Supporting Information

Brockarchaeota, a novel archaeal lineage capable of methylotrophy

Valerie De Anda[1], Lin-Xing Chen[2], Nina Dombrowski[1,3], Zheng-Shuang Hua[4,5], Hong-Chen Jiang[6], Jillian F. Banfield[2,7], Wen-Jun Li[3,8*], and Brett J. Baker[1*]

Table of Contents

Supporting Discussion

Genomic and environmental characteristics of Brockarchaeota spp.

Brockarchaeota metagenome reconstructed metagenomes (MAGs) are estimated to be 67-92% complete. The genome sizes of these MAGs range from 0.94 to 2.90 Mbp (average 1.85 Mbp) (Table 1). Two MAGs from Guaymas Basin (GB; B48_G17 and B27_G9, temperature 33.6 and 10.4°C respectively) were originally referred to as "CP5" in ref[1]. Although the GB genomes were obtained from a lower temperature environment compared with the Chinese MAGs (obtained from hots springs up to 86.5°C) (See Supporting Table 1) metagenomes from GB, were collected in close proximity to hydrothermal circulation and experience increases in temperature due to the dynamic nature of these sediments[1]. Although these environments are primarily anoxic, the hot spring (QC4) did have measurable oxygen (DO; 1.85mg/L) in agreement with a larger redox potential (59 mV) in this site, compared with more anoxic and reducing conditions found in the GD2_1 and QZM_A2 (-191 and -160 mV respectively). Low sulfide concentrations (up to 13.16 mg/L in B48_G17 from GB), and high sulfate were characteristic of GB sediments (up to 210 mg/L G27_G9) (See Supporting Table 1). Therefore, it seems that Brockarchaeota genomes thrive in geothermal ecosystems, preferably at high temperatures at the expense of inorganic and organic compounds under anoxic conditions. Consistently, several Brockarchaeota spp. encode reverse gyrase (Supporting Table 5) essential for life under hyper thermophilic conditions. However as other hyperthermophiles, Brockarchaeota genomes were recovered from anoxic environments ranging from 10 up to 86.5 °C.

Central metabolism

Glycolysis/Gluconeogenesis: reminiscent of early life and sugar flexibility

Brockarchaeota MAGs encode most of the steps of glycolysis via Embden-Meyerhof-Parnas (EMP) pathway as well as gluconeogenesis (Supporting Figure 5). However, unlike most sugar-utilizing (hyper)thermophilic Archaea analyzed so far, where the glyceraldehyde 3-phosphate (GAP) oxidation is carried out in one step without 1-3BPG

formation and, thus, without coupling of the oxidation of GAP to 3PG with the synthesis of ATP via substrate-level phosphorylation[2], Brockarchaeota MAGs are likely to obtain energy in this step by using two different strategies.

(a) The hot spring genomes encode the glyceraldehyde-3-phosphate dehydrogenase GAPDH [EC: 1.2.1.12]) and phosphoglycerate kinase PGK [EC:2.7.2.3]. The presence of these enzymes indicate that they could couple the reversible Pi -dependent oxidation of GAP with the concomitant generation of ATP via substrate-level phosphorylation, with 1,3BPG as an intermediate[2]. While a GAPDH/PGK might be involved in GAP oxidation in glucose catabolism in hot spring genomes.

(b) The deep sea GB genome B48_G17 bin was found to encode the NAD(P)+-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.59) that might be involved in gluconeogenesis/glycolysis.

Brockarchaeota genomes encode for most glycolytic enzymes including: ATP-dependent phosphofructokinase/diphosphate-dependent phosphofructokinase (K21071), 6-phosphofructokinase (K00850; only present in JZ-2.136) pyruvate kinase (K00873) and both GAPH/PGK and GPDHII, which might reflect their ability to degrade sugars, consistent with their broad spectrum of Carbohydrate-Active Enzymes (CAZymes) (see discussion below).

The phosphomannomutase/phosphoglucomutase (pgm) a key enzyme that serves as branch point in glucose metabolism for energy synthesis and cell surface construction providing substrates for the PPP pathway or the synthesis of polysaccharides from glucose-6-phosphate or mannose-6-phosphate[3] was present in most of the Brockarchaeota genomes (Supporting Figure 5)

Another difference between the hot spring and GB lineage is the presence of the heat-stabile fructose 1,6-bisphosphate aldolase/phosphatase [EC:4.1.2.13/3.1.3.11] (Fba, (K01622), only in the GB genome B48_G17 a key ancestral gluconeogenic enzyme that guarantees a unidirectional gluconeogenic pathway under conditions where the carbon flux does not need to be turned to sugar degradation[3]. The same reaction can be performed by fructose-bisphosphate aldolase/ 2-amino-3,7-dideoxy-D-threo-hept-6-

ulosonate synthase [EC:4.1.2.13/2.2.1.10] (K16306) detected in the hot spring genomes JZ-2.136 and JZZ-4.
Brockarchaeota from hot springs encode Fructose-1,6-bisphosphate aldolase (FBPA), class II (EC 4.1.2.13, K01624). Since this key enzyme for gluconeogenesis and glycolysis is widely distributed in bacteria and fungi, and rarely found in archaea, it has been suggested that it might be a HGT event from bacteria to haloarchaea[2] where they have been found.

Brockarchaeota MAGs display an incomplete tricarboxylic acid cycle (TCA) (Supporting Table 5). Only one bin encodes all the subunits for the key enzyme 2-oxoglutarate synthase (K00174, K00175, K00177), compared to the rest of the Brockarchaeota MAGs that encode one or two subunits. Furthermore, key enzymes β-oxidation pathway are absent in the Brockarchaeota genomes.

Pentose phosphate pathway (PPP) and ribulose monophosphate (RuMP) pathway

Although the complete Non-Oxidative Pentoses Phosphate Pathway (NOPPP) appears to be rare in Archaea, homologous of all four key enzymes; ribose-5-phosphate isomerase (RPI) ribose-5-phosphate-3-epimerase (RPE), transketolase (TK); transaldolase (TA) are present in Brockarchaeota (Supporting Figure 2). This suggests that these genomes are able to synthesize R5P and E4P via the reverse reactions of the NOPPP from F6P and glyceraldehyde 3-phosphate. Furthermore, Brockarchaeota MAGs encode both key enzymes of the ribulose monophosphate (RuMP) pathway: 3-hexulose-6-phosphate synthase (HPS) and 6-phospho-3-hexuloisomerase (PHI). The RuMP pathway was originally found in methylotrophic bacteria that are able to use C1 compounds as a sole source of carbon and energy, however, it is currently recognized as a widespread prokaryotic pathway for formaldehyde fixation and detoxification[4]. Formaldehyde was an essential building blocks 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 pectin[5]. In geothermally active systems such as in hydrothermal vents and hots springs, formaldehyde is expected to be present since C1 compounds are found in a variety of redox levels (i.e CH_4, CO_2, CO), and could be produced by autotrophic microorganisms such as *Metallosphaera yellowstonensis MK1*[6]. The RuMP

pathway functions as a highly efficient system for trapping free formaldehyde at relatively low concentrations. The presence of HPS and PHI in Brockarchaeota genomes suggest that formaldehyde can be fixed and detoxified via the RuMP pathway.

The presence Tungsten-dependent aldehyde ferredoxin oxidoreductase (AFO) (K03738) and RuMP indicates that Brockarchaeota can play a key role in controlling formaldehyde consumption and therefore maintaining viability of the microbial community in hot springs and deep-sea environments.

Catabolism of pentoses: xylose degradation

Interestingly, Brockarchaeota not only possesses the metabolic capability to degrade hexoses via EMP pathway, however, they encode key enzymes involved in the degradation of pentoses specifically xylose. Degradation of xylose in Brockarchaota can be possible by the presence of xylose isomerase *xylA* [EC:5.3.1.5] and xylulose kinase *xylB* [EC 2.7.1.17] (Supporting Figure 2). These enzymes are only being found in bacterial thermophiles and halophilic archaea that can grow by fermentation of complex compounds and can degrade xylose suggesting a similar behavior in Brockarchaeota genomes[7]

Fermentation

Reconstruction of Brockarchaeota metabolisms strongly support that this group is composed of hyperthermophile facultative and obligate anaerobic fermentative organism predicted to produce ATP by substrate level phosphorylation, producing acetate carbon dioxide and hydrogen as byproducts of their metabolism. Brockarchaeota are likely to produce H_2 during acetogenic fermentation. Brockarchaeota can potentially have the ability of fermentative H_2 production given the presence of the bidirectional group 3b and 3c [NiFe]-hydrogenases (Supporting Figure 6).

