

Brockarchaeota, a novel archaeal lineage capable of methylotrophy

Valerie De Anda

University of Texas at Austin <https://orcid.org/0000-0001-9775-0737>

Lin-Xing Chen

University of California, Berkeley

Nina Dombrowski

NIOZ, Royal Netherlands Institute for Sea Research, Department of Marine Microbiology and Biogeochemistry, and Utrecht University, P.O. Box 59, NL-1790 AB Den Burg <https://orcid.org/0000-0003-1917-2577>

Zhengshuang Hua

Sun Yat-sen University <https://orcid.org/0000-0002-2405-4228>

Hong-Chen Jiang

China University of Geosciences

Jillian Banfield

University of California, Berkeley <https://orcid.org/0000-0001-8203-8771>

Wen-Jun Li

Sun Yat-Sen University <https://orcid.org/0000-0002-1233-736X>

Brett Baker (✉ acidophile@gmail.com)

The University of Texas at Austin <https://orcid.org/0000-0002-5971-1021>

Article

Keywords: Brockarchaeota, archaeal lineage, anaerobic methylotrophy

Posted Date: July 14th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-39998/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Nature Communications on April 23rd, 2021. See the published version at <https://doi.org/10.1038/s41467-021-22736-6>.

Brockarchaeota, a novel archaeal lineage capable of methylotrophy

Valerie De Anda¹, Lin-Xing Chen², Nina Dombrowski^{1,3}, Zheng-Shuang Hua^{4,5}, Hong-Chen Jiang⁶, Jillian F. Banfield^{2,7}, Wen-Jun Li^{4,8*}, and Brett J. Baker^{1*}

¹ Dept. of Marine Sciences, Marine Science Institute, University of Texas Austin

² Department of Earth and Planetary Sciences, University of California, Berkeley, CA 94720, USA

³ NIOZ, Royal Netherlands Institute for Sea Research, Department of Marine Microbiology and Biogeochemistry, and Utrecht University, Den Burg, Netherlands

⁴ State Key Laboratory of Biocontrol and Guangdong Key Laboratory of Plant Resources, School of Life Sciences, Sun Yat-Sen University, Guangzhou 510275, People's Republic of China

⁵ Department of Biological Sciences, Dartmouth College, Hanover, NH, 03755, USA

⁶ State Key Laboratory of Biogeology and Environmental Geology, China University of Geosciences, Wuhan 430074, People's Republic of China

⁷ Department of Environmental Science, Policy and Management, University of California, Berkeley, CA 94720, USA

⁸ Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai 519000, People's Republic of China

*Corresponding authors emails acidophile@gmail.com and liwenjun3@mail.sysu.edu.cn

Single carbon (C1) compounds such as methanol, methylamines and formaldehyde are ubiquitous in nature and they are large components of the carbon cycle. In anoxic environments C1-utilizing microbes (methylotrophs) play an important role in controlling global carbon degradation. Currently described anaerobic methylotrophs are limited to methanogenic archaea, acetogenic bacteria, and sulfate-reducing bacteria. Here, we report the first archaeal lineage outside of methanogenic taxa that are capable of anaerobic methylotrophy. Phylogenetic analyses suggest these archaea form a new phylum within the TACK superphylum, which we propose be named Brockarchaeota. A survey revealed Brockarchaeota are globally distributed in geothermal springs. Metabolic inference from 15 metagenome-assembled genomes from hot springs and deep-sea sediments indicates that Brockarchaeota are strict anaerobes. They contain a variety C1 metabolisms including the methanol and trimethylamine methyltransferases system, the ribulose biphosphate pathway coupled with the non-oxidative pentoses phosphate pathway, and reductive glycine pathway. Brockarchaeota have an incomplete methyl-branch of the Wood-Ljungdahl pathway probably used for formaldehyde oxidation, since they lack several core genes involved in methanogenesis including methyl-CoM reductases. Brockarchaeota also appear to play an important role in the breakdown of plant-derived polysaccharides, especially cellulose, starch and xylan. Their broad distribution and their capacity to use both C1 compounds and complex polysaccharides via anaerobic metabolism suggest that the Brockarchaeota occupy previously overlooked roles in carbon cycling.

Introduction

Methylotrophs are organisms that are capable of using single-carbon (C1) compounds as a source of energy and carbon^{1,2}. C1 compounds, such as methanol and methylamines, are derived from a variety of sources such as phytoplankton, plants, and the decay of organic matter³⁻⁵. As a result, they are ubiquitous in oceans and atmosphere and are important components of the global carbon and nitrogen cycles⁴. In oxic environments, C1 compounds are generally converted to formaldehyde by methanol dehydrogenases (MDH) found in aerobic methylotrophs^{3,4}. In anoxic

settings, C1 compounds are used as substrate for methylotrophic-methanogenesis⁶⁻⁹ and sulfate reduction¹⁰. Anaerobic methylotrophs utilize the methyltransferase system (MT) to break and transfer the methyl residue of C1 compounds to coenzyme M (in the case of methanogens) or tetrahydrofolate (in acetogens and sulfate reducers)⁶⁻¹⁰ and couple this reaction to the Wood–Ljungdahl pathway (WLP) for energy conservation. Methylotrophic archaea include methanogenic orders (in Euryarchaeota): *Methanosarcinales*, *Methanobacteriales*, *Methanomassiliicoccales* and the recently discovered uncultured methylotrophic phylum, *Verstraetearchaeota*⁶. Methylotrophy has not been described in archaeal lineages outside of these methanogenic groups. Also, our understanding of C1-utilization is limited, as there is increasing evidence that suggests C1 compounds serve as energetic substrates to fuel non-methanogenic heterotrophic communities in the deep seafloor¹¹ and marine sediments¹³. However, little is known about the microorganisms or pathways mediating this process⁴. Here we characterize a new globally distributed archaeal lineages with metabolic pathways for C1 utilization from hot springs and the deep ocean.

Results

Genomic reconstruction

Metagenomic sequencing, assembly, and binning of sediments from seven terrestrial hot springs in Tibet (up to 70°C) and Tengchong Yunnan, China (up to 86 °C) and deep sea hydrothermally-heated Guaymas Basin (GB) sediments (10-34°C) resulted in the reconstruction of 15 draft metagenome-assembled genomes (MAGs) estimated to be 67-92% complete (Table 1). These MAGs range from 0.94 to 2.90 Mbp (average 1.85 Mbp) (Table 1). The two MAGs from GB (B48_G17 and B27_G9, temperature 33.6 and 10.4°C respectively) were originally designated as “CP5” in a prior study¹. Although the GB genomes were obtained from lower temperatures these sediments experience increases in temperature due to hydrothermal circulation¹². Thus, the organisms from which these genomes were derived likely prefer hot geothermal ecosystems, and anoxic conditions (Supporting Table 1).

Phylogeny and distribution in nature

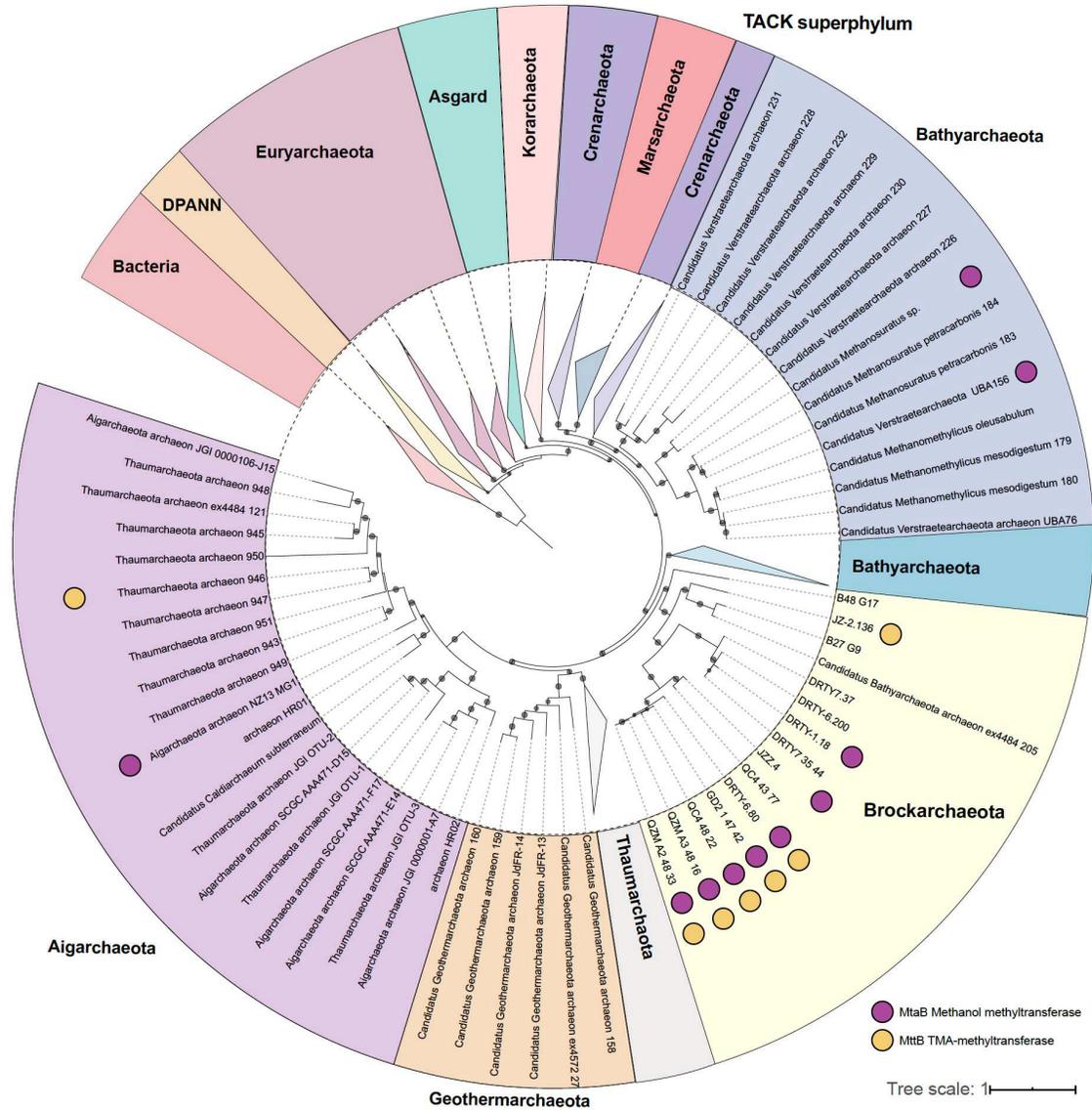
Phylogenetic analyses of these MAGs based on a concatenated alignment of 37 conserved proteins (see Methods), revealed they form a distinct group within the TACK superphylum, basal to *Aigarchaeota* and *Thaumarchaeota* (Figure 1). A comparison of average amino acid identities (AAI) across 250 available TACK genomes (Supporting Table 2) including the recently discovered phylum *Marsarchaeota*¹³, revealed that Brockarchaeota are distinct from neighboring phyla (*Geoarchaeota*, *Aigarchaeota* and *Thaumarchaeota*) and share up to 99% genome-wide nucleotide similarity to one another (Supporting Figure 1 and Supporting Table 3). The two GB MAGs (B48_G17 and B27_G9) are distinct from the hot springs at the AAI level (<50% similar to each other), and <45% AAI to members of *Geoarchaeota*, *Thaumarchaeota* and *Aigarchaeota*, which is consistent with their phylogenetic placement. Phylogeny of 16S rRNA genes also indicated that Brockarchaeota do not fall within any described archaeal phyla (Figure 2A), with < 78% similarity to other TACK members. Together, these results support the classification of these MAGs as a new phylum. We propose that the phylum be named “Brockarchaeota”, after Thomas Brock, an American microbiologist known for his groundbreaking research in hot springs microbiology.

