S13). Sixteen single-copy conserved proteins and the SMC protein were predicted from each analyzed MAG after careful curation. To check the phylogenetic conservation and potential HGT events of each marker gene, the trees of their protein sequences were constructed individually by IQ-TREE v2.1.2 with their own predicted models (table S14). The results support the notion that they are conserved phylogenetic marker genes for the class Bathyarchaeia and suitable for the next molecular dating analysis (fig. S10). The protein alignments of 16 conserved marker genes and the SMC-encoding gene were concatenated and analyzed by IQ-TREE v2.1.2 under LG+F+R10+C60 model with 1000 ultrafast bootstrap replicates and 1000 bootstrap replicates for SH-aLRT.

Divergence time estimation

The molecular dating analysis was conducted by using MCMCtree in paml version 4.9j with the WAG model and iteratively run until convergence (*101*). Three nodes were calibrated on the phylogenomic tree: (i) The predicted archaeal root age at ~4.29 to 3.8 Ga (*102*); (ii) the divergence time of Cyanobacteria is estimated at 2.5 to 3.0 Ga ago (*103*); and (iii) fossil evidence suggested that cyanobacterial clades Nostocales and Stigonematales already exist at ~2.0 to 1.2 Ga ago (*47*, *104*), which was defined as >1.2, >1.7, and >2.0 Ga, respectively, and calculated in the next divergence time estimation individually (*19*).

Alternative dating strategy for the lignin-degrading clade

Although the phylogenetic topology of the whole class Bathyarchaeia is rather robust on the basis of the marker set of 16 single-copy conserved proteins plus SMC protein, the phylogenies of the Bathyarchaeia lignin-degrading clade and close lineages show slight instability, which is attributed to the absence of marker genes in their MAGs (table S15). To address this issue, another 21 single-copy conserved marker genes were selected to supplement the marker gene set to provide enough phylogenetic information for this special clade (table S15). In this way, a total of 37 marker genes coding for proteins plus the SMC protein were used to construct the phylogenomic tree of the whole Bathyarchaeia lineage for estimating the divergence time of the lignin-degrading clade. The dating method and calibration time points are exactly the same as described before.

Supplementary Materials

This PDF file includes:
Figs. S1 to S10
Legends for tables S1 to S15

Other Supplementary Material for this manuscript includes the following:
Tables S1 to S15