Brockarchaeota encodes enzymes that indicate the potential to grow on organic substrates, which could be fermented to acetate (Supporting Figure 5). The ATP conserving step of sugar or pyruvate fermentation to acetate is catalyzed by the enzyme acetate-CoA ligase (ADP-forming) present in the hot spring genomes. Acetate can also be assimilated back to acetyl-CoA by means acetyl-CoA synthetase (ACS) [EC:62.1.1], suggesting that in the absence of other substrates, acetate can serve as a total source of

carbon and energy in hot springs Brockarchaeota genomes. Due to the acetogenic capability, we explore whether Brockarchaeota was similar to other homoacetogenic microorganisms, that can couple H_2-dependent CO_2 reduction with H2-producing fermentation. Homoacetogenic archaea such as Lokiarchaeota couple Wood–Ljungdahl pathway (WLP) with a variety of organic electron donors (for example, sugars, ethanol and lactate)[8], allowing them ecological advantages. However, Brockarchaeota lacks the enzymes of methanogenesis (Figure 3) including the methyl and carbonyl branches (Supporting Figure 7), discarding any couple mechanism for H2 producing fermentation with WLP.

Brockarchaeota encode the four subunits of the pyruvate ferredoxin oxidoreductase (PFO).Although the presence of this enzyme allows to couple the (WL) pathway with the reductive tricarboxylic acid (rTCA) in autotrophic anaerobes, the lack of key marker genes including ATP-citrate-lyase (aclAB), and carbon monoxide dehydrogenase/acetyl-CoA synthase (CODH/ACS) indicates that Brockarchaeota cannot support autotrophic metabolism.

PFO is found in strict anaerobes and organisms that experience anoxia (including algae). PFO acts by coupling pyruvate oxidation to H_2 production ultimately generating acetyl-CoA. PFO are classified into four groups based on their subunit compositions. Consistent with the metabolic repertoire, Brockarchaeota encodes the third type of PFO, that encodes four subunits. This type of PFOs has been found in hyperthermophile anaerobes, including the fermentative archaea *Pyrococcus furiosus* and *Thermococcus litoralis*, the sulfate-reducing archaeon *Archaeoglobus fulgidus*, and the fermentative bacterium *Thermotoga maritima*[9].

Brockarchaeota genomes encode several proteins that possess the nitronate monooxygenase domain (PF03060). As it was suggested recently in ref[10] this family may represent a novel NADH: quinone reductases, that could function in the re-oxidation of NAD(P)H generated during growth of Brockarchaeota on organic substrates as has been suggested for some fermentative members of the Asgard phylum.

Brockarchaeota from hot springs can generate Na^+ gradient using different proton pumps allowing for Na^+-driven synthesis of ATP via a V-type ATPase. For example, they encode the natA sodium transport system ATP-binding protein [EC:7.2.2.4] (Figure 3), that is active in the presence of ethanol[11] and the K(+)-stimulated pyrophosphate-energized sodium pump encode membrane-bound proton-translocating inorganic

pyrophosphatases (HppA), which might have a crucial role in anaerobic metabolism[12] and may allow energy conservation from hydrolysis of pyrophosphate in Brockarchaeota couple with the electron bifurcation complex. The extent to which this mechanism is associated with energy conservation mechanism is still unclear. Brockarchaeota genomes appear to encode a wide repertoire of ATPases: i) the plasma-membrane proton-efflux P-type ATPase [EC:7.1.2.1] (K01535) (only present in the hot spring genomes), ii) $Zn2+/Cd^{2+}$-exporting ATPase [EC:7.2.2.12 7.2.2.21] (K01534) only present in DRTY7.37 and finally the V/A-type H+/Na+-transporting ATPase (present in GB and hot spring genomes). The existence of ATPase in Brockarchaeota suggests that members of this latter group have the additional ability to couple acetogenic fermentation to membrane potential generation of a transmembrane ion gradient across the membrane.

Although the enzymatic repertoire differs between MAGs from GB and hot springs suggesting that they are characterized by different anaerobic physiological lifestyles, this could also be the reflection of low number of genomes from deep-sea sediments. 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.

Detoxification or energy conservation mechanisms

Arsenate

The genetic system for arsenic resistance is present in Brockarchaeota genomes. This detoxification enzymatic system acts by decreasing the intracellular arsenic concentration that pumps out the arsenate that enters the cell, thus preventing the metals from accumulating and denaturing proteins[13]. The intracellular thioredoxin dependent arsenate reductase (ArsC, K03741) that catalyzes the reduction of arsenate to arsenite, is present in most hot spring genomes (Figure 3). The non-toxic arsenite is then exported from the cell by the arsenite efflux transporter (asrA, K01551 and asrB K03893). The presence of this energy-dependent efflux process related detoxification enzymes, could indicate that hot springs Brockarchaeota genomes could use arsenate as terminal electron

acceptor, since some anaerobic bacteria like *Desulfovibrio strain Ben-RA* can reduce arsenate in the presence of asrC-like genes[14] Subsequently, molecular genetics and biochemical studies with *D. alaskensis* G20 demonstrated that ArsC functions as an efficient arsenate reductase with electrons delivered from thioredoxin[13]

Mercury

Geothermal ecosystems such as shallow, and deep-sea vents, volcanoes, geysers, hot springs, and fumaroles are known natural sources of mercury (Hg)[15]. Hg resisting microorganisms are known to be enriched in deep-sea hydrothermal vents and in terrestrial geothermal springs. Three hot springs Brockarchaeota genomes (DRTY735_44, DRTY-1.18 and DRTY.37) encode mercuric reductase (MerA) [EC:1.16.1.1] which plays transforms Hg (II) to Hg(0)[15]. The presence of this enzyme has been suggested to be a reflection detoxifying Hg under the different environmental naturally occurring mercury-rich geothermal environments.

Sulfur

Thiosulfate is a significant intermediate in the sulfur cycle of anoxic marine and freshwater sediments, where it is involved in reduction, oxidation, and disproportionation pathways[16] JZ-2.136 is the only bin that encodes Rhodanese (PF00581) implicated in the disproportionation of thiosulfate for energy conservation purposes[17]. Biochemical characterization of *Aquifex aeolicus*, isolated from an hydrothermal system indicate that this bacterium can use either tetrathionate ($S_4O_6^{2-}$), polysulfide or thiosulfate ($S_2O_3^{2-}$) as electron donors using the rhodanese pathway, however despite numerous studies, the physiological role of this pathway remain unclear and are still widely debated[17], and therefore the role of Rhodanase in Brockarchaeota remains unclear. Similar to other heterotrophic fermentative hyperthermophilic archaea[18,19] Brockarchaeota might be able to reduce elemental sulfur during fermentative growth and produce H_2S due to the presence of [NiFe] Group 3b hydrogenases (Supporting Figure 6). During carbohydrate fermentation in the absence of sulfur, [NiFe] Group 3b hydrogenase can catalyze the production of H_2 with NADPH or NAD(P)H as the electron donor. However, in the presence of sulfur, Brockarchaeota from might have the ability to reduce sulfur using H_2 or organic substrates as electron donors, a widespread physiology

in hyperthermophilic archaea living in geothermally active environments (volcanic habitats, hots springs or marine sediments).

Hydrogen

Hydrogen is found at significant levels in most hydrothermal systems due to volcanic outgassing or abiotic production[20]. Since [NiFe] Group 3b hydrogenase can also catalyze the reverse reaction of H_2 oxidation using NADP+ or NAD(P)+ as an electron acceptor[21], it is possible that Brockarchaeota might be the use of hydrogen gas as electron donor.

Brockarchaeota from hot springs might switch metabolism depending on the availability of external electron acceptors, is the presence of the oxygen-tolerant group 3d [NiFe]-hydrogenases (Supporting Figure 6). This group is suggested to interconvert electrons between NAD(P)H and H_2 depending on the availability of electron acceptors. Consistently, 3d [NiFe]-hydrogenases are found to be abundant at metagenomic level in hot springs where it suggested that this could be the reason why microbial communities in these geothermally active systems are relatively stable despite pO_2 fluctuations[22].

Butanol degradation

Brockarchaeota genomes from hot springs encode a butanol dehydrogenase BDH (Supporting Figure 3) that catalyzes the reversible conversion of butyraldehyde to butanol, oxidizing NAD(P)H. Brockarchaeota BDHs are homologues to obligately anaerobic, thermophilic bacteria that can degrade complex of complex plant saccharides such as xylan (i.e *Caldicoprobacter oshimai*[23] *Hungateiclostridium thermocellum*[24]) or cellulose (*Hungateiclostridium alkalicellulosi).* To investigate whether butanol was being oxidized or produced by Brockarchaeota, we look for the genes involved in production of butanol in two model organisms. First we look for the genes described in the obligately anaerobic bacterium *Clostridium acetobutylicum*[25,26] since it is one of the few organisms known to produce butanol as a major fermentation product. We also analyze the pathway involved in butanol and isopropanol production described in *Saccharomyces cerevisiae*[27]. We found that Brockarchaeota genomes lacks the key enzymes involved in the fermentation of pyruvate to butanol (butanal dehydrogenase, butyryl-coA dehydrogenase, enoyl-CoA

dehydratase, 3-hydroxyacyl-CoA dehydrogenase). However, most of the genomes encode a putative aldehyde dehydrogenase (PF00171) that could convert butyraldehyde to butyric acid. Also, we found a putative enoyl-CoA hydratase/isomerase protein that is only encoded in bin JZ-1.89, that could be involved in further processing of the butyric acid molecule. Our results suggest an alternative pathway for butanol oxidation that still remains unresolved (Supporting Figure 4).