Table 1. General information on the fifteen novel Brockarchaeota MAGs.

Genome	Origin	Size (Mb)	Scaffolds (number)	Protein coding genes (number)	Predicted size (Mb)	GC (%)	Compl. (%)	Red. (%)	Largest scaffold (Mb)	Strain hetero
B48_G17	Deep sea sediment (12-15 cm, 33.6 °C) Vent 2 in Dombrowski <i>et al</i> 2018	0.79	113	925	0.94	51	84	4.05	49,468	14.29
B27_G9	Deep sea sediment (0-3 cm, 10.4 °C) Vent 2 in Dombrowski <i>et al</i> 2018	0.99	223	1341	1.48	42	67	1.46	17,768	0
DRTY7	Hot spring sediment DiReTiYanQu-7 (collected in Jan, 2016) in Tengchong county, Yunnan, China (55.8 °C)	0.83	133	910	0.99	34	84	1.94	23,553	0
QC4_43	Hot spring QuCai village, Tibet, China (69.5 °C)	1.18	36	1332	1.48	42	80	2.43	135,049	0
QC4_48	Hot spring QuCai village, Tibet, China (69.5 °C)	1.67	87	1807	1.99	48	84	9.06	78,542	12.5
GD2_1	Hot spring GuDui geothermal area, Tibet, China (61.8 °C)	2.14	156	2326	2.40	48	89	1.94	82,330	0
QZM_A2	Hot spring QuZhuoMu village, Tibet, China (63.1 °C)	2.19	173	2447	2.67	48	82	1.94	96,973	0
QZM_A3	Hot spring QuZhuoMu village, Tibet, China (62.9 °C)	1.96	218	2281	2.51	47	78	5.83	70,810	0
DRTY-1.18	Hot spring sediment DiReTiYanQu-1 (collected in May, 2017) in Tengchong county, Yunnan, China (67 °C)	1.62	55	1582	1.84	34	88	1.94	117,455	0
DRTY-6.200	Hot spring sediment DiReTiYanQu-6 (collected in May, 2017) in Tengchong county, Yunnan, China (60 °C)	0.55	129	663	0.96	34	57	0	11,975	0
DRTY-6.80	Hot spring sediment DiReTiYanQu-6 (collected in May, 2017) in Tengchong county, Yunnan, China (60 °C)	2.32	55	2264	2.67	47	87	0	266,602	0
DRTY-7.37	Hot spring sediment DiReTiYanQu-7 (collected in Jan, 2016) in Tengchong county, Yunnan, China (55.8 °C)	1.03	189	1131	1.18	35	87	0.97	23,662	0
JZ-1.89	Hot spring sediment Jinze-1 (collected in May, 2017) in Tengchong county, Yunnan, China (86.5 °C)	1.48	88	1680	2.90	41	51	0.97	147,883	0
JZ-2.136	Hot spring sediment Jinze-2 (collected in May, 2017) in Tengchong county, Yunnan, China (63 °C)	2.04	215	2234	2.43	58	84	0.97	40,656	0
JZ-2.4	Hot spring sediment Jinze-2 (collected in Jan, 2016) in Tengchong county, Yunnan, China (75 °C)	1.30	61	1412	1.41	40	92	0	130,091	0

105

Genome ID, origin, number of scaffolds, number of protein-coding genes, guanine-cytosine (GC) content, estimated completeness (Compl.), estimated gene redundancy (Red.), and strain heterogeneity (Strain hetero.) are shown.



110 **Figure 1. Comparison of phylogeny and distribution of methyltransferase system in Brockarchaeota**
and other members of TACK superphylum. Phylogeny generated using iqtree1.6.1 using a concatenation
of 37 conserved single-copy protein-coding genes described in ref¹⁴. Bootstrap values were calculated using
non-parametric bootstrapping with 100 replicates (represented by gray circles, only bootstrap >70 are
shown). The presence methanol methyltransferase MtaB (PF12176) and trimethylamine methyltransferase
115 MttB (PF06253) are shown in the outer circles. The annotation was conducted with MEBS¹⁵ details can be found
in Supporting Table 5.

120 Interestingly, only three 16S rRNA gene sequences with similarity (92-96%) to
Brockarchaeota sequences have been described in PCR-based surveys, highlighting the inherent
bias for primer choice in diversity studies. Therefore, we searched publicly available
metagenomic databases to examine the geographic distribution of this phylum. Notably, we
almost exclusively found 16S rRNA gene sequences related to Brockarchaeota in sequence data
generated from other hot springs from around the world (China, USA, South Africa; Figure 2A),
revealing Brockarchaeota are globally distributed in hot springs (Figure 2B). Three sequences,
125 which cluster together, were recovered from lake sediments in Rwanda and the Gulf of Boni in

Indonesia (28°C) (see Supporting Table 4), suggesting that some Brockarchaeota are mesophilic as well.

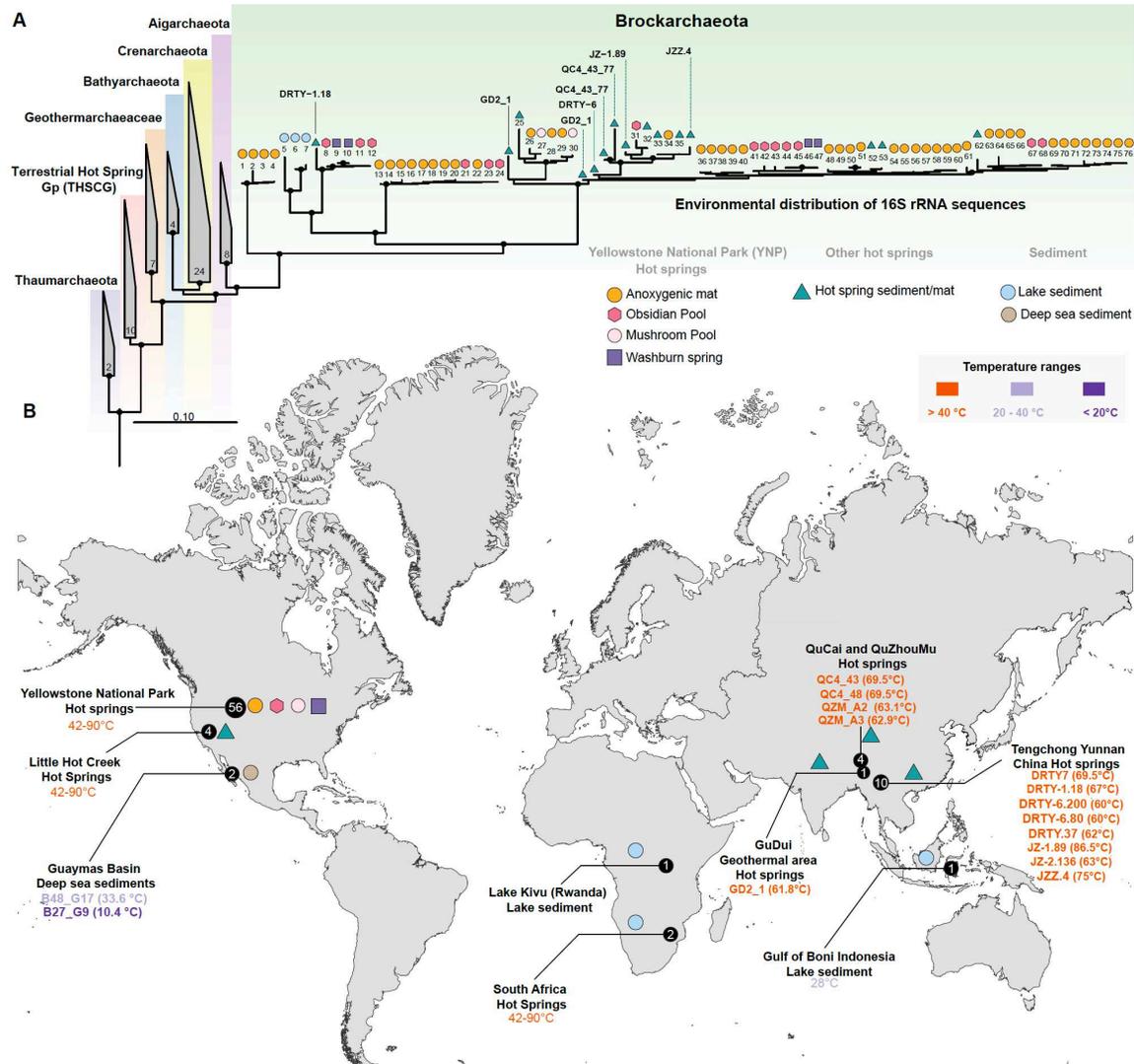


Figure 2. Location of samples from which Brockarchaeota genomes and 16S rRNA gene sequences have been recovered. (A) 16S rRNA gene tree of sequences derived from metagenomic and rRNA-based diversity surveys (NCBI accessions EU924237, KX213943, and KX213897). The eight complete 16S rRNA gene sequences of Brockarchaeota genomes described in this study are shown in their respective names. Black circles in the tree represent 100 bootstrap values using RAXML with 100 replicates. Environmental information of each sequence is shown as obtained from Integrated Microbial Genomes and Microbiomes database. The number of the sequences and the corresponding metadata are described in Supporting Table 4. (B) Geographic localization from which Brockarchaeota sequences were obtained. The size of the circle corresponds to the total number of Brockarchaeota-related sequences in each geographic location. The specific MAGs obtained in this study are shown in the map according to their temperature range.

Utilization of C1 compounds and central carbon metabolism

To begin to understand the metabolism of the Brockarchaeota we compared the predicted proteins encoded by these genomes with a variety of functional databases (see Methods). This revealed a unique type of anaerobic methylotrophic metabolism. They contain the methyltransferase system (MT), that has been shown to be essential for anaerobic methylotrophy¹⁶ and is composed of three components. The first is a methyltransferase involved in breaking the C-O bond of specific methylated compounds (MtaB for methanol, MtmB for monomethylamine, MtbB for dimethylamine, and MttB for trimethylamine). The second is the transferring methyl residue to the second component, a corrinoid protein (Figure 3). The third is a methylated methylamine-specific corrinoid protein (MtaA for methanol, MtbA for methylamines), that is involved in transferring the methyl-group from the corrinoid protein to coenzyme M in methanogens, or tetrahydrofolate in acetogens. Brockarchaeota from hot springs encode proteins predicted to be methanol-CoM methyltransferases (MtaB) and trimethylamine-corrinoid protein methyltransferase (MttB) for the utilization of methanol and trimethylamine (TMA), respectively. Brockarchaeota also encode a putative B12-binding corrinoid protein, that once methylated, can act as a substrate for the third component of the MT system that is lacking in Brockarchaeota genomes. Another undescribed protein may be involved in the transfer of the methylated compound from the corrinoid protein to unknown methyl carrier (See details in Supporting Table 5). The distribution of methyltransferases among a current set of TACK superphylum genomes indicate methanol-MT system is unique feature of Brockarchaeota within the TACK superphylum (Figure 1). Prior to this, methanol-utilization is currently thought to be limited to Euryarchaeota and TACK (Verstraetearchaeota⁶, Korarchaeota¹⁷) archaea and some bacteria (Firmicutes and Deltaproteobacteria)¹⁸.