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. F. Inagaki, M. Suzuki, K. Takai, H. Oida, T. Sakamoto, K. Aoki, K. H. Nealson, K. Horikoshi, Microbial communities associated with GeologicalHorizons in coastal subseafloor sediments from the sea of Okhotsk. *Appl. Environ. Microbiol.* **69**, 7224–7235 (2003).
2. J. Meng, J. Xu, D. Qin, Y. He, X. Xiao, F. Wang, Genetic and functional properties of uncultivated MCG archaea assessed by metagenome and gene expression analyses. *ISME J.* **8**, 650–659 (2014).
3. K. G. Lloyd, L. Schreiber, D. G. Petersen, K. U. Kjeldsen, M. A. Lever, A. D. Steen, R. Stepanauskas, M. Richter, S. Kleindienst, S. Lenk, A. Schramm, B. B. Jorgensen, Predominant archaea in marine sediments degrade detrital proteins. *Nature* **496**, 215–218 (2013).
4. M. Fillol, J. C. Auguet, E. O. Casamayor, C. M. Borrego, Insights in the ecology and evolutionary history of the miscellaneous crenarchaeotic group lineage. *ISME J.* **10**, 665–677 (2016).
5. Y. He, M. Li, V. Perumal, X. Feng, J. Fang, J. Xie, S. M. Sievert, F. Wang, Genomic and enzymatic evidence for acetogenesis among multiple lineages of the archaeal phylum bathyarchaeota widespread in marine sediments. *Nat. Microbiol.* **1**, 16035 (2016).
6. Y. L. Qi, P. N. Evans, Y. X. Li, Y. Z. Rao, Y. N. Qu, S. Tan, J. Y. Jiao, Y. T. Chen, B. P. Hedlund, W. S. Shu, Z. S. Hua, W. J. Li, Comparative genomics reveals thermal adaptation and a high metabolic diversity in "*Candidatus* Bathyarchaeia". *mSystems* **6**, e0025221 (2021).
7. S. Diamond, P. F. Andeer, Z. Li, A. Crits-Christoph, D. Burstein, K. Anantharaman, K. R. Lane, B. C. Thomas, C. Pan, T. R. Northen, J. F. Banfield, Mediterranean grassland soil c-n compound turnover is dependent on rainfall and depth, and is mediated by genomically divergent microorganisms. *Nat. Microbiol.* **4**, 1356–1367 (2019).
8. K. Kubo, K. G. Lloyd, J. F. Biddle, R. Amann, A. Teske, K. Knittel, Archaea of the miscellaneous crenarchaeotal group are abundant, diverse and widespread in marine sediments. *ISME J.* **6**, 1949–1965 (2012).
9. Z. Zhou, J. Pan, F. Wang, J. D. Gu, M. Li, Bathyarchaeota: Globally distributed metabolic generalists in anoxic environments. *FEMS Microbiol. Rev.* **42**, 639–655 (2018).
10. C. Rinke, M. Chuvochina, A. J. Mussig, P. A. Chaumeil, A. A. Davin, D. W. Waite, W. B. Whitman, D. H. Parks, P. Hugenholtz, A standardized archaeal taxonomy for the genome taxonomy database. *Nat. Microbiol.* **6**, 946–959 (2021).
11. L. M. Seyler, L. M. McGuinness, L. J. Kerkhof, Crenarchaeal heterotrophy in salt marsh sediments. *ISME J.* **8**, 1534–1543 (2014).
12. H. Na, M. A. Lever, K. U. Kjeldsen, F. Schulz, B. B. Jørgensen, UnculturedDesulfobacteraceaeand crenarchaeotal group C3 incorporate ^{13}C-acetate in coastal marine sediment. *Environ. Microbiol. Rep.* **7**, 614–622 (2015).
13. T. Yu, W. Wu, W. Liang, M. A. Lever, K. U. Hinrichs, F. Wang, Growth of sedimentary bathyarchaeota on lignin as an energy source. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 6022–6027 (2018).
14. D.-D. Lin, Y.-F. Liu, L. Zhou, S.-Z. Yang, J.-D. Gu, B.-Z. Mu, Stimulation of bathyarchaeota in enrichment cultures by syringaldehyde, 4-hydroxybenzaldehyde and vanillin under anaerobic conditions. *Int. Biodeter. Biodegr.* **171**, 105409 (2022).
15. P. N. Evans, D. H. Parks, G. L. Chadwick, S. J. Robbins, V. J. Orphan, S. D. Golding, G. W. Tyson, Methane metabolism in the archaeal phylum bathyarchaeota revealed by genome-centric metagenomics. *Science* **350**, 434–438 (2015).
16. Y. Wang, G. Wegener, J. Hou, F. Wang, X. Xiao, Expanding anaerobic alkane metabolism in the domain of archaea. *Nat. Microbiol.* **4**, 595–602 (2019).
17. Y. Yang, C. Zhang, T. M. Lenton, X. Yan, M. Zhu, M. Zhou, J. Tao, T. J. Phelps, Z. Cao, The evolution pathway of ammonia-oxidizing archaea shaped by major geological events. *Mol. Biol. Evol.* **38**, 3637–3648 (2021).