Methylotrophy

Methylotrophs are a phylogenetically can utilize C1 compounds of a more reduced form than that of CO_2 as a sole source of energy and carbon. Among the reduced C1 compounds, methane, methanol, methylamines and formaldehyde are ubiquitous in nature in the form of methyl esters of pectin and of the methoxyl group of lignin[28]. C1 compounds are found at a variety of redox levels in marine hydrothermal vents (e.g., methane, CO, and CO_2).

The reduced C1 compounds are first oxidized to formaldehyde, which the branch point between further oxidation to CO_2 for energy conservation purposes and assimilation. Brockarchaeota exhibit one of the three assimilation pathways described for methylotrophic organisms (the ribulose bisphosphate (RuBP) pathway (Supporting Figure 2), that also functions as a highly efficient system for trapping free formaldehyde at relatively low concentrations. Due to the presence of RuBP pathway we search for the methylotrophic metabolism in Brockarchaeota genomes.

Aerobic methylotrophy

Brockarchaeota lack the main the main enzymatic steps involved in aerobic methylotrophy including the PQQ-linked methanol dehydrogenases (MDH) and NAD-linked MDH found in aerobic methylotrophic bacteria[27] also, they lack methane monooxygenases (MMO) found in methanotrophic organisms[28]. Brockarchaeota lacks the methylamine dehydrogenase pathway that oxidizes C1 compounds to formate (Figure 10 left side panel). Also, they lack the essential genes for the alternative pathway that is found in organisms that can grow on C1 compounds and lack methylamine dehydrogenase. This alternative pathway proceeds via N-methyl-L-glutamate synthase complex (mgsABC)). Only one Brockarchaeota genome encodes mgsC and but lacks the rest of the subunits for the complex as well as the enzymatic steps for the pathway. Due

to the presence of glycine hydroxymethyltransferase (SHMT), is likely that formaldehyde can react with glycine to form L-serine which can be converted to pyruvate via threonine dehydratase [EC 4.3.1.19] and serve as be an entry point to the central metabolism in the form of pyruvate.

Anaerobic methylotrophy

So far, 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[29–31]. These organisms use a methyltransferase MT system that involves breaking and transferring the methyl residue to a C1 carrier, which can be coenzyme M in the case of methanogens or tetrahydrofolate in acetogens. Brockarchaeota from hot springs encode MT system for the utilization of methanol and methylamine, methanol-CoM methyltransferases (MtaAB) and trimethylamine-corrinoid protein transferases (MttB) respectively. So far, no methylotrophic members of the archaea domain have been described outside methanogenic groups. Despite the presence of MT system, unlike phylogenetically related archaea such as Verstraetearchaeota[32]; Brockarchaeota genomes do not possess the common core marker genes specific to methanogenesis including Methyl-coenzyme M reductase (MCR) (Figure 3).

JZ-2_136 is the only genome encoding a maximum of three methanogenesis marker genes described in ref[33] and listed in Supporting Table 6. These marker genes are also present in some other non-methanogenic TACK genomes such as m23_Soluble P-type ATPase (arCOG01579), m25_arCOG04853 and m37 m37_DUF2119. GD2_1 _47_42 is the only genome that encodes trahydromethanopterin S-methyltransferase subunit A mtrA (m27). Genomes from Yunnan area DRTY-1_18, DRTY7 _35_44 and DRTY7_37 encode m17_Zn-ribbon protein present in Verstratearchaeota, Bathyarchaeota and known methanogenic genomes.

Another indication that methylotrophy in Brockarchaeota does not proceed via methanogenesis is the complete lack of the carbonyl branch and incomplete methyl branch of the Wood–Ljungdahl (WL) pathway (Supporting Figure 7).

JZ2-136 is the only bin encoding two key enzymes present in the tetrahydromethanopterin (H4 MPT) pathway: the formylmethanofuran dehydrogenase (MRF-dehydrogenase) and the F_{420}-dependent methylene-H_4MPT dehydrogenase (MTD)

[EC:1.5.98.1] (PF01993). Some genomes from hot springs encode the WL-methyl tetrahydrofolate (H4F) pathway.

In the case of the H_4MPT pathway, is unclear the role of MRF-dehydrogenase in JZ2-136. This complex performs the first step in the Wolf cycle to reduce CO_2 to formylmethano-furan (MFR-CHO)[34]. This reaction is endergonic and dependent on reduced ferredoxin (Fd_{red}). In obligate CO_2 reducing methanogens this step is coupled to a flavin-based electron bifurcation complex FEBC involving a soluble MvhAGD:HdrABC[35]. However, JZ2-136 only encode one heterodisulfide reductase subunit A2 [K03388] (Supporting Table 5), that belongs to [NiFe] Group 3: Cofactor-coupled bidirectional [NiFe] hydrogenases, not likely to be involved in FEBC in this Brockarchaeota genome. In methanogenic archaea, the formyl group of the MFR-CHO is then transferred to H_4MPT to yield CH_3-H_4MPT which ultimately is transferred to coenzyme M (HS-CoM) by the methyl-H_4MPT: coenzyme M methyltransferase (MtrA-H) complex[29] to generate a proton translocation gradient that drives ATP synthesis. Only MAGs from hot springs encode the A and H subunits of Mtr complex which are associated with the methyl transfer from methyltetrahydromethanopterin to coenzyme M[36] . However, all Brockarchaeota MAGs lack the MtrBCDEFG subunits, indicating that there is no coupling site for methyl-coenzyme M formation associated to an electrochemical gradient concomitant to ATP formation.

In the case of Brockarchaeota MAGs that display incomplete H4F methyl WL pathway, they encode methylene-H4F reductase (NADPH) [EC:1.5.1.20] and Formate--tetrahydrofolate ligase [EC:6.3.4.3]. In other archaea methylene-H4F reductase are suggested to form a complex with HdrCBA and MvhD proteins, indicating some sort of electron bifurcation. In the case of Brockarchaeota, methylene-H4F Reductase do not form a transcription unit with hdrCBA, mvhD and metF (Supporting Table 9) however, they are associated with an uncharacterized membrane-associated protein/domain (COG2512), ferredoxin (COG1145) and thioredoxin (COG0492). Yet, we can suggest that the presence of 3C[NiFe]-hydrogenase in hot spring genomes including JZ-236 () can indicate the presence of a functional complex with heterodisulfide reductase by reducing ferredoxin and heterodisulfide during H_2 oxidation. The presence of 4G [NiFe] in JZ-136 and B48_G17 suggest that these organisms can couple $ferredoxin_{red}$ oxidation to proton reduction leading to a proton gradient driving ATP synthesis and regenerating H_2

recycled for a probable bifurcating complex[22,37]. ATP synthesis can be catalyzed by the ATP synthase dependent on the Na^+ and H^+ gradients. Yet, this is all speculative and no clear evidence supports this mechanism in Brockarchaeota yet.

Overall, Brockarchaeota genomes do not present the canonical methylotrophic mechanism described for phylogenetically related genomes, and incomplete WL pathway.

Degradation of complex carbohydrates

Brockarchaeota genomes display a wide variety of genes that encode for enzymes that break down glycans (CAZymes; Carbohydrate-Active Enzymes) divided in the following classes: Carbohydrate-binding modules (CBM), Glycoside Hydrolases (GH) Glycoside Transferases (GT) and Polysaccharide Lyases (PL). A total of 301 CAZYmes-related genes were identified among the 15 Brockarchaeota genomes (Supporting Table 8). The hot spring ones are the most metabolically versatile, encoding from 8-58 CAZYmes per genome, while GB genomes encode 4-10 (B48_G17, B27_G9 respectively). The genome DRTY.6.80 exhibits 58 CAZymes-related genes, that is more than the average of Brockarchaeota CAZymes-related genes (around 27 genes per genome).