To our knowledge, no methylotrophic members of the archaea domain have been described outside methanogenic groups (Euryarchaeota and Verstraetearchaeota). Despite the presence of MT system, unlike phylogenetically related archaea such as Verstraetearchaeota⁶; Brockarchaeota genomes do not possess the common core marker genes specific to methanogenesis including Methyl-coenzyme M reductase (MCR) (Figure 4 and Supporting Table 6). To ensure that the MCRs are not missing due to incomplete MAGs, we searched metagenomic datasets from each of communities from which they were obtained for *mcr* genes belonging to Brockarchaeota and did not find any (See supporting Table 7 and Supporting Discussion). In addition to lacking MCR genes, they encode lack a complete Wood-Ljungdahl pathway essential for substrate utilization an energy conservation in methanogenic archaea. Furthermore, Brockarchaeota lack a methanol dehydrogenase (MDH) system that has been recently described in deep-sea sulfate reducing bacteria encoding methyltransferases¹⁰.

The lack of previously described anaerobic methylotrophic pathways, raises the question of how C1 compounds are assimilated by Brockarchaeota, or if they are being used for energy conservation or biosynthetic purposes. Most of the Brockarchaeota encode all four enzymes of the non-oxidative pentoses phosphate pathway (NOPPP) and both key enzymes of the ribulose monophosphate pathway (RuMP) including 3-hexulose-6-phosphate synthases (HPS) and 6-phospho-3-hexuloisomerases (PHI) (Supporting Figure 2). The presence of complete NOPPP, RuMP and reductive glycine pathway (rGlyP) suggests that these compounds are being assimilated into pyruvate, which can be converted to acetate for ATP formation at substrate-level phosphorylation.

The RuMP pathway was originally described in methylotrophic bacteria, which use C1 compounds as a sole source of carbon and energy, however, it is now recognized as a widespread pathway for formaldehyde fixation and detoxification¹⁹. Formaldehyde was an essential building block to synthesize sugars on early Earth and is ubiquitous in nature, produced through the degradation of compounds containing methyl- or methoxyl-groups, e.g., lignin and pectin²⁰. In geothermally environments autotrophic microbes are thought to produce formaldehyde from a

variety of C1 compounds (i.e. CH₄, CO₂, CO)²¹. The RuMP pathway functions as an efficient system for trapping free formaldehyde at relatively low concentrations. The presence of HPS and PHI in Brockarchaeota suggests that formaldehyde can be fixed and detoxified via the RuMP pathway. Furthermore, formaldehyde can potentially be oxidized to CO₂ or formate due to the presence of genes coding tungsten-dependent aldehyde ferredoxin oxidoreductase (AFOR). However, this enzyme can oxidize a wide range of different aldehydes derived from organic carbon degradation (e.g., of peptides) and may also be involved in electron transport reactions²². The presence of potential oxidative and assimilative formaldehyde pathways in Brockarchaeota may be an advantage for their survival in hot springs and hydrothermal deep-sea environments

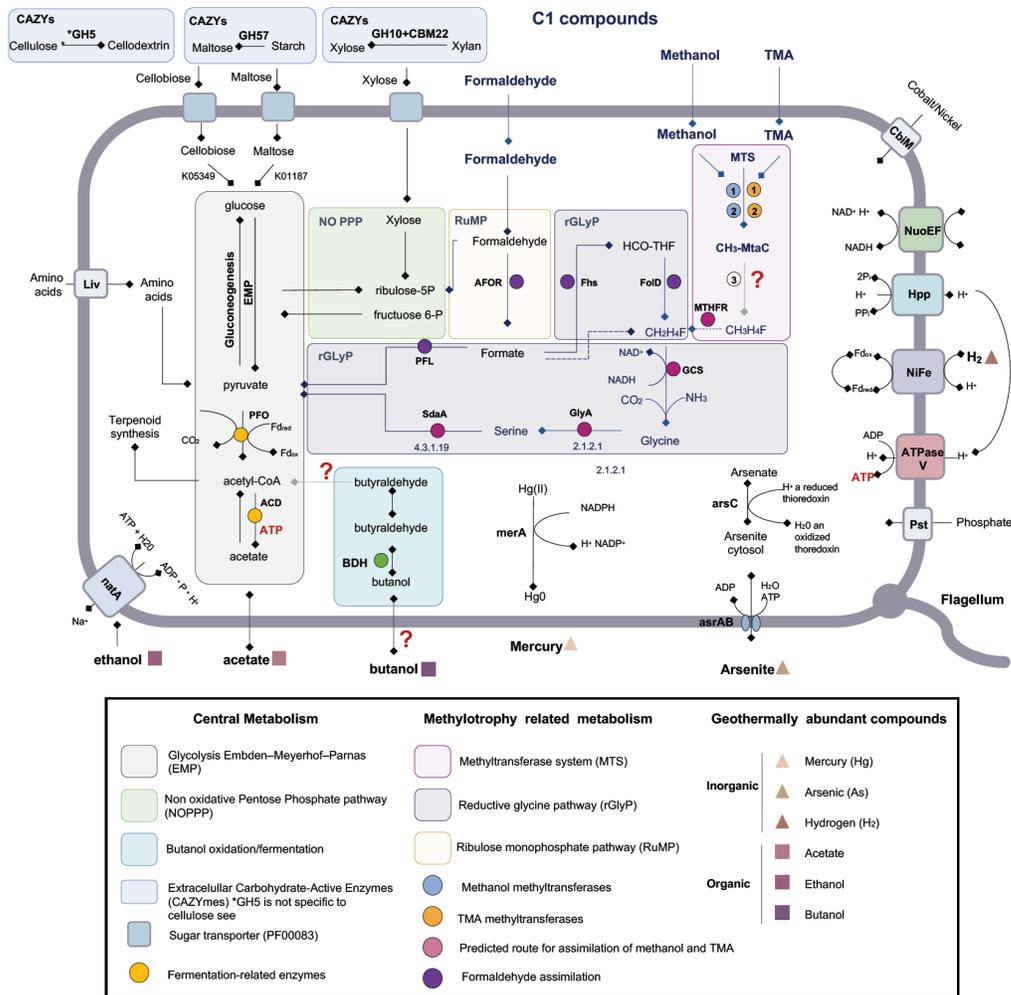


Figure 3. Overview of potential metabolic capabilities of Brockarchaeota phylum. Brockarchaeota is predicted to produce ATP by substrate-level phosphorylation by fermentation of complex organic compounds, pyruvate and acetate as carbon and energy sources via glycolysis (EMP) by the concerted action of PFO and ACD, that might represent the major energy-conserving reaction in this lineage. Formaldehyde can be fixed and detoxified via the ribulose monophosphate (RuMP) pathway and enter to central metabolism or be oxidized to formate by AFOR which can enter to the rGLyP pathway via FhS and FoID action C1 compounds such as methanol and TMA can be assimilated via MT system and MTHFR, generating glycine, and following a linear route to central metabolism through rGLyP by subsequent conversion of serine and pyruvate. Geothermally abundant compounds are showed in colored circles Each shaded pathway can be seen in more detail in Supporting Discussion. Dashed arrow indicates that spontaneous condensation of

formaldehyde with THF occur but is a very minor contribution to assimilation fluxes according to ref³⁶. For full names and copy numbers of the genes in number see Supporting Discussion. Abbreviations. Pathways: Embden-Meyerhof-Parnas (EMP), Non-Oxidative Pentoses Phosphate Pathway (NOPPP), ribulose monophosphate (RuMP), reductive glycine pathway (rGlyP). Enzymes: pyruvate ferredoxin oxidoreductase (PFO), and acetate-CoA ligase (ADP-forming) (ACD), pyruvate formate lyase (PFL), tungsten-dependent aldehyde ferredoxin oxidoreductase (AFOR), tetrahydrofolate (THF) ligase (FhS), methenyl-THF cyclohydrolase/methylene-THF dehydrogenase (FolD), methylene-THF reductase MTHFR (MetMethyltransferase system (MT), *glyA*, serine hydroxymethyltransferase (GlyA), serine deaminase (SdaA), glycine-cleavage system (GCS). Compounds: Single carbon compounds (C1), trimethylamine (TMA). The three main components of the methyltransferase system are explained in the main text.

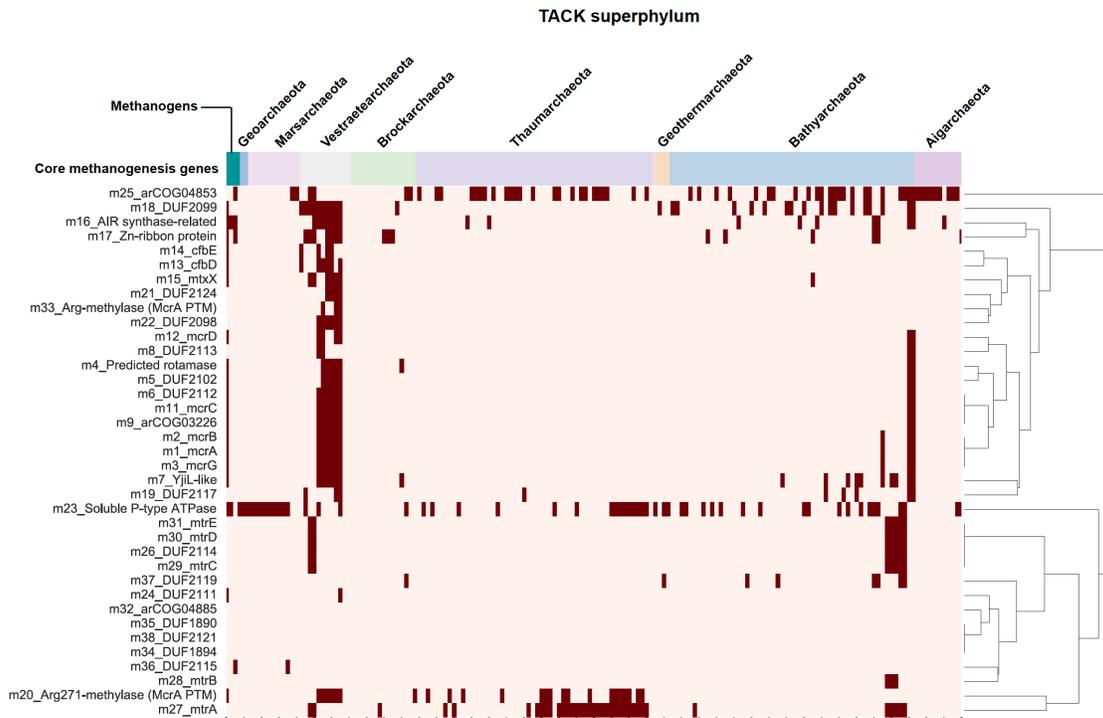


Figure 4. Presence and absence profile of common core marker genes specific to methanogenesis across the TACK superphylum. Genes include those associated to methanogenesis, anaerobic methanotrophic and short-chain alkane-oxidizing archaea described in Greening et al. 2016³⁴ (see Supporting Table 6 for specific details of the genomes used in this study and the specific marker genes named with m followed with a number in the figure). For comparison purposes the following known methanogenic archaea were included: *Candidatus* Methanoplasma termitum that lacks the entire pathway for CO₂ reduction to methyl coenzyme M and produces methane by hydrogen-dependent reduction of methanol or methylamine (Methanomassiliicoccales), *Methanosphaera stadtmanae* (Methanobacteriales) that can generate methane only by the reduction of methanol with H₂ and is dependent on acetate as a carbon source, and *Methanosarcina acetivorans* (Methanosarcinales).

Some Brockarchaeota genomes code tetrahydrofolate (THF) ligase (*fhs*), methenyl-THF cyclohydrolase/methylene-THF dehydrogenase (*folD*), and methylene-THF reductase MTHFR (*metF*), which are key enzymes of the methyl-branch of the Wood–Ljungdahl pathway (Supporting Table 5). Nonetheless, given the absence of key carbonyl branch genes, Brockarchaeota appears to be incapable of assimilating C1 compounds via the WL pathway used by other methylotrophic archaea. Yet, the presence of genes for the rGLYP pathway (*gcvP*, glycine dehydrogenase; *gcvT*, aminomethyltransferase; *gcvH*, lipoate-binding protein; *lpd*, dihydrolipoyl dehydrogenase; *glyA*,

245 serine hydroxymethyltransferase; and *sdaA*, serine deaminase) in Brockarchaeota genomes, offers another alternative route for assimilation of C1 compounds not yet described in methylotrophic archaea. Subsequent conversion of C1 compounds in Brockarchaeota can proceed via methylene-THF derivatives, either by MTHFR or FoldD (in the case of formaldehyde) to generate 5,10-CH₂-tetrahydrofolate which can be assimilated via glycine-cleavage system (GCS)²³, followed by the assimilation into central metabolism with the conversion of serine into pyruvate (see figure 3).