18. Y. Wang, G. Wegener, T. A. Williams, R. Xie, J. Hou, F. Wang, X. Xiao, A methylotrophic origin of methanogenesis and early divergence of anaerobic multicarbon alkane metabolism. *Sci. Adv.* **7**, eabd7180 (2021).
19. Y. Wang, R. Xie, J. Hou, Z. Lv, L. Li, Y. Hu, H. Huang, F. Wang, The late archaean to early proterozoic origin and evolution of anaerobic methane-oxidizing archaea. *mLife* **1**, 96–100 (2022).
20. S. Nayfach, S. Roux, R. Seshadri, D. Udwary, N. Varghese, F. Schulz, D. Wu, D. Paez-Espino, I. M. Chen, M. Huntemann, K. Palaniappan, J. Ladau, S. Mukherjee, T. B. K. Reddy, T. Nielsen, E. Kirton, J. P. Faria, J. N. Edirisinghe, C. S. Henry, S. P. Jungbluth, D. Chivian, P. Dehal, E. M. Wood-Charlson, A. P. Arkin, S. G. Tringe, A. Visel, I. M. D. Consortium, T. Woyke, N. J. Mouncey, N. N. Ivanova, N. C. Kyrpides, E. A. Eloe-Fadrosh, A genomic catalog of earth's microbiomes. *Nat. Biotechnol.*, 499–509 (2021).
21. P. A. Chaumeil, A. J. Mussig, P. Hugenholtz, D. H. Parks, GTDB-tk: A toolkit to classify genomes with the genome taxonomy database. *Bioinformatics* **36**, 1925–1927 (2019).
22. P. Yarza, P. Yilmaz, E. Pruesse, F. O. Glockner, W. Ludwig, K. H. Schleifer, W. B. Whitman, J. Euzeby, R. Amann, R. Rossello-Mora, Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat. Rev. Microbiol.* **12**, 635–645 (2014).
23. P. Wang, T. Zhang, S. Chen, X. Li, D. Lai, S. Gao, W. Xie, C. Zhang, Niche specificity and potential terrestrial organic carbon utilization of benthic bathyarchaeota in a eutrophic subtropic estuarine system. *Chem. Geol.* **556**, 119839 (2020).
24. K. Anantharaman, C. T. Brown, L. A. Hug, I. Sharon, C. J. Castelle, A. J. Probst, B. C. Thomas, A. Singh, M. J. Wilkins, U. Karaoz, E. L. Brodie, K. H. Williams, S. S. Hubbard, J. F. Banfield, Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. *Nat. Commun.* **7**, 13219 (2016).
25. X. Feng, Y. Wang, R. Zubin, F. Wang, Core metabolic features and hot origin of bathyarchaeota. *Engineering* **5**, 498–504 (2019).
26. X. Yin, G. Zhou, M. Cai, Q. Z. Zhu, T. Richter-Heitmann, D. A. Aromokeye, Y. Liu, R. Nimzyk, Q. Zheng, X. Tang, M. Elvert, M. Li, M. W. Friedrich, Catabolic protein degradation in marine sediments confined to distinct archaea. *ISME J.* **16**, 1617–1626 (2022).
27. Q. P. Liu, G. Sulzenbacher, H. Yuan, E. P. Bennett, G. Pietz, K. Saunders, J. Spence, E. Nudelman, S. B. Levery, T. White, J. M. Neveu, W. S. Lane, Y. Bourne, M. L. Olsson, B. Henrissat, H. Clausen, Bacterial glycosidases for the production of universal red blood cells. *Nat. Biotechnol.* **25**, 454–464 (2007).
28. A. Levasseur, E. Drula, V. Lombard, P. M. Coutinho, B. Henrissat, Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol. Biofuels* **6**, 41 (2013).
29. A. Channing, A review of active hot-spring analogues of rhynie: Environments, habitats and ecosystems. *Philos. Trans. R Soc. Lond. B Biol. Sci.* **373**, 20160490 (2018).
30. J. J. Grzymski, A. M. Dussaq, The significance of nitrogen cost minimization in proteomes of marine microorganisms. *ISME J.* **6**, 71–80 (2012).
31. J. G. Bragg, C. L. Hyder, Nitrogen versus carbon use in prokaryotic genomes and proteomes. *Proc. Biol. Sci.* **271**, S374–S377 (2004).
32. S. J. Biller, P. M. Berube, D. Lindell, S. W. Chisholm, Prochlorococcus: The structure and function of collective diversity. *Nat. Rev. Microbiol.* **13**, 13–27 (2015).
33. H. L. Kornberg, H. A. Krebs, Synthesis of cell constituents from C2-units by a modified tricarboxylic acid cycle. *Nature* **179**, 988–991 (1957).
34. J. A. Serrano, M. Camacho, M. J. Bonete, Operation of glyoxylate cycle in halophilic archaea: Presence of malate synthase and isocitrate lyase in haloferax volcanii. *FEBS Lett.* **434**, 13–16 (1998).
35. M. Lübben, K. Morand, Novel prenylated hemes as cofactors of cytochrome oxidases. Archaea have modified hemes a and o. *J. Biol. Chem.* **269**, 21473–21479 (1994).