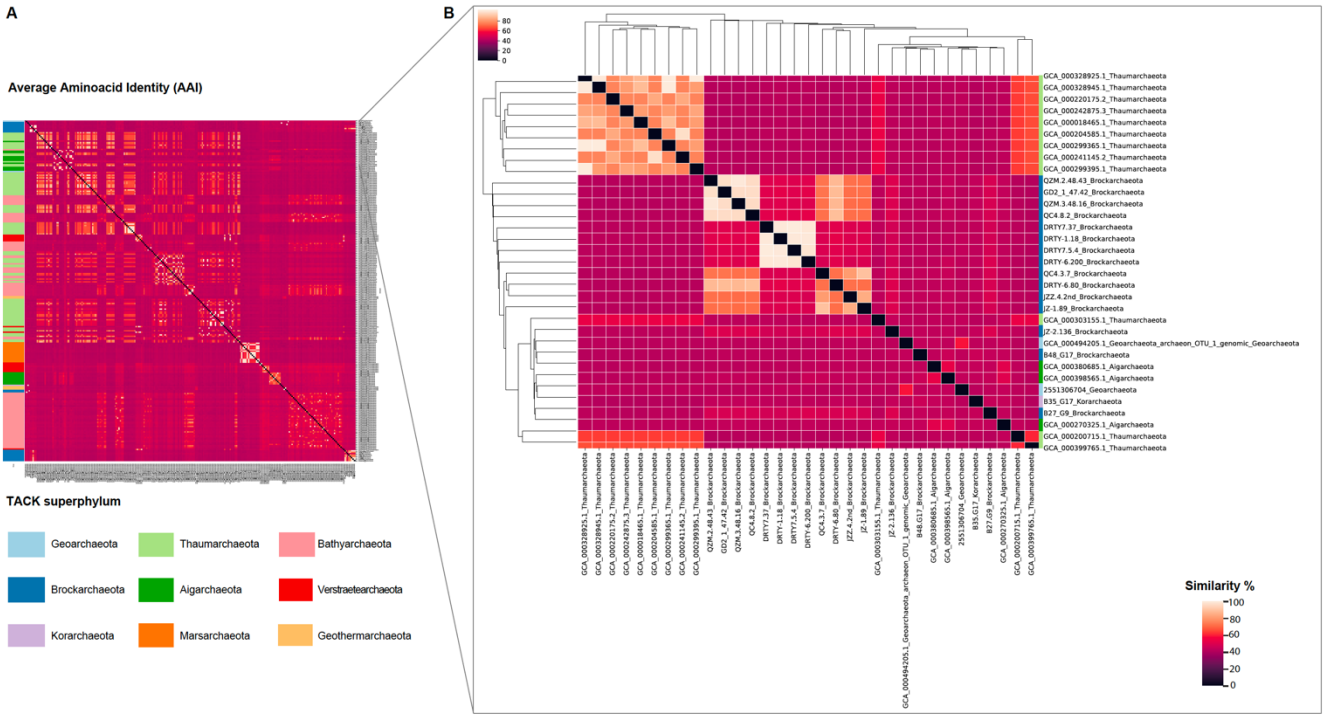
Consistent with the analysis of central metabolism, the most diverse enzymatic class among present in Brockarchaeota genomes is GH. This group of CAZYmes is common in heterotrophs and they catalyze the hydrolysis of glycosidic bonds in primary metabolism, allowing the use of complex carbohydrates as energy sources[38] The predicted extracellular CAZYme belong to the GH class, and they indicate that Brockarchaeota could have a key role in plant-that indicated a key role in biomass degradation .

Five different types of Glycosyl transferases were detected among Brockarchaeota genomes. Interestingly, type GT4 and GT2 which are the families with the highest number of enzymes in hyperthermophilic Archaea were also present in high numbers in Brockarchaeota. GT4 accounts for a total of 29 genes in the 15 MAGs, some like DRTY.6.80 encode up to 5 copies, while GT2 was present in both GB genomes and 4 hot spring MAGs. These families include activities involved in cellulose, chitin, and sucrose biosynthesis, but also in N-glycosylation Supporting Table 8.

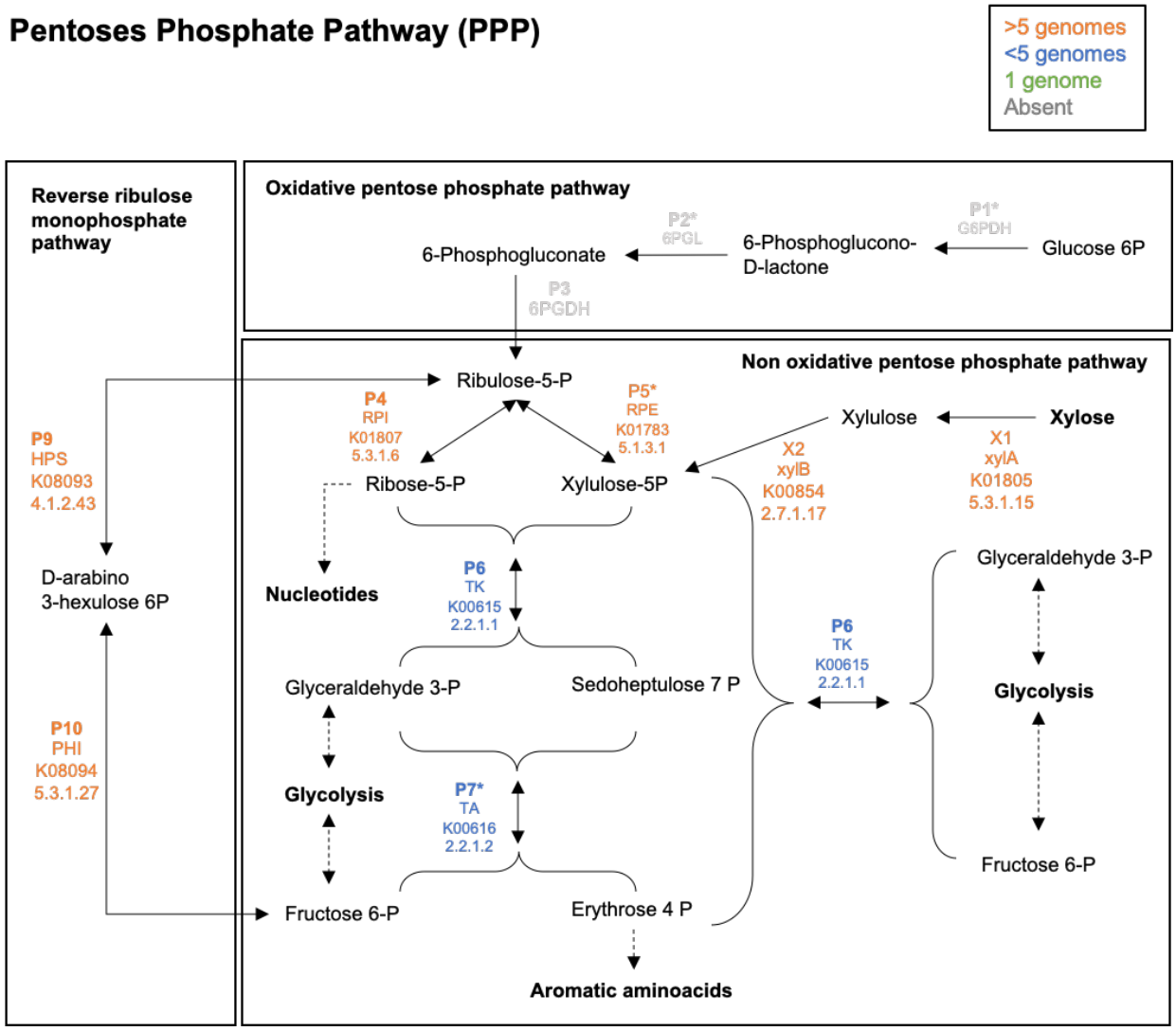
The genome size of known complex-carbon degrading microorganisms isolated from similar environments ranges from 2-4Mb. For example *Caldicellulosiruptor kronotskyensis*[39] isolated from hot springs (2.8M), the anaerobic, xylanolytic, extremely thermophilic bacterium, *Caldicoprobacter oshimai*[23] (2.7 Mb) and the xylan-degrading bacterium *Petroclostridium xylanilyticum*[24] (3.8Mb). The genus Caldicellulosiruptor is globally distributed, being isolated from terrestrial geothermal hot springs in Russia, Iceland, Yellowstone National Park and New Zealand and from solar-heated mud flats in Owens Lake, CA. With optimal growth temperatures ranging from 70 to 78°C[40,41]. All members of the Caldicellulosiruptor genus have similar genome size ranging from 2.4 to 2.97 Mb. Similarly, Brockarchaeota genome size ranges from 0.94-2.9 Mb.