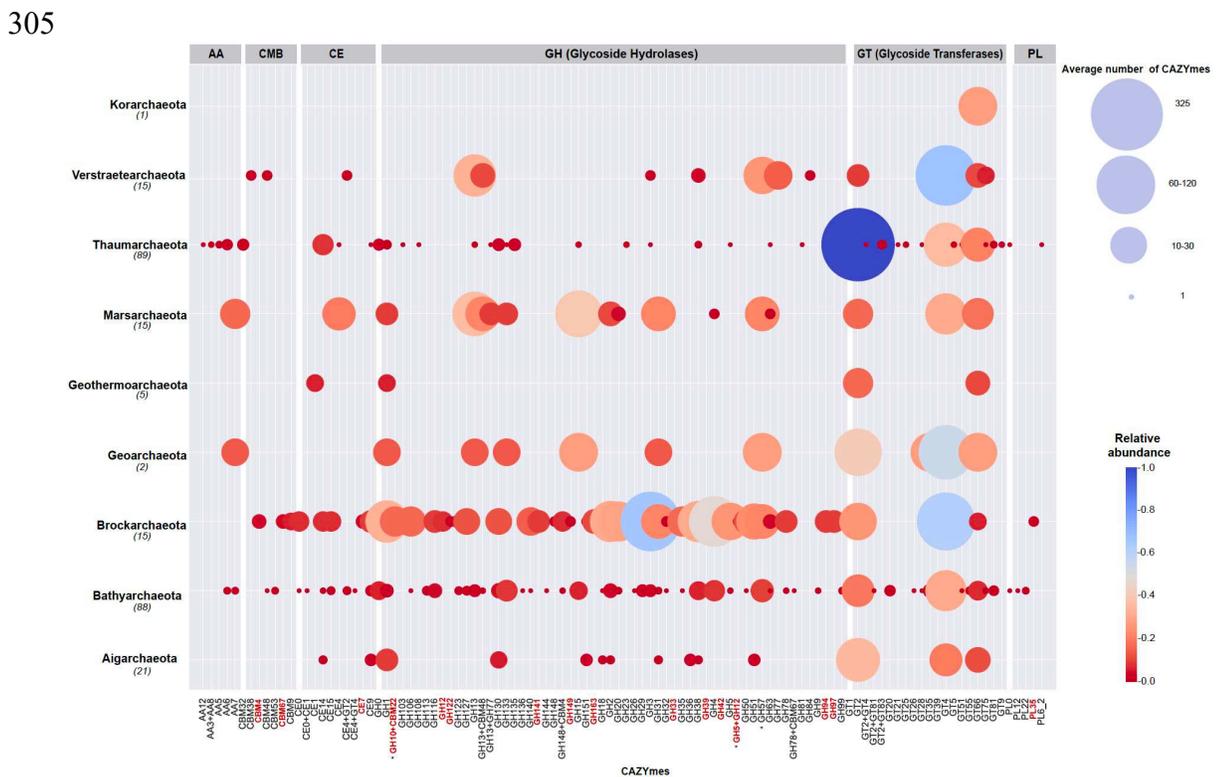
250 A phylogenetic analysis of alcohol dehydrogenases from hot spring genomes revealed that they encode a butanol dehydrogenase BDH (Supporting Figure 3) that catalyzes the reversible conversion of butyraldehyde to butanol. Brockarchaeota BDH's are homologues to sequences from obligately anaerobic, thermophilic bacteria that can degrade complex plant saccharides such as xylan (i.e *Caldicoprobacter oshimai*²⁴ and *Hungateiclostridium thermocellum*²⁵) or cellulose
255 (*Hungateiclostridium alkalicellulosi*). To investigate if Brockarchaeota can oxidize or produce butanol, we searched for genes involved in production of butanol in two model organisms; *Clostridium acetobutylicum*^{26,27} which is one of the few organisms that produces butanol as a fermentation product, and *Saccharomyces cerevisiae*²⁸ involved in butanol and isopropanol production. We found that Brockarchaeota genomes lack the key enzymes involved in the
260 fermentation of pyruvate to butanol (butanal dehydrogenase, butyryl-CoA dehydrogenase, enoyl-CoA dehydratase, 3-hydroxyacyl-CoA dehydrogenase). However, most of the genomes code a putative aldehyde dehydrogenase that could convert butyraldehyde to butyric acid. Also, we found a putative enoyl-CoA hydratase/isomerase protein that is coded by one bin (JZ-1.89), which could be involved in further converting butyric acid to acetyl-CoA. Our results suggest an alternative
265 pathway for butanol oxidation that still remains unresolved (Supporting Figure 4).

Pathways for the utilization of extracellular organic carbon

In addition to anaerobic methylotrophy, Brockarchaeota may be able to degrade a variety of organic carbon compounds. They may utilize hexoses via Embden-Meyerhof-Parnas (EMP)
270 pathway (Supporting Figure 5) and pentoses (xylose isomerase *xylA* and xylulose kinase *xylB*) via the isomerase pathway (Supporting Figure 2). These enzymes were previously only found in bacterial thermophiles and halophilic archaea that ferment complex compounds and degrade xylose suggesting a similar physiology in Brockarchaeota²⁹. Once assimilated into the cell, carbon complex compounds could enter the central metabolism and be converted to acetate and H₂ via
275 acetogenic fermentation. The ATP conserving step of either sugar or pyruvate fermentation to acetate could be catalyzed by acetate-CoA ligase in the hot spring genomes. Acetate can also be assimilated to acetyl-CoA by acetyl-CoA synthetase (ACS), thus acetate might be a source of carbon and energy in the absence of other substrates in hot spring Brockarchaeota. The presence of pyruvate ferredoxin oxidoreductase (PFO) that couples pyruvate oxidation to H₂ production,
280 generating acetyl-CoA, could support fermentative metabolism via degradation of either acetate, pyruvate, hexoses or pentoses. Brockarchaeota genomes appear to code a wide repertoire of ATPases such as the plasma-membrane proton-efflux P-type ATPase (only present in the hot spring genomes), Zn²⁺/Cd²⁺-exporting ATPase (present in DRTY7.37), and finally the V/A-type H⁺/Na⁺-transporting ATPase (in most of the genomes). The existence of ATPase in Brockarchaeota suggests that members of these genotypes have the additional ability to couple acetogenic
285 fermentation to membrane potential generation of a transmembrane ion gradient across the membrane.

To complement their ability to degrade xylenes, Brockarchaeota also contain a relatively high number of carbohydrate-active enzymes (average of 27 CAZYmes per genome) which is 3
290 times what has been observed in other TACK archaea (Supporting Table 8 and Figure 5). Ten of the 15 Brockarchaeota genomes have genes with similarity to α -L-fucosidase involved in the degradation of xyloglucan, which is the major component of hemicellulose in plant-cell walls³⁰. All the hot spring genomes encode GH3 family proteins for detrital cellulose degradation, plant

295 and bacterial cell wall remodeling, energy metabolism, and pathogen defense³¹. The hot spring
 300 genotypes contain a wider repertoire of CAZymes than the deep-sea GB genomes. Among these
 are four predicted to be extracellular glycoside hydrolases, which are involved in the breakdown of
 high molecular-weight plant-derived polysaccharides, primarily xylanes, cellulose, and starch.
 Comparison of the CAZymes across the TACK superphylum revealed that 17 extracellular
 enzymes, including enzymes for the degradation xylanes, are unique to Brockarchaeota (Supporting
 Table 8). The diversity and abundance of CAZymes in members of the TACK superphylum
 highlights that that despite the low number of sequenced Brockarchaeota genomes (15 described
 in this study) compared to Thaumarchaeota (89), the former encodes a wider arrange and in wider
 abundance that their phylogenetically related counterparts.



310 **Figure 5. Annotated carbohydrate-active enzymes (CAZymes) encoded by the total number of genomes belonging to the TACK superphylum including Brockarchaeota.** The total number of CAZymes per phylum was normalized by the total number of genomes described for each phylum shown in parenthesis. Auxiliary activities (AA), carbohydrate-binding module (CBM), carbohydrate esterases (CE), glycoside-hydrolases (GH), glycoside transferases (GT), polysaccharide lyases (PL). Unique Brockarchaeota CAZymes are shown in red. Extracellular CAZymes in Brockarchaeota are shown in asterisks described in Supporting Table 8 sheet 3.

315 **Hydrogen metabolism**

320 Hydrogen is also abundant in hydrothermal systems due to volcanic processes³². Brockarchaeota might be able to use 3b [NiFe]-hydrogenases for H₂ oxidation with NADP⁺ or NAD(P)⁺ as an electron acceptor³³. The hot spring genomes also encode oxygen-tolerant group 3d [NiFe]-hydrogenases, which may allow them to transfer electrons between NAD(P)H and H₂ depending on the availability of electron acceptors (Supporting Figure 6). Group 3d [NiFe]-hydrogenases are abundant in metagenomes from hot springs where microbial communities are relatively stable

despite partial pressure of oxygen fluctuations³⁴. Group [NiFe] 3b hydrogenases may also make it possible for these archaea to reduce elemental sulfur to H₂S during fermentative growth. During carbohydrate fermentation in the absence of sulfur, Group 3b [NiFe]-hydrogenases might catalyze the production of H₂ with NADPH or NAD(P)H as the electron donor. Therefore, Brockarchaeota might have the ability to reduce sulfur, using H₂ or organic substrates as electron donors, which is common in hyperthermophilic archaea living in geothermally active environments³⁵

330 Discussion

Brockarchaeota gene contents suggests they are facultative or obligate anaerobic fermentative organisms that produce acetate, CO₂, and H₂ as byproducts (see Supporting Information for details). Brockarchaeota are also able to degrade complex carbon compounds such as xylenes and perform anaerobic methylotrophy. Xylenes are a major structural polysaccharide in plant cells, and is the second most abundant polysaccharide in nature, accounting for approximately one-third of all renewable organic carbon on Earth after cellulose^{37,38}. This suggests that Brockarchaeota are key players in organic matter degradation in hot springs and deep-sea sediments. Furthermore, Brockarchaeota have unique pathways for non-methanogenic methylotrophy which has not been described in the archaea domain prior to this and has not been attributed to any organism broadly distributed in nature. This gives them a unique ecological position in nature, to degrade abundant methylamines in anoxic environments (Figure 6).