36. A. C. Doxey, D. A. Kurtz, M. D. Lynch, L. A. Sauder, J. D. Neufeld, Aquatic metagenomes implicate thaumarchaeota in global cobalamin production. *ISME J.* **9**, 461–471 (2015).
37. N. Kanno, S. Kato, T. Itoh, M. Ohkuma, S. Shigeto, Resonance raman analysis of intracellular vitamin B_{12} analogs in methanogenic archaea. *Anal. Sci. Adv.* **3**, 165–173 (2022).
38. T. Yu, H. Hu, X. Zeng, Y. Wang, D. Pan, F. Wang, Cultivation of widespread bathyarchaeia reveals a novel methyltransferase system utilizing lignin-derived aromatics. *bioRxiv* 2022. 10.23.513425 (2022).
39. I. Sanchez-Andrea, I. A. Guedes, B. Hornung, S. Boeren, C. E. Lawson, D. Z. Sousa, A. Bar-Even, N. J. Claassens, A. J. M. Stams, The reductive glycine pathway allows autotrophic growth of desulfovibrio desulfuricans. *Nat. Commun.* **11**, 5090 (2020).
40. V. De Anda, L. X. Chen, N. Dombrowski, Z. S. Hua, H. C. Jiang, J. F. Banfield, W. J. Li, B. J. Baker, Brockarchaeota, a novel archaeal phylum with unique and versatile carbon cycling pathways. *Nat. Commun.* **12**, 2404 (2021).
41. D. Zou, J. Pan, Z. Liu, C. Zhang, H. Liu, M. Li, The distribution of bathyarchaeota in surface sediments of the pearl river estuary along salinity gradient. *Front. Microbiol.* **11**, 285 (2020).
42. J. Sun, M. A. Mausz, Y. Chen, S. J. Giovannoni, Microbial trimethylamine metabolism in marine environments. *Environ. Microbiol.* **21**, 513–520 (2019).
43. E. Trembath-Reichert, Y. Morono, A. Ijiri, T. Hoshino, K. S. Dawson, F. Inagaki, V. J. Orphan, Methyl-compound use and slow growth characterize microbial life in 2-km-deep subseafloor coal and shale beds. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E9206–E9215 (2017).
44. Z. Zhang, Y. Wang, P. Zheng, J. Sun, Promoting lignin valorization by coping with toxic C1 byproducts. *Trends Biotechnol.* **39**, 331–335 (2021).
45. G. C. Zhuang, T. D. Pena-Montenegro, A. Montgomery, K. S. Hunter, S. B. Joye, Microbial metabolism of methanol and methylamine in the gulf of mexico: Insight into marine carbon and nitrogen cycling. *Environ. Microbiol.* **20**, 4543–4554 (2018).
46. G. Zhuang, L. Xu, Q. Liang, X. Fan, Z. Xia, S. B. Joye, F. Wang, Biogeochemistry, microbial activity, and diversity in surface and subsurface deep-sea sediments of south china sea. *Limnol. Oceanogr.* **64**, 2252–2270 (2019).
47. J. M. Wolfe, G. P. Fournier, Horizontal gene transfer constrains the timing of methanogen evolution. *Nat. Ecol. Evol.* **2**, 897–903 (2018).
48. P. Chowdhury, J. A. Mulder, P. A. Cawood, S. Bhattacharjee, S. Roy, A. N. Wainwright, O. Nebel, S. Mukherjee, Magmatic thickening of crust in non-plate tectonic settings initiated the subaerial rise of earth's first continents 3.3 to 3.2 billion years ago. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2105746118 (2021).
49. S. V. Lalonde, K. O. Konhauser, Benthic perspective on earth's oldest evidence for oxygenic photosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 995–1000 (2015).
50. J. S. Boden, K. O. Konhauser, L. J. Robbins, P. Sánchez-Baracaldo, Timing the evolution of antioxidant enzymes in cyanobacteria. *Nat. Commun.* **12**, 4742 (2021).
51. N. D. Greber, N. Dauphas, A. Bekker, M. P. Ptáˇek, I. N. Bindeman, A. Hofmann, Titanium isotopic evidence for felsic crust and plate tectonics 3.5 billion years ago. *Science* **357**, 1271–1274 (2017).