A genomic analysis of the potential CAZymes of members of the TACK superphylum, including Brockarchaeota (Supporting Table 8) revealed the unique potential of Brockarchaeota for xylan degradation among TACK superphylum. The 3 unique CBM: (CBM67, CBM4 and CBM9), are associated with xylanases. Among the 13 unique GH (GH42, GH163, GH33, GH122, GH148, GH39, GH94, GH141, GH97, GH10+CBM22 GH149), two of them are predicted to be extracellular (GH10+CBM22, GH5+GH12) and with xylanase activity in Brockarchaeota. The carbohydrate esterase CE7 unique in Brockarchaeota catalyzes the hydrolysis of acetyl groups from polymeric xylan[42].

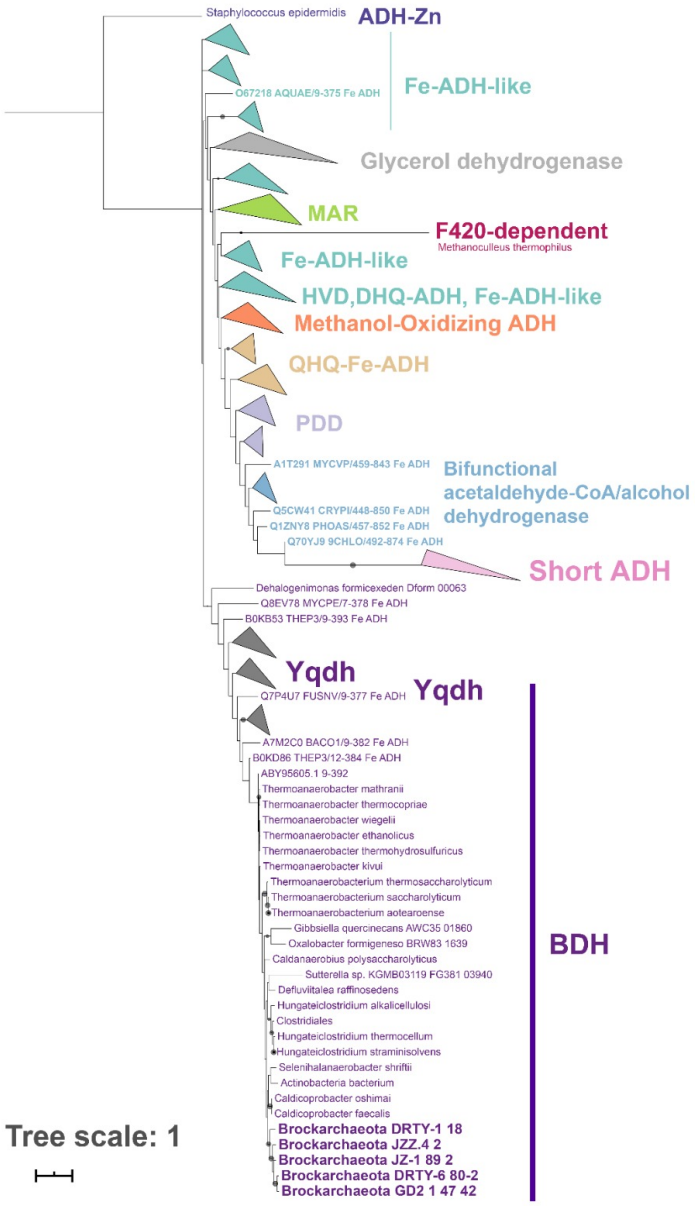
Supporting figures



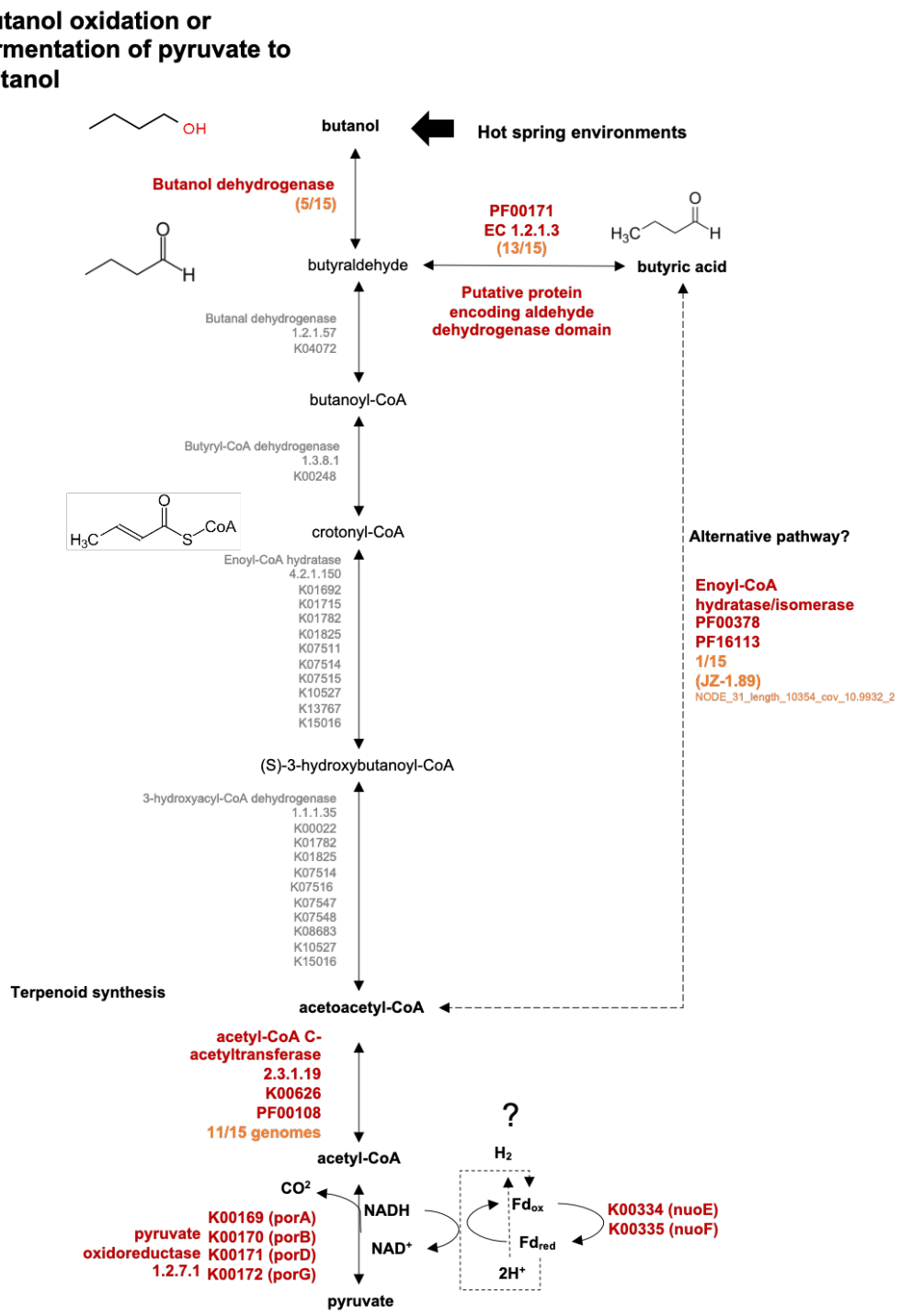
Supporting Figure 1. Average amino acid identity (AAI) comparison of Brockarchaeota genomes and phylogenetically related TACK phyla A) Hierarchical clustering heatmap based on average amino acids identity for each genome pair of genomes described in Supporting Table 2. Genome self-comparisons are shown in the black square diagonal. Analysis was done using compareM. B) Hierarchical clustering heatmap based on average amino acids identity for each genome pair of closely related genomes o Brockarchaeota based on panel A. Selected genomes include Geoarchaeota, Aigarchaeota and Thaumarchaeota phyla. The color code indicates the AAI (%) from 0-100 (red to white respectively).