The protein repertoire of GB and hot springs genomes have some important distinctions that likely reflects different anaerobic physiologies. GB genomes appear to be obligately fermenting organisms that rely mostly on substrate-level phosphorylation since they lack all the complexes for the respiratory chain with exception of the ATPase. In contrast, hot spring genomes appear to have mechanisms to increase their ATP yield including the use of geothermally derived inorganic substrates as possible terminal electron acceptors such as mercury (Hg), arsenic (As) and hydrogen (H₂). Deep-sea hydrothermal vents, hot springs, and fumaroles are natural sources of Hg³⁹, H₂³⁴ Arsenic⁴⁰ and sulfur⁴¹. Three hot springs Brockarchaeota genomes (DRTY735_44, DRTY-1.18 and DRTY.37) encode mercuric reductase (MerA), which detoxifies Hg (II) to Hg(0)³⁹. Brockarchaeota also appear to have the genetic capacity for As-resistance including arsenate reductase (ArsC) for the reduction of arsenate to arsenite, which is present in most of the hot spring genomes, and arsenite efflux transporters (*asrA* and *asrB*) for cell export (Figure 3). The presence of this energy-dependent efflux process related detoxification proteins, could also indicate that Brockarchaeota in hot springs genomes could use arsenate as terminal electron acceptor, as seen in bacteria^{42,43}

The discovery of Brockarchaeota genomes from sediments around the world, overlooked by conventional rRNA gene diversity approaches, highlights the need for further exploration of subsurface microbial communities. The addition of these genomes to public databases, like other recently described novel archaeal lineages⁴⁴⁻⁴⁶, will enhance their detection in future environmental studies. A lack of recognition of their existence prior to this, limited our ability to fully describe sediment community structure and function. Given their broad distribution, and versatile carbon metabolism, they are likely key players in global carbon cycling. However, this first description is limited to genomic characterization, thus culturing or in activity measurements are needed to confirm their physiological activities⁴⁷. Overall, the description of this new phylum enhances our understanding of biodiversity of archaea and suggests they are mediating unique roles in anoxic carbon cycling.

370

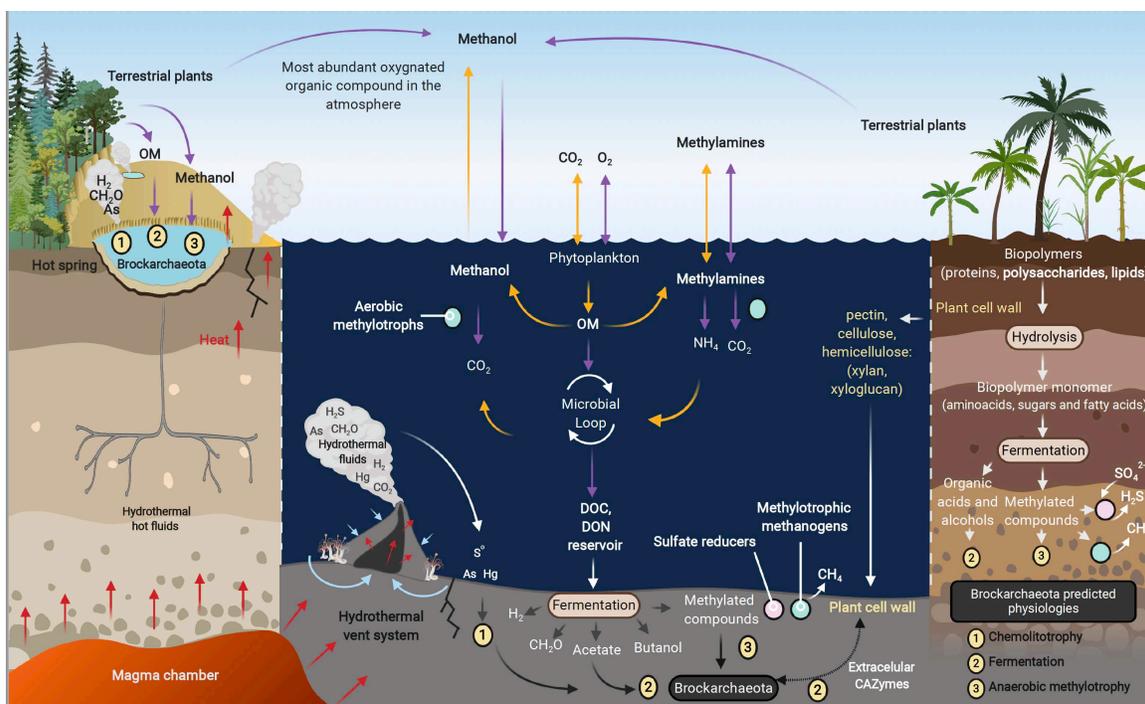


Figure 6. The role of Brockarchaeota in the anaerobic carbon cycle. Single carbon (C1)-methylated compounds, such as methanol or methylamines, are utilized biologically as carbon and energy sources in the ocean and deep-sea sediments resulting in a considerable carbon reservoir. The biodegradation of organic carbon in the water column and subsurface is a source of C1-methylated compounds. The utilization of methyl compounds as precursors in methane synthesis is confined to a small group of methylotrophic methanogens (i.e. Verstraetearchaeota). The only described anaerobic methylotrophs include members of methanogenic archaea, acetogenic bacteria, and sulfate-reducing bacteria. These organisms compete for C1 compounds geochemically produced in anoxic settings. Brockarchaeota may recycle C1 in anoxic environments without methane formation and may be sequestered in deep sea sediments and hot springs. Orange and purple arrows represent sources and sinks, respectively. Organic Matter (OM) includes dissolved and particulate organic matter feeding the microbial loop (Adapted from Evans et al., 2019 and Zhuang et al., 2018).

Methods

Metagenomic assembly and binning. Two MAGs (B48_G17 and B27_G9) were obtained from Guaymas Basin sediments (Gulf of California; 27°N0.388, 111°W24.560) and were obtained as part of a larger study of these hydrothermal marine sediments¹². Both samples were collected from the same location but G9 was sampled from 0-3 cm and G17 from 12-15 cm depth. The sediment cores from which these two MAGs were binned from were collected during Alvin dive 4571_4 in 2009 using polycarbonate cores (45-60 cm in length, 6.25 cm interior diameter), subsampled into 395 cm layers under N₂ gas in the ship's laboratory and immediately frozen at -80°C. Details on the sampling site and metagenomic sequencing effort is provided in Dombrowski et al., 2018.

Briefly, total DNA from ≥10 g of sediment from each sample was extracted using the MoBio PowerMax soil kit using the manufacturer's instructions and adjusted to a final concentration of 10 ng/μl of each sample (using a total amount of 100 ng). Libraries for paired-end Illumina (HiSeq-2500 1TB) sequencing were prepared by the Joint Genome Institute (JGI). Sequencing was performed on an Illumina HiSeq 2500 machine using the paired-end 2x125 bp run-

type mode. All runs combined provided a total of ~280 gigabases of sequencing data. Quality control and sequence assembly was performed by JGI. For further binning, only scaffolds ≥ 2000 bps were included.

405 Metagenomic binning was performed on individual assemblies using the binning tools ESOM, Anvi'o (v2.2.2)⁴⁸ and Metabat (v1)⁴⁹. For ESOM bins were extracted using getClassFasta.pl and the command -loyal 51. Anvi'o was run with default parameters and metabat was run using the following settings: --minProb 75 --minContig 2000 --minContigByCorr 2000. Results from the three different binning tools were combined using DAS Tool (version 1.0) as follows: DAS_Tool.sh -i Anvio_contig_list.tsv, Metabat_contig_list.tsv, ESOM_contig_list.tsv -l Anvio, Metabat, ESOM -c scaffolds.fasta --write_bins 1. The accuracy of the binning approach was evaluated by calculating the percentage of completeness and contamination using CheckM lineage_wf (v1.0.5).

410 Six additional MAGs (QC4_43, QC4_48, GD2_1_47_42, QZM_A2, QZM_A3, and DRTY7) were recovered from hot springs in Tibet and Yunnan, China collected in August of 2016 in several hot springs. Sequencing was done on an Illumina HiSeq4000 (Beijing Novogene Bioinformatics Technology Co., Ltd). The Chinese MAGs were assembled using metaSPADES (version 3.10.1), with a k-mer set of "21, 33, 55, 77, 99, 127". For each sample only scaffolds larger than 2500 bp were binned using MetaBAT (v.1) with default parameters, considering both tetranucleotide frequencies (TNF) and scaffold coverage information. The scaffolds from the obtained bins and the unbinned scaffolds were visualized using ESOM with a minimum length of 2500 bp and maximum length of 5000 bp as previously described⁵⁰ and the bins were modified by removing any out-of-range scaffolds (indicated by sequence points) or adding any unbinned scaffolds using ESOM related scripts⁴³. Also, scaffolds ≥ 1000 bp from each sample were uploaded to ggKbase (<http://ggkbase.berkeley.edu/>), and the bins from ESOM analyses were evaluated and modified manually at ggKbase based on GC content, coverage and taxonomic information of scaffolds.

Phylogenetic analyses. A phylogenetic tree was generated as recently described in ref⁴⁶. Briefly, 430 36 conserved marker proteins were extracted using phylosift⁵¹, in a genomic dataset containing 3,549 archaeal genomes including Brockarchaeota, and 40 bacterial genomes. An alignment of the 36 individual proteins extracted from a total of 3,599 genomes was generated using MAFFT (algorithm autoselection) with a BLOSUM62 scoring and contains 4,962 characters after masking gaps present in at least 50% of the taxa. A tree was constructed with IQtree (v1.6.11) with a best fit LG+F+R10 model selected using the Bayesian Information Criterion (BIC) and bootstraps are based on 1000 replicated trees. The bacterial genomes were used as an outgroup. The 16S rRNA sequences were extracted from Brockarchaeota genomes using Barnap (<https://github.com/Victorian-Bioinformatics-Consortium/barnap>) and used for a 16S rRNA gene phylogeny that included sequences derived from metagenomic surveys (NCBI accession 440 EU924237, KX213943, and KX213897) and the IMG database. The rRNA phylogeny was generated using RAxML within the ARB software package (v. 2.5b). using default parameters.

Metabolic predictions. Gene predictions for individual genomes were performed using Prodigal⁵² (V2.6.2, default settings). Predicted genes of individual genomes were further characterized using a combination of several databases: KofamKOALA⁵³, Interproscan v5.31.70⁵⁴, HydDB⁵⁵, dbCAN2⁵⁶, MEBS⁵⁷ and METABOLIC⁵⁸. For KofamKOALA only hits above the predefined threshold for individual KOs were selected. Hydrogenases were extracted using the reference database described in ref^{34,55} where there was conflict, the protein was manually reanalyzed using BLAST against non-redundant protein database, and genomic organization and annotation was

450 confirmed using a web-based tool Operon Mapper⁵⁹. The detected hydrogenases were used to
generate a phylogenetic tree as previously described in ref⁴⁵. Hits for key metabolic marker genes
were verified across different databases KofamKOALA, PFAMv31 and TIGRFAMs and HydDB
and were further verified using BLASTP using the NCBI web server tool. Genes encoding for
455 carbohydrate degradation enzymes described in the Carbohydrate-Active enZymes (CAZymes)
database⁶⁰ were identified by only retaining hits recovered by ≥ 2 tools. Protein localization of the
selected CAZymes was determined with the command line version of Psort (V3.0)⁶¹ using the
options -a and -terse for archaeal genomes in tabular format files. Finally, the presence of specific
protein families was obtained with MEBS. The annotation was performed in a genomic dataset of
460 250 publicly available TACK genomes (Supporting Table 2) that were also used for the CAZymes
annotation.

Methyl coenzyme M reductase screening.

The *mcrA* gene was identified using GraftM v0.10.2⁶² across metagenome assemblies where
Brockarchaeota genomes from hot springs were detected⁶³. The *mcrA*-containing scaffolds with
465 sequence length < 2.5 Kbp were discarded since scaffolds with short length were not used during the
genome binning step. The taxonomic information of the corresponding bins which contain *mcrA* genes
were determined using either GTDBtk v0.3.2⁶⁴ or phylogenetic placement (as reported in
Supplementary Table 9). The *mcrABGCD* genes were identified in metagenome assemblies from deep-
sea assemblies previously described in ref¹² (Guay17 and Guay9; IMG genome ID 3300014887 and
470 3300013103 respectively).