52. N. J. Planavsky, S. A. Crowe, M. Fakhraee, B. Beaty, C. T. Reinhard, B. J. W. Mills, C. Holstege, K. O. Konhauser, Evolution of the structure and impact of earth's biosphere. *Nat. Rev. Earth Environ.* **2**, 123–139 (2021).
53. M. M. Tice, D. R. Lowe, Photosynthetic microbial mats in the 3,416-myr-old ocean. *Nature* **431**, 549–552 (2004).
54. A. C. Allwood, M. R. Walter, B. S. Kamber, C. P. Marshall, I. W. Burch, Stromatolite reef from the early archaean era of australia. *Nature* **441**, 714–718 (2006).
55. T. Djokic, M. J. Van Kranendonk, K. A. Campbell, M. R. Walter, C. R. Ward, Earliest signs of life on land preserved in ca. 3.5 Ga hot spring deposits. *Nat. Commun.* **8**, 15263 (2017).
56. C. F. Chyba, P. J. Thomas, L. Brookshaw, C. Sagan, Cometary delivery of organic molecules to the early earth. *Science* **249**, 366–373 (1990).
57. B. Dhuime, C. J. Hawkesworth, P. A. Cawood, C. D. Storey, A change in the geodynamics of continental growth 3 billion years ago. *Science* **335**, 1334–1336 (2012).
58. B. Dhuime, A. Wuestefeld, C. J. Hawkesworth, Emergence of modern continental crust about 3 billion years ago. *Nat. Geosci.* **8**, 552–555 (2015).
59. M. Tang, K. Chen, R. L. Rudnick, Archean upper crust transition from mafic to felsic marks the onset of plate tectonics. *Science* **351**, 372–375 (2016).
60. M. Garçon, Episodic growth of felsic continents in the past 3.7 Ga. *Sci. Adv.* **7**, eabj1807 (2021).
61. M. J. Van Kranendonk, R. H. Smithies, W. L. Griffin, D. L. Huston, A. H. Hickman, D. C. Champion, C. R. Anhaeusser, F. Pirajno, N. M. W. Roberts, M. V. Kranendonk, S. Parman, S. Shirey, P. D. Clift, in *Continent formation through time* (Geological Society of London, 2015; https://doi.org/10.1144/SP389.12), vol. 389, p. 0.
62. I. N. Bindeman, D. O. Zakharov, J. Palandri, N. D. Greber, N. Dauphas, G. J. Retallack, A. Hofmann, J. S. Lackey, A. Bekker, Rapid emergence of subaerial landmasses and onset of a modern hydrologic cycle 2.5 billion years ago. *Nature* **557**, 545–548 (2018).
63. X. M. Liu, R. L. Rudnick, Constraints on continental crustal mass loss via chemical weathering using lithium and its isotopes. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 20873–20880 (2011).
64. B. Rasmussen, I. R. Fletcher, A. Bekker, J. R. Muhling, C. J. Gregory, A. M. Thorne, Deposition of 1.88-billion-year-old iron formations as a consequence of rapid crustal growth. *Nature* **484**, 498–501 (2012).
65. D. Hibbett, R. Blanchette, P. Kenrick, B. Mills, Climate, decay, and the death of the coal forests. *Curr. Biol.* **26**, R563–R567 (2016).
66. G. Janusz, A. Pawlik, J. Sulej, U. Swiderska-Burek, A. Jarosz-Wilkolazka, A. Paszczynski, Lignin degradation: Microorganisms, enzymes involved, genomes analysis and evolution. *FEMS Microbiol. Rev.* **41**, 941–962 (2017).
67. J. M. Robinson, Lignin, land plants, and fungi: Biological evolution affecting phanerozoic oxygen balance. *Geology* **18**, 607 (1990).
68. D. Floudas, M. Binder, R. Riley, K. Barry, R. A. Blanchette, B. Henrissat, A. T. Martínez, R. Otillar, J. W. Spatafora, J. S. Yadav, A. Aerts, I. Benoit, A. Boyd, A. Carlson, A. Copeland, P. M. Coutinho, R. P. de Vries, P. Ferreira, K. Findley, B. Foster, J. Gaskell, D. Glotzer, P. Górecki, J. Heitman, C. Hesse, C. Hori, K. Igarashi, J. A. Jurgens, N. Kallen, P. Kersten, A. Kohler, U. Kües, T. K. A. Kumar, A. Kuo, K. LaButti, L. F. Larrondo, E. Lindquist, A. Ling, V. Lombard, S. Lucas, T. Lundell, R. Martin, D. J. McLaughlin, I. Morgenstern, E. Morin, C. Murat, L. G. Nagy, M. Nolan, R. A. Ohm, A. Patyshakuliyeva, A. Rokas, F. J. Ruiz-Dueñas, G. Sabat, A. Salamov, M. Samejima, J. Schmutz, J. C. Slot, F. St. John, J. Stenlid, H. Sun, S. Sun, K. Syed, A. Tsang, A. Wiebenga, D. Young, A. Pisabarro, D. C. Eastwood, F. Martin, D. Cullen, I. V. Grigoriev, D. S. Hibbett, The paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science* **336**, 1715–1719 (2012).