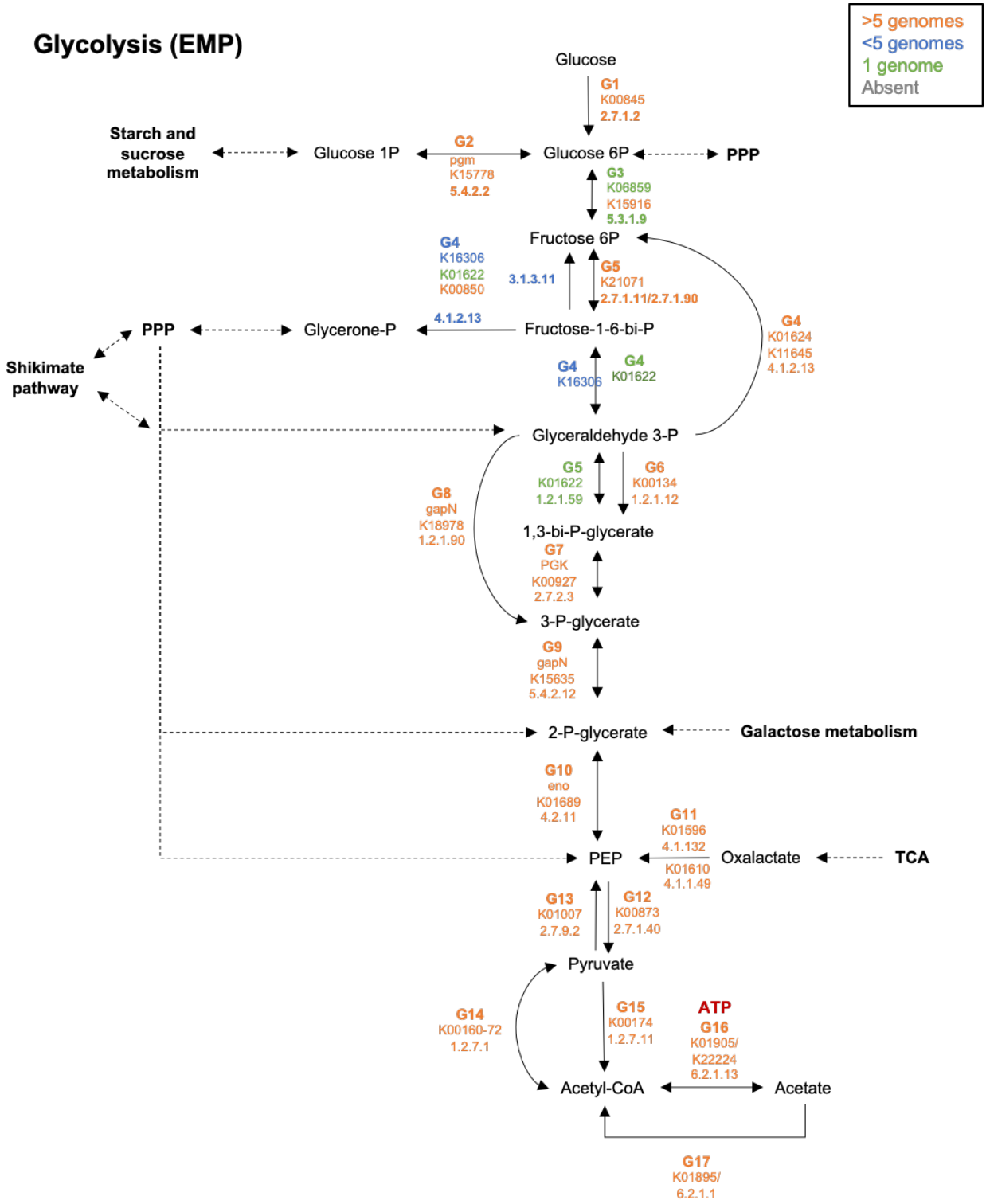
Supporting Figure 2. Pathways for pentose formation in Brockarchaeota. Asterisks indicate enzymes not identified or common in archaea according to ref[2] . Enzymatic steps are found in Supplementary Table 7. Abbreviations: G6DPH, glucose-6-phosphate dehydrogenase; 6PGL, 6-hosphogluconate-D-lactone; 6PGDH 6-phosphogluconate dehydrogenase; RPI, ribose-5-phosphate isomerase; RPE ribose-5-phosphate-3-epimerae; TK, transketolase; TA, transaldolase; HPS,3-hexulose-6-phosphate synthase, PHI, 5-phospho-3-3hexuloisomerase



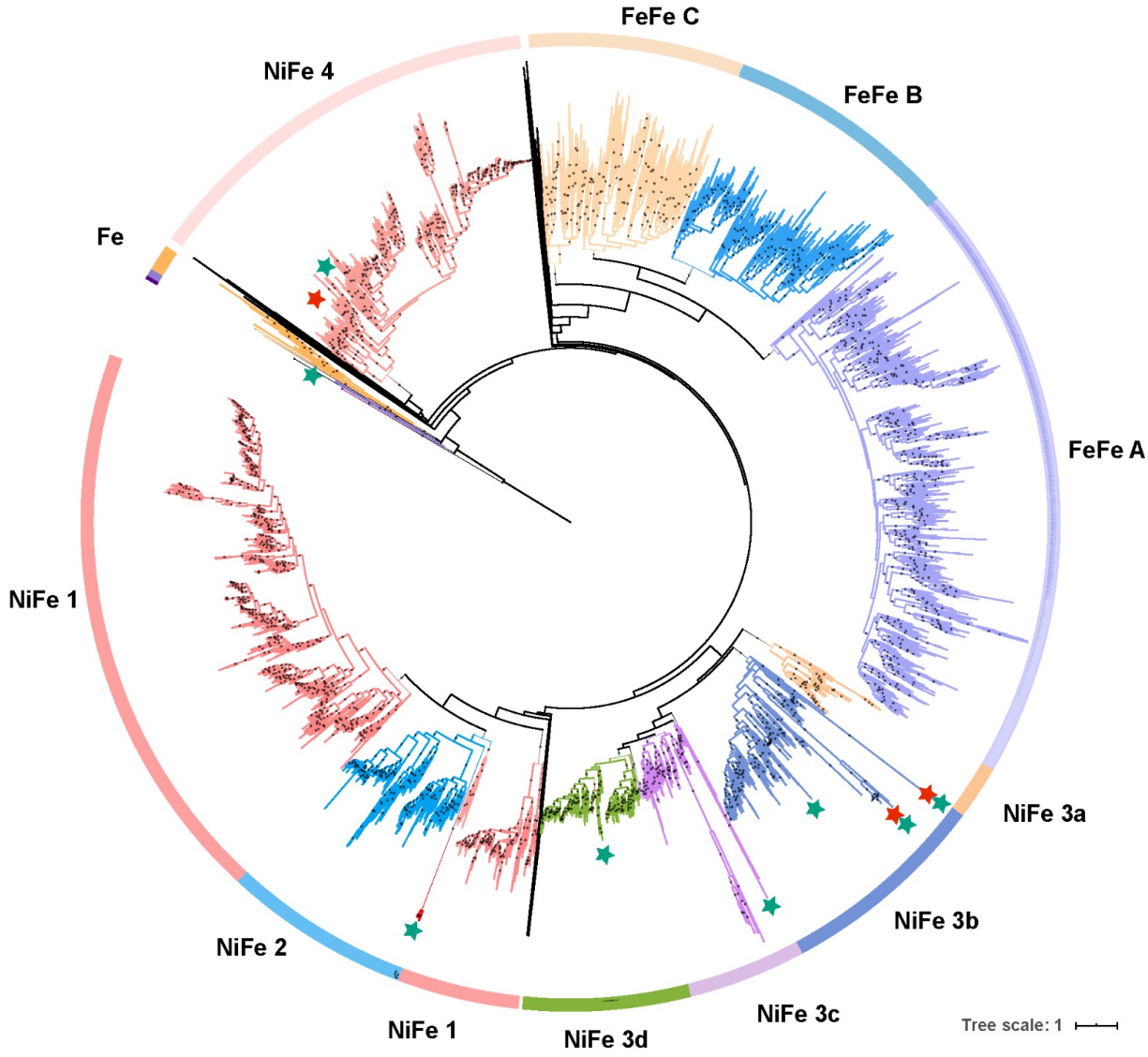
S**upporting Figure 3.** Phylogeny and classification of alcohol dehydrogenases (ADH) amino acid sequences found in Brockarchaeota genomes. The sequences were obtained from a BLASTp analysis using ADHs from Brockarchaeota as query sequences. Reference ADHs were obtained as follows: short chain ADH (PF00106), Fe-containing ADH (PF00465) methanol dehydrogenase was obtained from *Desulfotomaculum kuznetsovii* according to ref[31]. Sequences described in ref [43] were obtained from Uniprot and KEGG databases. A total of 301 sequences, including 5 from Brockarchaeota MAGs were aligned using MAFFT v7.450 (default parameters). The alignment was masked (50% Gaps) in Geneious Prime 2020.0.5 and manually refined. The phylogenetic tree was generated using a maximum likelihood-based approach using RAxML v8.2.10, called as: raxmlHPC-PTHREADS-AVX -f a -m PROTGAMMAAUTO -N autoMRE -p 12345 -x 12345 -s



Supporting Figure 4. Proposed pathway for butanol oxidation in Brockarchaeota genomes from hot springs where butanol dehydrogenase (BDH) was present. Brockarchaeota lack the set of enzymes involved in the fermentation of pyruvate to butanol, yet they encode a putative aldehyde dehydrogenase (PF00171) that could convert butyraldehyde to butyric acid. Red colors indicate the present enzymes in Brockarchaeota genomes. Orange colors indicate the number of Brockarchaeota genomes encoding that enzymatic step. The steps involved in further conversions of butyric acid still remain unresolved.