Data availability

The final assembled and annotated genomic sequences of Brockarchaeota from deep sea sediments
(B27_G9 and B48_G17) have been deposited in NCBI under BioProject ID PRJNA362212:
475 BioSample id SAMN09215183 and SAMN09214986 respectively. Sequence data and sample
information of Brockarchaeota from hot springs are available at NCBI under Bio Project ID
PRJNA544494.

Author contributions

480 J.F.B, W-J.L., and B.J.B conceived the study. H.C.J. and Z.S.H performed sampling at Tibet and
Yunnan hot springs. L.X.C. and N.D. reconstructed and curated the genomes. V.D.A and B.J.B.
generated the phylogenetic trees. V.D.A., N.D., and B.J.B. generated the metabolic reconstructions.
V.D.A and B.J.B performed the genomic comparisons and physiology inference. V.D.A. and B.J.B.
wrote the manuscript with input for all the other authors.

Acknowledgements

We thank Dr. Thomas Brock for a career of transformative microbiological research and graciously
allowing us to name Brockarchaeota in his honor. The work conducted by the U.S. Department of
Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported by the Office
490 of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231 provided
to ND. This work was funded by a National Science Foundation DEB: Systematics and Biodiversity
Sciences (grant number 1753661) provided to BJB. This work was supported by Science and
Technology Infrastructure work project of China Ministry of Science and Technology (No.
2015FY110100), National Natural Science Foundation of China (Nos. 91951205 and 91751206).
495 We thank Andreas for providing the sediments from Guaymas Basin. Sampling in Guaymas Basin
was supported by NSF Awards OCE-0647633 to Dr. Andreas P. Teske.

500 **References**

1. Chistoserdova, L. Modularity of methylotrophy, revisited. *Environ. Microbiol.* **13**, 2603–2622 (2011).
- 505 2. Chistoserdova, L. & Kalyuzhnaya, M. G. Current Trends in Methylotrophy. *Trends Microbiol.* **26**, 703–714 (2018).
3. Sun, J., Mausz, M. A., Chen, Y. & Giovannoni, S. J. Microbial trimethylamine metabolism in marine environments. *Environ. Microbiol.* **21**, 513–520 (2018).
4. Zhuang, G. C., Peña-Montenegro, T. D., Montgomery, A., Hunter, K. S. & Joye, S. B. Microbial metabolism of methanol and methylamine in the Gulf of Mexico: insight into
510 marine carbon and nitrogen cycling. *Environ. Microbiol.* **20**, 4543–4554 (2018).
5. Zhuang, G.-C., Montgomery, A. & Joye, S. B. Heterotrophic metabolism of C1 and C2 low molecular weight compounds in northern Gulf of Mexico sediments: controlling factors and implications for organic carbon degradation. *Geochim. Cosmochim. Acta* **247**, 243–260 (2019).
- 515 6. Vanwonterghem, I. *et al.* Methylotrophic methanogenesis discovered in the archaeal phylum Verstraetearchaeota. *Nat. Microbiol.* **1**, 1–9 (2016).
7. Lazar, C. S. *et al.* Genomic evidence for distinct carbon substrate preferences and ecological niches of Bathyarchaeota in estuarine sediments. *Environ. Microbiol.* **18**, 1200–1211 (2016).
- 520 8. Zhuang, G. Methylotrophic methanogenesis and potential methylated substrates in marine sediment. (University of Bremen, 2014).
9. Richards, M. A. *et al.* Exploring hydrogenotrophic methanogenesis: A genome scale metabolic reconstruction of *Methanococcus maripaludis*. *J. Bacteriol.* **198**, 3379–3390 (2016).
- 525 10. Sousa, D. Z. *et al.* The deep-subsurface sulfate reducer *Desulfotomaculum kuznetsovii* employs two methanol-degrading pathways. *Nat. Commun.* **9**, (2018).
11. Trembath-Reichert, E. *et al.* Methyl-compound use and slow growth characterize microbial life in 2-km-deep seafloor coal and shale beds. *Proc. Natl. Acad. Sci.* **114**, E9206–E9215 (2017).
- 530 12. Dombrowski, N., Teske, A. P. & Baker, B. J. Extensive metabolic versatility and redundancy in microbially diverse, dynamic Guaymas Basin hydrothermal sediments. *Nat. Commun.* doi:10.1038/s41467-018-07418-0.
13. Jay, Z. J. *et al.* Marsarchaeota are an aerobic archaeal lineage abundant in geothermal iron oxide microbial mats. *Nat. Microbiol.* **3**, 732–740 (2018).
- 535 14. Darling, A. E. *et al.* PhyloSift: phylogenetic analysis of genomes and metagenomes. *PeerJ* **2**, e243 (2014).
15. De Anda, V. *et al.* MEBS, a software platform to evaluate large (meta)genomic collections according to their metabolic machinery: unraveling the sulfur cycle Authors. *Gigascience* **6**, 1–17 (2017).
- 540 16. Fricke, W. F. *et al.* The Genome Sequence of *Methanosphaera stadtmanae* Reveals Why This Human Intestinal Archaeon Is Restricted to Methanol and H₂ for Methane Formation and ATP Synthesis †. **188**, 642–658 (2006).
17. McKay L, Dlakic M, Fields M, Jay Z, Eren M, Delmont T, Klingelsmith K, Rusch D, I. W. Co-occurring genomic capacity for anaerobic methane and dissimilatory sulfur metabolisms discovered in the Korarchaeota. *Nat. Microbiol.* (2019) doi:10.1038/s41564-019-0362-4.
- 545 18. Muñoz-Velasco, I. *et al.* Methanogenesis on Early Stages of Life: Ancient but Not Primordial. *Orig. Life Evol. Biosph.* 1–14 (2019) doi:10.1007/s11084-018-9570-9.
19. Orita, I. *et al.* The ribulose monophosphate pathway substitutes for the missing pentose

- 550 phosphate pathway in the archaeon *Thermococcus kodakaraensis*. *J. Bacteriol.* **188**, 4698–4704 (2006).
20. Orita, I. *et al.* The archaeon *Pyrococcus horikoshii* possesses a bifunctional enzyme for formaldehyde fixation via the ribulose monophosphate pathway. *J. Bacteriol.* **187**, 3636–3642 (2005).
- 555 21. Moran, J. J. *et al.* Formaldehyde as a carbon and electron shuttle between autotroph and heterotroph populations in acidic hydrothermal vents of Norris Geyser Basin, Yellowstone National Park. *Extremophiles* **20**, 291–299 (2016).
22. Bräsen, C., Esser, D., Rauch, B. & Siebers, B. Carbohydrate metabolism in archaea: Current insights into unusual enzymes and pathways and their regulation. *JAMA Ophthalmol.* **132**, 326–331 (2014).
- 560 23. Lokanath, N. K., Kuroishi, C., Okazaki, N. & Kunishima, N. Purification, crystallization and preliminary crystallographic analysis of the glycine-cleavage system component T-protein from *Pyrococcus horikoshii* OT3. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 1450–1452 (2004).
- 565 24. Yokohama, H., Wagner, I. D. & Wiegel, J. *Caldicoprobacter oshimai* gen. nov., sp. nov., an anaerobic, xylanolytic, extremely thermophilic bacterium isolated from sheep faeces, and proposal of *Caldicoprobacteraceae* fam. nov. *Int. J. Syst. Evol. Microbiol.* **60**, 67–71 (2010).
25. Zhang, X. *et al.* *Petroclostridium xylanilyticum* gen. Nov., sp. nov., a xylan-degrading bacterium isolated from an oilfield, and reclassification of clostridial cluster iii members into four novel genera in a new hungateiclostridiaceae fam. nov. *Int. J. Syst. Evol. Microbiol.* **68**, 3197–3211 (2018).
- 570 26. Girbal, L., Croux, C., Vasconcelos, I. & Soucaille, P. Regulation of metabolic shifts in *Clostridium acetobutylicum* ATCC 824. *FEMS Microbiol. Rev.* **17**, 287–297 (1995).
- 575 27. Qi, F. *et al.* Improvement of butanol production in *Clostridium acetobutylicum* through enhancement of NAD(P)H availability. *J. Ind. Microbiol. Biotechnol.* **45**, 993–1002 (2018).
28. Branduardi, P., Longo, V., Berterame, N. M., Rossi, G. & Porro, D. A novel pathway to produce butanol and isobutanol in *Saccharomyces cerevisiae*. *Biotechnol. Biofuels* **6**, 1 (2013).
- 580 29. Johnsen, U. & Schönheit, P. Novel xylose dehydrogenase in the halophilic archaeon *Haloarcula marismortui*. *J. Bacteriol.* **186**, 6198–6207 (2004).
30. Ravachol, J. *et al.* Mechanisms involved in xyloglucan catabolism by the cellulosome-producing bacterium *Ruminiclostridium cellulolyticum*. *Sci. Rep.* **6**, 1–17 (2016).
- 585 31. Macdonald, S. S., Blaukopf, M. & Withers, S. G. N-acetylglucosaminidases from CAZy family GH3 are really glycoside phosphorylases, thereby explaining their use of histidine as an acid/Base catalyst in place of glutamic acid. *J. Biol. Chem.* **290**, 4887–4895 (2015).
32. Jenney, F. E. & Adams, M. W. W. Hydrogenases of the model hyperthermophiles. *Ann. N. Y. Acad. Sci.* **1125**, 252–266 (2008).
- 590 33. Van Haaster, D. J., Silva, P. J., Hagedoorn, P. L., Jongejan, J. A. & Hagen, W. R. Reinvestigation of the steady-state kinetics and physiological function of the soluble NiFe-hydrogenase I of *Pyrococcus furiosus*. *J. Bacteriol.* **190**, 1584–1587 (2008).
34. Greening, C. *et al.* Genomic and metagenomic surveys of hydrogenase distribution indicate H₂ is a widely utilised energy source for microbial growth and survival. *ISME J.* **10**, 761–777 (2016).
- 595 35. Stetter, K. O. Hyperthermophiles in the history of life. *Philos. Trans. R. Soc. B Biol. Sci.* **361**, 1837–1842 (2006).
36. Adam, P. S., Borrel, G. & Gribaldo, S. An archaeal origin of the Wood–Ljungdahl H4MPT branch and the emergence of bacterial methylotrophy. *Nat. Microbiol.* **4**, (2019).