69. M. P. Nelsen, W. A. DiMichele, S. E. Peters, C. K. Boyce, Delayed fungal evolution did not cause the paleozoic peak in coal production. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 2442–2447 (2016).
70. X. Zhao, Y. Yu, M. E. Clapham, E. Yan, J. Chen, E. A. Jarzembowski, X. Zhao, B. Wang, Early evolution of beetles regulated by the end-permian deforestation. *eLife* **10**, e72692 (2021).
71. D. D. McKenna, S. Shin, D. Ahrens, M. Balke, C. Beza-Beza, D. J. Clarke, A. Donath, H. E. Escalona, F. Friedrich, H. Letsch, S. Liu, D. Maddison, C. Mayer, B. Misof, P. J. Murin, O. Niehuis, R. S. Peters, L. Podsiadlowski, H. Pohl, E. D. Scully, E. V. Yan, X. Zhou, A. Slipinski, R. G. Beutel, The evolution and genomic basis of beetle diversity. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 24729–24737 (2019).
72. S. Seibold, W. Rammer, T. Hothorn, R. Seidl, M. D. Ulyshen, J. Lorz, M. W. Cadotte, D. B. Lindenmayer, Y. P. Adhikari, R. Aragon, S. Bae, P. Baldrian, H. Barimani Varandi, J. Barlow, C. Bassler, J. Beauchene, E. Berenguer, R. S. Bergamin, T. Birkemoe, G. Boros, R. Brandl, H. Brustel, P. J. Burton, Y. T. Cakpo-Tossou, J. Castro, E. Cateau, T. P. Cobb, N. Farwig, R. D. Fernandez, J. Firn, K. S. Gan, G. Gonzalez, M. M. Gossner, J. C. Habel, C. Hebert, C. Heibl, O. Heikkala, A. Hemp, C. Hemp, J. Hjalten, S. Hotes, J. Kouki, T. Lachat, J. Liu, Y. Liu, Y. H. Luo, D. M. Macandog, P. E. Martina, S. A. Mukul, B. Nachin, K. Nisbet, J. O'Halloran, A. Oxbrough, J. N. Pandey, T. Pavlicek, S. M. Pawson, J. S. Rakotondranary, J. B. Ramanamanjato, L. Rossi, J. Schmidl, M. Schulze, S. Seaton, M. J. Stone, N. E. Stork, B. Suran, A. Sverdrup-Thygeson, S. Thorn, G. Thyagarajan, T. J. Wardlaw, W. W. Weisser, S. Yoon, N. Zhang, J. Muller, The contribution of insects to global forest deadwood decomposition. *Nature* **597**, 77–81 (2021).
73. M. Niu, X. Fan, G. Zhuang, Q. Liang, F. Wang, Methane-metabolizing microbial communities in sediments of the haima cold seep area, northwest slope of the south china sea. *FEMS Microbiol. Ecol.* **93**, fix101 (2017).
74. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
75. D. Li, C. M. Liu, R. Luo, K. Sadakane, T. W. Lam, MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de bruijn graph. *Bioinformatics* **31**, 1674–1676 (2015).
76. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
77. J. Hou, S. M. Sievert, Y. Wang, J. S. Seewald, V. P. Natarajan, F. Wang, X. Xiao, Microbial succession during the transition from active to inactive stages of deep-sea hydrothermal vent sulfide chimneys. *Microbiome* **8**, 102 (2020).
78. D. D. Kang, J. Froula, R. Egan, Z. Wang, MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ.* **3**, e1165 (2015).
79. Y.-W. Wu, B. A. Simmons, S. W. Singer, MaxBin 2.0: An automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* **32**, 605–607 (2016).
80. C. M. K. Sieber, A. J. Probst, A. Sharrar, B. C. Thomas, M. Hess, S. G. Tringe, J. F. Banfield, Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nat. Microbiol.* **3**, 836–843 (2018).