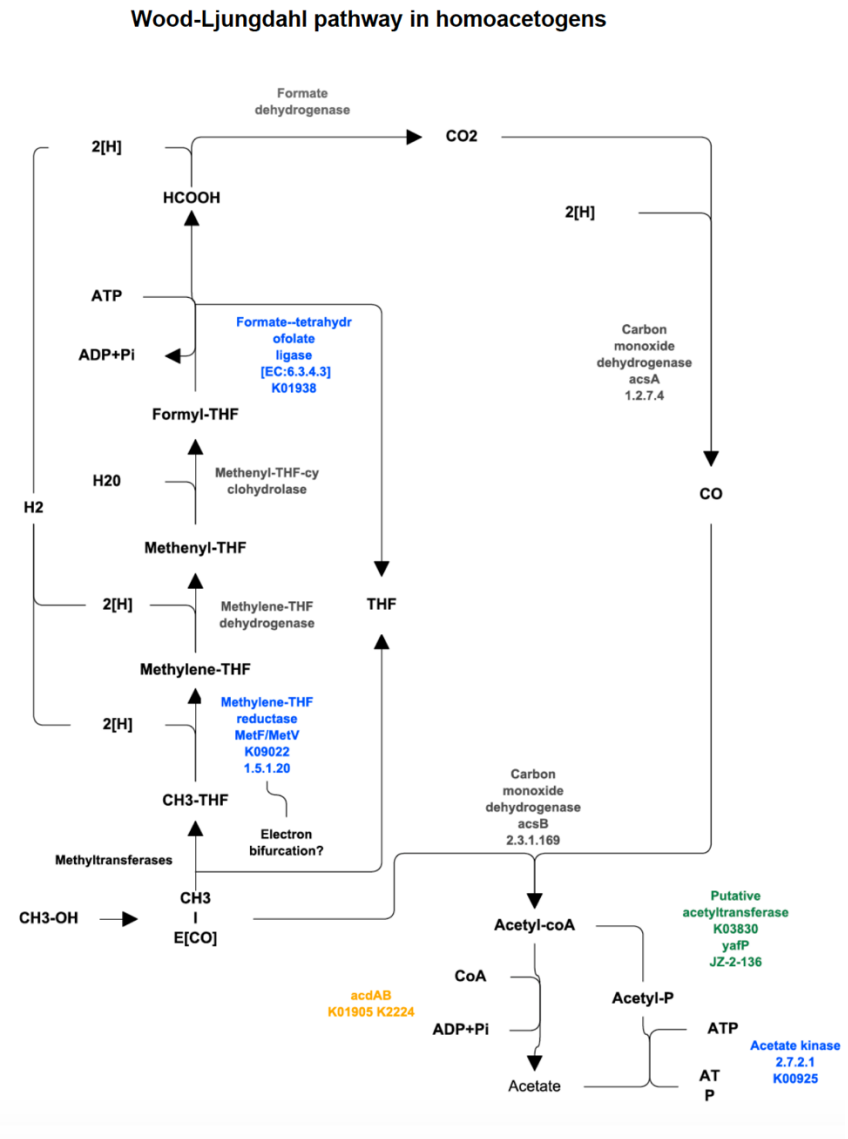


Supporting Figure 5. Glycolytic Embden-Meyerhof (EM) pathway found in Brockarchaeota genomes. Phosphomannomutase/Phosphoglucomutase (pgm) [EC:5.4.2.2] is a branch point between central metabolic pathways and the polysaccharide synthesis pathways. pgi glucose-6-phosphate isomerase, archaeal [EC:5.3.1.9] was found only present in JZ-2-136. K01622 is only present in B48_G17.



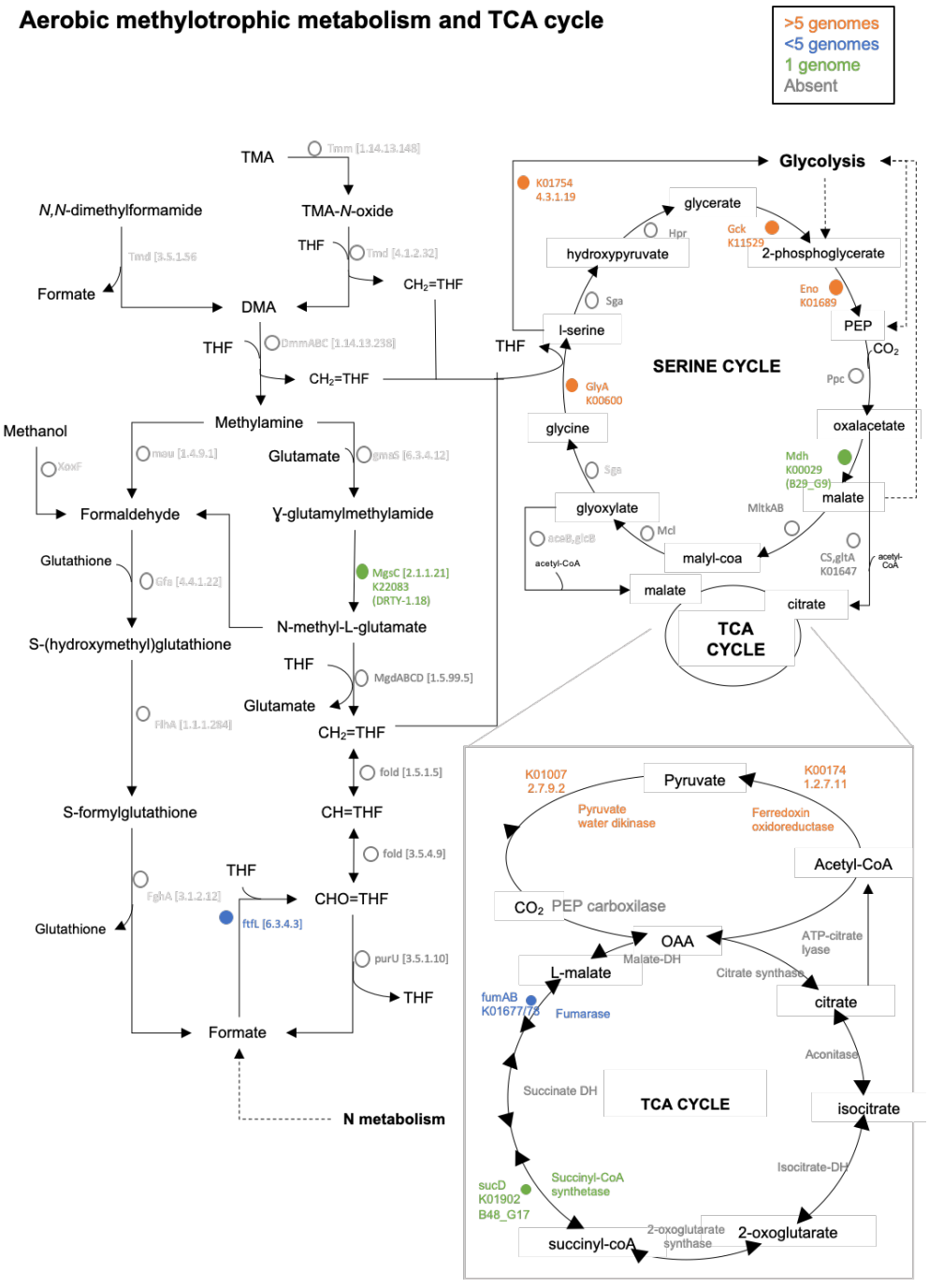
Supporting Figure 6. Classification and phylogeny of hydrogenases found in Brockarchaeota genomes. Maximum likelihood tree aligned with a dataset of hydrogenases belonging to each class encircled and colored according to each of the IV groups of hydrogenases described in ref[22,37]. Branches with red and green stars indicate hydrogenases found in Brockarchaeota genomes from deep sea sediments and hots springs respectively. A total of 41 hydrogenases were detected in Brockarchaeota MAGs and were aligned with 3250 reference hydrogenases using MAFFT v7.450 (default parameters) and redefined with MUSCLE v3.8.425 (default parameters). The alignment was masked (50% Gaps) in Geneious Prime 2020.0.5 and manually refined.The phylogenetic tree was generated using a maximum likelihood-based approach using RAxML

v8.2.10, called as: raxmlHPC-PTHREADS-AVX -f a -m PROTGAMMAAUTO -N autoMRE -p 12345 -x 12345 -s .