- 600 37. Collins, T., Gerday, C. & Feller, G. Xylanases, xylanase families and extremophilic xylanases. *FEMS Microbiol. Rev.* **29**, 3–23 (2005).
38. Schädel, C., Richter, A., Blöchl, A. & Hoch, G. Hemicellulose concentration and composition in plant cell walls under extreme carbon source-sink imbalances. *Physiol. Plant.* **139**, 241–255 (2010).
- 605 39. Wang, Y. *et al.* Environmental Conditions Constrain the Distribution and Diversity of Archaeal merA in Yellowstone National Park, Wyoming, U.S.A. *Microb. Ecol.* **62**, 739–752 (2011).
40. Chen, S. *et al.* The Great Oxidation Event expanded the genetic repertoire of arsenic metabolism and cycling. (2020) doi:10.1073/pnas.2001063117.
- 610 41. Rogers, K. L. & Schulte, M. D. Organic sulfur metabolisms in hydrothermal environments. *Geobiology* **10**, 320–32 (2012).
42. Rabus, R., Venceslau, S. S., Lars, W., Wall, J. D. & Pereira, I. A. C. A Post-Genomic View of the Ecophysiology, Catabolism and Biotechnological Relevance of Sulfate-Reducing Prokaryotes. in *Adv Microb Physiol* vol. 66 55–321 (2015).
- 615 43. Nunes, C. I. P. *et al.* ArsC3 from *Desulfovibrio alaskensis* G20, a cation and sulfate-independent highly efficient arsenate reductase. *J. Biol. Inorg. Chem.* **19**, 1277–1285 (2014).
44. Baker, B. J. *et al.* Genomic inference of the metabolism of cosmopolitan subsurface Archaea, Hadesarchaea. *Nat. Microbiol.* **1**, (2016).
- 620 45. Seitz, K. W. *et al.* New Asgard archaea capable of anaerobic hydrocarbon cycling. *Nat. Commun.* **10**, 1822 (2019).
46. Baker, B. J. *et al.* Diversity , ecology and evolution of Archaea. *Nat. Microbiol.* **29**, (2020).
47. Hatzenpichler, R., Krukenberg, V., Spietz, R. L. & Jay, Z. J. Next-generation physiology approaches to study microbiome function at single cell level. *Nat. Rev. Microbiol.* **18**, 241–256 (2020).
- 625 48. Eren, A. M. *et al.* Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* **3**, e1319 (2015).
49. Kang, D. D., Froula, J., Egan, R. & Wang, Z. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* **3**, e1165 (2015).
- 630 50. Dick, G. J. *et al.* Community-wide analysis of microbial genome sequence signatures. *Genome Biol.* **10**, (2009).
51. Darling, A. E. *et al.* PhyloSift: phylogenetic analysis of genomes and metagenomes. *PeerJ* **2**, e243 (2014).
- 635 52. Hyatt, D. *et al.* Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, 119 (2010).
53. Aramaki, T. *et al.* KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* **36**, 2251–2252 (2020).
- 640 54. Jones, P. *et al.* InterProScan 5: Genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240 (2014).
55. Søndergaard, D., Pedersen, C. N. S. & Greening, C. HydDB: A web tool for hydrogenase classification and analysis. *Sci. Rep.* **6**, 1–8 (2016).
56. Zhang, H. *et al.* DbCAN2: A meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* **46**, W95–W101 (2018).
- 645 57. De Anda, V. *et al.* MEBS, a software platform to evaluate large (meta)genomic collections according to their metabolic machinery: Unraveling the sulfur cycle. *Gigascience* **6**, 1–17 (2017).
58. Zhou, Z., Tran, P., Liu, Y., Kieft, K. & Anantharaman, K. METABOLIC: A scalable high-

- 650 throughput metabolic and biogeochemical functional trait profiler based on microbial
genomes. *bioRxiv* 761643 (2019) doi:10.1101/761643.
59. Taboada, B., Estrada, K., Ciria, R. & Merino, E. Operon-mapper: a web server for precise
operon identification in bacterial and archaeal genomes. *Bioinformatics* **34**, 4118–4120
(2018).
- 655 60. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. The
carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42**, 490–495
(2014).
61. Yu, N. Y. *et al.* PSORTb 3.0: Improved protein subcellular localization prediction with
refined localization subcategories and predictive capabilities for all prokaryotes.
660 *Bioinformatics* **26**, 1608–1615 (2010).
62. Boyd, J. A., Woodcroft, B. J. & Tyson, G. W. GraftM : a tool for scalable ,
phylogenetically informed classification of genes within metagenomes. **46**, (2018).
63. Hua, Z. S. *et al.* Insights into the ecological roles and evolution of methyl-coenzyme M
reductase-containing hot spring Archaea. *Nat. Commun.* **10**, 1–11 (2019).
- 665 64. Chaumeil, P., Mussig, A. J., Parks, D. H. & Hugenholtz, P. Genome analysis GTDB-Tk :
a toolkit to classify genomes with the Genome Taxonomy Database. **36**, 1925–1927
(2020).

shown). The presence methanol methyltransferase MtaB (PF12176) and trimethylamine methyltransferase MttB (PF06253) are shown in the outer circles. The annotation was conducted with MEBS15 details can found in Supporting Table 5.

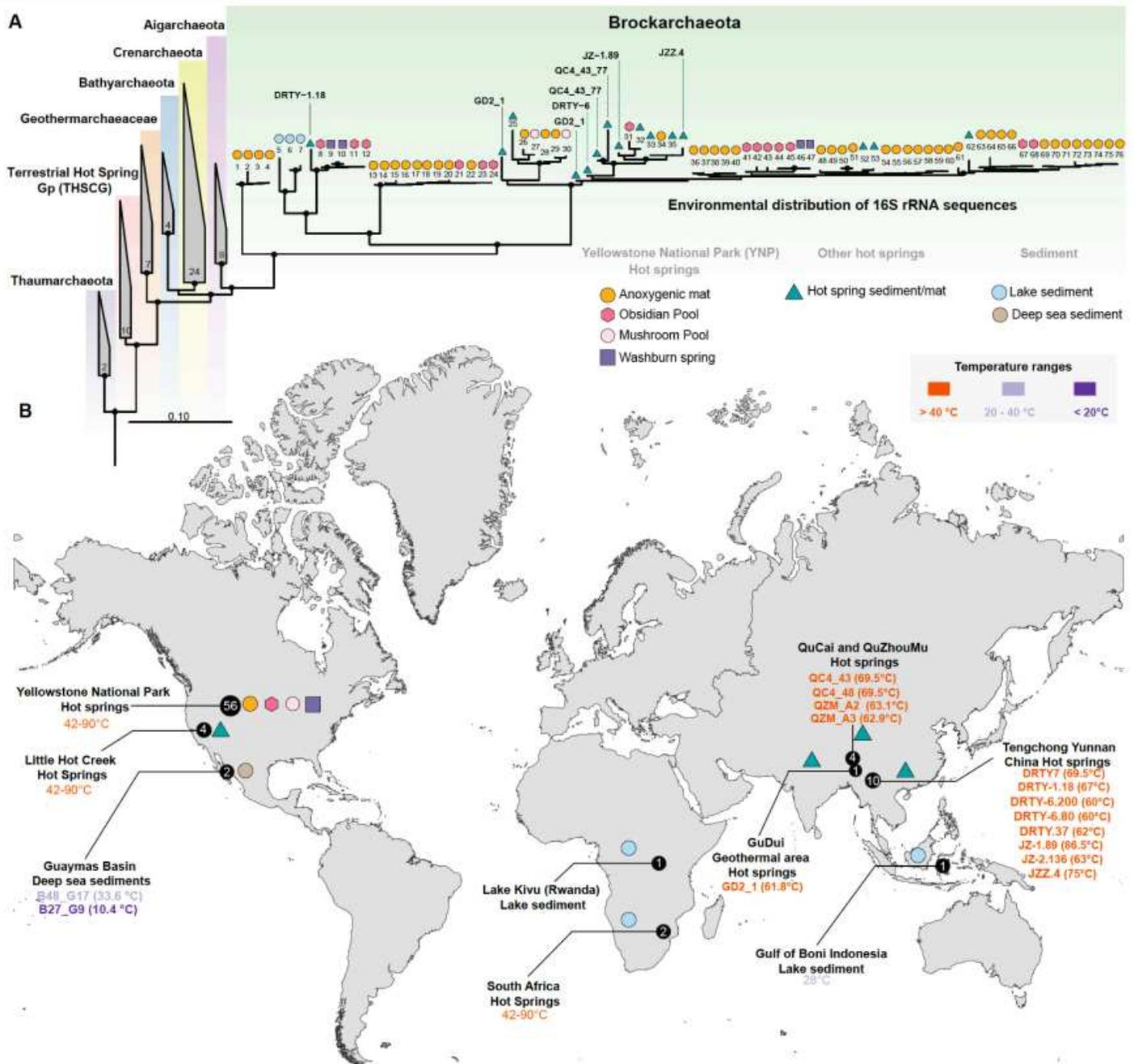


Figure 2

Location of samples from which Brockarchaeota genomes and 16S rRNA gene sequences have been recovered. (A) 16S rRNA gene tree of sequences derived from metagenomic and rRNA-based diversity surveys (NCBI accessions EU924237, KX213943, and KX213897). The eight complete 16S rRNA gene sequences of Brockarchaeota genomes described in this study are shown in their respective names. Black circles in the tree represent 100 bootstrap values using RAxML with 100 replicates. Environmental information of each sequence is shown was obtained from Integrated Microbial Genomes and

Microbiomes database. The number of the sequences and the corresponding metadata are described in Supporting Table 4. (B) Geographic localization from which Brockarchaeota sequences were obtained. The size of the circle corresponds to the total number of Brockarchaeota-related sequences in each geographic location. The specific MAGs obtained in this study are shown in the map according to their temperature range. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