81. D. H. Parks, M. Imelfort, C. T. Skennerton, P. Hugenholtz, G. W. Tyson, CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**, 1043–1055 (2015).
82. C. Jain, R. L. Rodriguez, A. M. Phillippy, K. T. Konstantinidis, S. Aluru, High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* **9**, 5114 (2018).
83. S. Capella-Gutiérrez, J. M. Silla-Martínez, T. Gabaldón, trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).
84. A. Stamatakis, RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
85. B. Q. Minh, H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. von Haeseler, R. Lanfear, IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
86. M. J. Pallen, A. Telatin, A. Oren, The next million names for archaea and bacteria. *Trends Microbiol.* **29**, 289–298 (2021).
87. M. Chuvochina, C. Rinke, D. H. Parks, M. S. Rappe, G. W. Tyson, P. Yilmaz, W. B. Whitman, P. Hugenholtz, The importance of designating type material for uncultured taxa. *Syst. Appl. Microbiol.* **42**, 15–21 (2019).
88. International code of nomenclature of prokaryotes. *Int. J. Syst. Evol. Microbiol.* **69**, S1–S111 (2019).
89. X. Feng, X. Chu, Y. Qian, M. W. Henson, V. C. Lanclos, F. Qin, S. Barnes, Y. Zhao, J. C. Thrash, H. Luo, Mechanisms driving genome reduction of a novel roseobacter lineage. *ISME J.* **15**, 3576–3586 (2021).
90. D. Hyatt, G.-L. Chen, P. F. LoCascio, M. L. Land, F. W. Larimer, L. J. Hauser, Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, 119 (2010).
91. M. Kanehisa, Y. Sato, M. Kawashima, M. Furumichi, M. Tanabe, KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44**, D457–D462 (2016).
92. J. Mistry, S. Chuguransky, L. Williams, M. Qureshi, G. A. Salazar, E. L. L. Sonnhammer, S. C. E. Tosatto, L. Paladin, S. Raj, L. J. Richardson, R. D. Finn, A. Bateman, Pfam: The protein families database in 2021. *Nucleic Acids Res.* **49**, D412–D419 (2021).
93. M. Shaffer, M. A. Borton, B. B. McGivern, A. A. Zayed, S. L. La Rosa, L. M. Solden, P. Liu, A. B. Narrowe, J. Rodriguez-Ramos, B. Bolduc, M. C. Gazitua, R. A. Daly, G. J. Smith, D. R. Vik, P. B. Pope, M. B. Sullivan, S. Roux, K. C. Wrighton, DRAM for distilling microbial metabolism to automate the curation of microbiome function. *Nucleic Acids Res.* **48**, 8883–8900 (2020).
94. B. Buchfink, C. Xie, D. H. Huson, Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60 (2015).
95. V. Lombard, H. Golaconda Ramulu, E. Drula, P. M. Coutinho, B. Henrissat, The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42**, D490–D495 (2014).
96. J. Xu, H. Zhang, J. Zheng, P. Dovoedo, Y. Yin, eCAMI: Simultaneous classification and motif identification for enzyme annotation. *Bioinformatics* **36**, 2068–2075 (2020).
97. H. Zhang, T. Yohe, L. Huang, S. Entwistle, P. Wu, Z. Yang, P. K. Busk, Y. Xu, Y. Yin, dbCAN2: A meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* **46**, W95–W101 (2018).
98. N. D. Rawlings, A. J. Barrett, P. D. Thomas, X. Huang, A. Bateman, R. D. Finn, The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Res.* **46**, D624–D632 (2018).
99. J. J. Almagro Armenteros, K. D. Tsirigos, C. K. Sonderby, T. N. Petersen, O. Winther, S. Brunak, G. von Heijne, H. Nielsen, SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* **37**, 420–423 (2019).
100. G. J. Szollosi, W. Rosikiewicz, B. Boussau, E. Tannier, V. Daubin, Efficient exploration of the space of reconciled gene trees. *Syst. Biol.* **62**, 901–912 (2013).
101. Z. Yang, PAML: A program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* **13**, 555–556 (1997).
102. P. P. Sheridan, K. H. Freeman, J. E. Brenchley, Estimated minimal divergence times of the major bacterial and archaeal phyla. *Geomicrobiol. J.* **20**, 1–14 (2003).
103. P. M. Shih, J. Hemp, L. M. Ward, N. J. Matzke, W. W. Fischer, Crown group oxyphotobacteria postdate the rise of oxygen. *Geobiology* **15**, 19–29 (2017).
104. B. Amard, J. Bertrand-Sarfati, Microfossils in 2000 ma old cherty stromatolites of the franceville group, gabon. *Precambrian Res.* **81**, 197–221 (1997).
105. P. C. J. Donoghue, C. J. Harrison, J. Paps, H. Schneider, The evolutionary emergence of land plants. *Curr. Biol.* **31**, R1281–R1298 (2021).

Acknowledgments: We are thankful to P. Hugenholtz for valuable suggestions in GTDB classification and nomenclature and to S. Ishaq for advice in etymology. We thank C. Li, Y. Wang, and Y. Guo from SJTU HPC for useful discussion of software. We are grateful to the researchers who shared sequence data on the NCBI (https://ncbi.nlm.nih.gov/) and to the U.S. Department of Energy Joint Genome Institute (http://jgi.doe.gov/) for providing genome files in collaboration with the scientific community. **Funding:** This work is supported by the National Natural Science Foundation of China [NSFC grant nos. 41921006 (to F.W.), 92251303 (to F.W.), 92051116 (to Y.W.), and 42272354 (to Y.W.)] and Australian Research Council Future Fellowship no. FT210100521 (to B.J.W.). **Author contributions:** Conceptualization: J.H., F.W., and Y.W. Methodology: J.H., Y.W., and B.J.W. Investigation: P.Z., N.Y., T.Y., L.L., and M.N. Visualization: J.H. Supervision: F.W. and Y.W. Writing—original draft: J.H. Writing—review and editing: J.H., F.W., Y.W., and K.K. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The MAGs generated by our laboratory in the current study have been submitted under NCBI BioProject (https://ncbi.nlm.nih.gov/bioproject/) accession number PRJNA869847 and eLMSG (an eLibrary of Microbial Systematics and Genomics; https://biosino.org/elmsg/index) under accession numbers LMSG_G000011401.1 to LMSG_G000011440.1. All MAGs from the GEM (Genome from Earth's Microbiome) catalog are publicly available at https://genome.jgi.doe.gov/portal/GEMs/GEMs.home.html. The specific accession number for each MAG analyzed in this study is provided in table S2.

Submitted 19 November 2022
Accepted 1 June 2023
Published 5 July 2023
10.1126/sciadv.adf5069

Science Advances

Taxonomic and carbon metabolic diversification of Bathyarchaeia during its coevolution history with early Earth surface environment

Jialin Hou, Yinzhao Wang, Pengfei Zhu, Na Yang, Lewen Liang, Tiantian Yu, Mingyang Niu, Kurt Konhauser, Ben J. Woodcroft, and Fengping Wang

Sci. Adv., **9** (27), eadf5069.
DOI: 10.1126/sciadv.adf5069