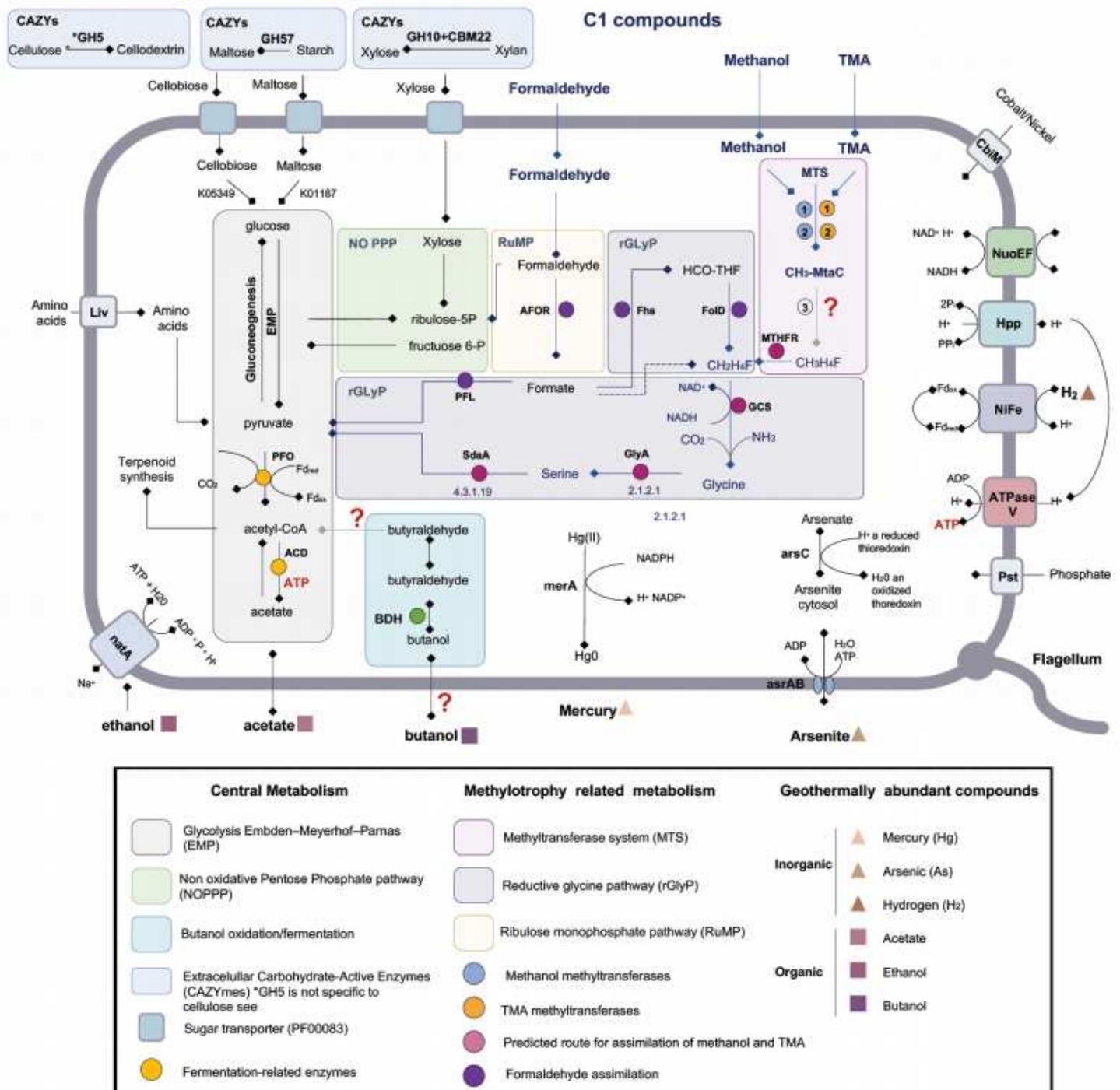


Figure 3

Overview of potential metabolic capabilities of Brockarchaeota phylum. Brockarchaeota is predicted to produce ATP by substrate-level phosphorylation by fermentation of complex organic compounds, pyruvate and acetate as carbon and energy sources via glycolysis (EMP) by the concerted action of PFO and ACD, that might represent the major energy-conserving reaction in this lineage. Formaldehyde can be fixed and detoxified via the ribulose monophosphate (RuMP) pathway and enter to central metabolism or be oxidized to formate by AFOR which can enter to the rGLyP pathway via FhS and FOLD action C1 compounds such as methanol and TMA can be assimilated via MT system and MTHFR, generating glycine, and following a linear route to central metabolism through rGLyP by subsequent conversion of serine and pyruvate. Geothermally abundant compounds are showed in colored circles Each shaded pathway can be seen in more detail in Supporting Discussion. Dashed arrow indicates that spontaneous condensation of formaldehyde with THF occur but is a very minor contribution to assimilation fluxes according to ref36. For full names and copy numbers of the genes in number see Supporting Discussion. Abbreviations. Pathways: Embden-Meyerhof-Parnas (EMP), Non-Oxidative Pentoses Phosphate Pathway (NOPPP), ribulose monophosphate (RuMP), reductive glycine pathway (rGLyP). Enzymes: pyruvate ferredoxin oxidoreductase (PFO), and acetate-CoA ligase (ADP-forming) (ACD), pyruvate formate lyase (PFL), tungsten-dependent aldehyde ferredoxin oxidoreductase (AFOR), tetrahydrofolate (THF) ligase (FhS), methenyl-THF cyclohydrolase/methylene-THF dehydrogenase (FOLD), methylene-THF reductase MTHFR (MetFmethyltransferase system (MT), glyA, serine hydroxymethyltransferase (GlyA), serine deaminase (SdaA), glycine-cleavage system (GCS). Compounds: Single carbon compounds (C1), trimethylamine (TMA). The three main components of the methyltransferase system are explained in the main text.

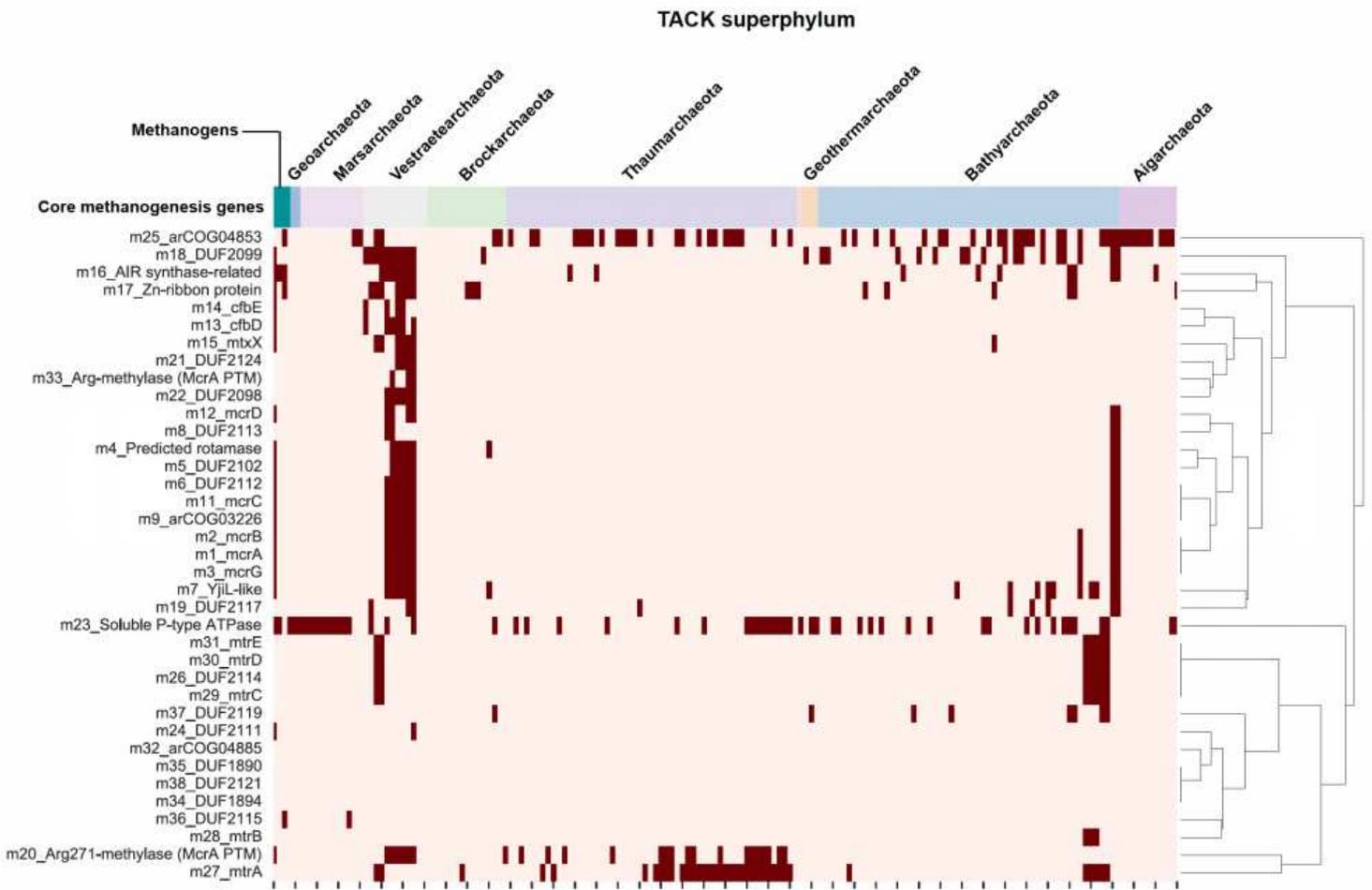


Figure 4

Presence and absence profile of common core marker genes specific to methanogenesis across the TACK superphylum. Genes include those associated to methanogenesis, anaerobic methanotrophic and short-chain alkane-oxidizing archaea described in Greening et al. 201634 (see Supporting Table 6 for specific details of the genomes used in this study and the specific marker genes named with m followed with a number in the figure). For comparison purposes the following known methanogenic archaea were included: *Candidatus Methanoplasma termitum* that lacks the entire pathway for CO₂ reduction to methyl coenzyme M and produces methane by hydrogen-dependent reduction of methanol or methylamine (Methanomassiliicoccales), *Methanosphaera stadtmanae* (Methanobacteriales) that can generate methane only by the reduction of methanol with H₂ and is dependent on acetate as a carbon source, and *Methanosarcina acetivorans* (Methanosarcinales).

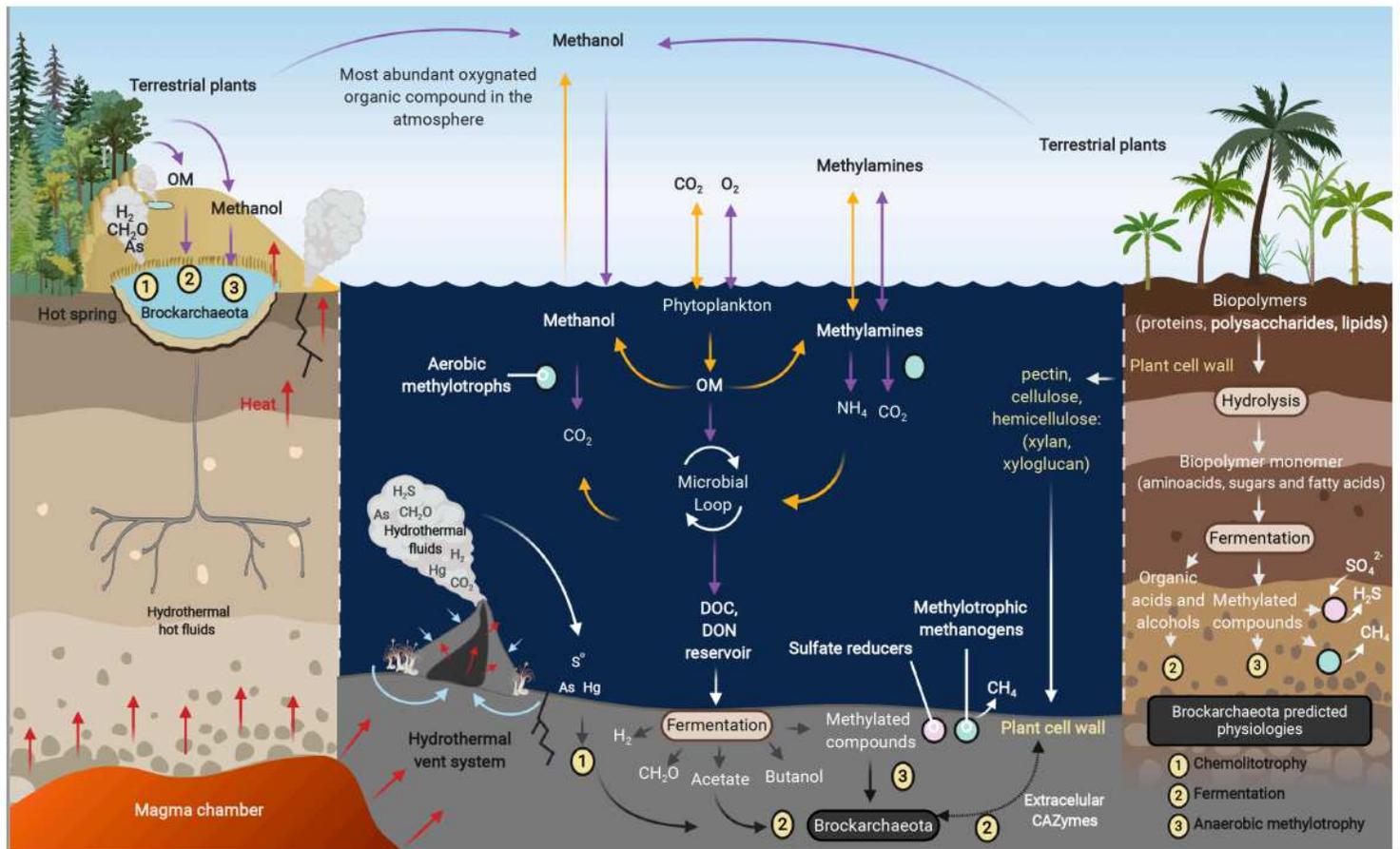


Figure 6

The role of Brockarchaeota in the anaerobic carbon cycle. Single carbon (C1)-methylated compounds, such as methanol or methylamines, are utilized biologically as carbon and energy sources in the ocean and deep-sea sediments resulting in a considerable carbon reservoir. The biodegradation of organic carbon in the water column and subsurface is a source of C1-methylated compounds. The utilization of methyl compounds as precursors in methane synthesis is confined to a small group of methylotrophic methanogens (i.e. Verstraetearchaeota). The only described anaerobic methylotrophs include members of methanogenic archaea, acetogenic bacteria, and sulfate-reducing bacteria. These organisms compete for C1 compounds geochemically produced in anoxic settings. Brockarchaeota may recycle C1 in anoxic environments without methane formation and may be sequestered in deep sea sediments and hot springs. Orange and purple arrows represent sources and sinks, respectively. Organic Matter (OM) includes dissolved and particulate organic matter feeding the microbial loop (Adapted from Evans et al., 2019 and Zhuang et al., 2018).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Description.pdf](#)