Supporting Figure 7. Canonical pathway of acetate synthesis from methanol in acetogens. Acetogens compete directly with hydrogenotrophic methanogenic archaea or interact syntrophically with acetotrophic methanogens that use H_2 and CO_2 to produce methane. The formation of acetate from methanol is only possible if other carbon containing compounds more oxidized than methanol - such as formate, CO, or CO_2 - are present. The first step of methanol conversion is catalyzed by a methyltransferase (MT) system present in Brockarchaeota. Tetrahydrofolate can serve as methyl acceptor in this reaction. The carbonyl and methyl branches of the Wood–Ljungdahl (WL) pathway are shown in the right and left panel respectively. The WL pathway is used in the reductive direction for energy conservation and autotrophic carbon assimilation in described

acetogens., however Brockarchaeota genomes lack the carbonyl branch required for carbon assimilation and the methyl branch is incomplete. Abbreviations: THF, tetrahydrofolate; CoA coenzyme A; CH3-E[Co], methyl coronoid (bound). Modified from ref [44]



Supporting Figure 8. Enzymatic steps involved in aerobic methylotrophy described in ref[45]. Malate dehydrogenase (oxaloacetate-decarboxylating) (NADP+) [EC:1.1.1.40] K00029 is only present only in one B27_G9 genome

Legends for Supporting tables

Supporting Table 1. *Excel File*. Physicochemical characteristics of samples where the MAGs in this study were obtained. n.a: not available data

Supporting Table 2. *Excel File* Selected 250 publicly available TACK genomes including 15 Brockarchaeota genomes reported in this study.

Supporting Table 3 *Excel File.* Sheet 1. Average amino acid identity (AAI) comparison of Brockarchaeota genomes and phylogenetically related TACK phyla. Sheet 2. Selected TACK genomes closely related to Brockarchaeota genomes. (Geoarchaeota, Aigarchaeota, Korarchaeota and Thaumarchaeota). See Supplementary Figure 1. The color code indicates the AAI (%) from 0-100 (green to red respectively)

Supporting Table 4. *Excel file* Metadata and geographic locations of 16S rRNA gene sequences related to Brockarchaeota genomes obtained from the Integrated Microbial Genomes and Microbiomes (IMG/M) database (https://img.jgi.doe.gov/).

Supporting Table 5. *Excel File*, Metabolism sheet: Overview of key metabolic genes found in Brockarchaeota genomes using KofamKOALA. Only hits above the predefined threshold for individual KOs were selected. Raw sheet contains the raw annotations obtained with KofamKOALA including the specific scaffolds for each annotation. MT system sheet contains the selected genes for the MT system and folD proteins detected with Interproscan v5.31-70.0. Metabolic Marker genes sheet contains the refence and absence profile of selected marker genes obtained from METABOLIC[46]. MEBS completeness sheet shows the completeness of PFAM domains in TACK genomes showed in Figure 1. MEBS mapping File sheet contains the full repertoire of protein domains involved in specific metabolic processes.

Supporting Table 6. *Excel file.* Description of the common core marker genes specific to methanogenic, anaerobic methanotrophic and short-chain alkane-oxidizing archaea recently described[34], across the TACK superphylum. Sheet 2 includes the raw data with the scaffolds encoding the specific marker genes and corresponding E values.

Supporting Table 7. *Excel file.* Metagenomic screening of Methyl-coenzyme M reductase (MCR) genes in metagenomic assemblies from which Brockarchaeota genomes where recovered

Supporting Table 8. *Excel File*. Total number of carbohydrate-active enzymes (CAZymes) detected in Brockarchaeota and TACK superphylum. Sheet 1. Number of total CAZYmes in Brockarchaeota genomes. Sheet 2 contains the raw output from the dbCAN webserver and PSORT results. Sheet 3 Comparison of unique and shared CAZYmes across members of the TACK superphylum. Only hits detected in more than two tools where selected for domain assignment. Total number of carbohydrate-binding modules (CBMs),

carbohydrate esterases (CE), glycoside hydrolases GH) and polysaccharide lyases (PL). Subcellular localization for CEs, GHs and PLs was determined using PSORTb.

Supporting Table 9. *Excel file*. Predicted transcription units of the scaffolds were methylenetetrahydrofolate reductase (NADPH) [EC:1.5.1.20] was found in Brockarchaeota genomes (JZZ-4 and JZ-1.89). The operon prediction based on intergenic distance of neighboring genes as well as the functional relationships of their protein coding products was computed with Operon Mapper [47]

References for supporting information.

1. 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.
2. 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).
3. Say, R. F. & Fuchs, G. Fructose 1,6-bisphosphate aldolase/phosphatase may be an ancestral gluconeogenic enzyme. *Nature* **464**, 1077–1081 (2010).
4. Orita, I. *et al.* The ribulose monophosphate pathway substitutes for the missing pentose phosphate pathway in the archaeon Thermococcus kodakaraensis. *J. Bacteriol.* **188**, 4698–4704 (2006).
5. 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).
6. 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).
7. Johnsen, U. & Schönheit, P. Novel xylose dehydrogenase in the halophilic archaeon Haloarcula marismortui. *J. Bacteriol.* **186**, 6198–6207 (2004).
8. Orsi, W. D. *et al.* Metabolic activity analyses demonstrate that Lokiarchaeon exhibits homoacetogenesis in sulfidic marine sediments. *Nat. Microbiol.* **5**, 248–255 (2020).
9. Mai, X. & Adams, M. W. W. Characterization of a fourth type of 2-keto acid-oxidizing enzyme from a hyperthermophilic archaeon: 2-Ketoglutarate ferredoxin oxidoreductase from Thermococcus litoralis. *J. Bacteriol.* **178**, 5890–5896 (1996).
10. Spang, A. *et al.* Proposal of the reverse flow model for the origin of the eukaryotic cell based on comparative analyses of Asgard archaeal metabolism. *Nat. Microbiol.* **4**, 1138–1148 (2019).
11. Cheng, J., Guffanti, A. A. & Krulwich, T. A. A two-gene ABC-type transport system that extrudes Na+ in Bacillus subtilis is induced by ethanol or protonophore. *Mol. Microbiol.* **23**, 1107–1120 (1997).
12. Pereira Cardoso, I. *et al.* A comparative genomic analysis of energy metabolism in sulfate reducing bacteria and archaea. **2**, 1–22 (2011).
13. 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).
14. 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 Sulphate-Reducing Prokaryotes. in *Adv Microb Physiol* vol. 66 55–321 (2015).
15. 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).
16. Stoffels, L., Krehenbrink, M., Berks, B. C. & Unden, G. Thiosulfate reduction in salmonella enterica is driven by the proton motive force. *J. Bacteriol.* **194**, 475–485 (2012).
17. Aussignargues, C. *et al.* Rhodanese functions as sulfur supplier for key enzymes in sulfur energy metabolism. *J. Biol. Chem.* **287**, 19936–19948 (2012).
18. Tóth, A., Takács, M., Groma, G., Rákhely, G. & Kovács, K. L. A novel NADPH-dependent oxidoreductase with a unique domain structure in the hyperthermophilic Archaeon, Thermococcus litoralis. *FEMS Microbiol. Lett.* **282**, 8–14 (2008).
19. Ma, K., Weiss, R. & Adams, M. W. W. Characterization of hydrogenase II from the hyperthermophilic archaeon Pyrococcus furiosus and assessment of its role in sulfur reduction. *J. Bacteriol.* **182**, 1864–1871 (2000).
20. Jenney, F. E. & Adams, M. W. W. Hydrogenases of the model hyperthermophiles. *Ann. N. Y. Acad. Sci.* **1125**, 252–266 (2008).
21. 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).
22. Greening, C. *et al.* Genomic and metagenomic surveys of hydrogenase distribution indicate H 2 is a widely utilised energy source for microbial growth and survival. *ISME J.* **10**, 761–777 (2016).
23. 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).
24. 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).
25. Girbal, L., Croux, C., Vasconcelos, I. & Soucaille, P. Regulation of metabolic shifts in Clostridium acetobutylicum ATCC 824. *FEMS Microbiol. Rev.* **17**, 287–297 (1995).
26. 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).
27. 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).
28. Chistoserdova, L. & Kalyuzhnaya, M. G. Current Trends in Methylotrophy. *Trends Microbiol.* **26**, 703–714 (2018).
29. Thauer, R. K., Kaster, A. K., Seedorf, H., Buckel, W. & Hedderich, R. Methanogenic archaea: Ecologically relevant differences in energy conservation. *Nat. Rev. Microbiol.* **6**, 579–591 (2008).
30. Yanagawa, K. *et al.* Biogeochemical Cycle of Methanol in Anoxic Deep-Sea Sediments. *Microbes Environ.* **31**, 190–193 (2016).
31. Sousa, D. Z. *et al.* The deep-subsurface sulfate reducer Desulfotomaculum kuznetsovii employs two methanol-degrading pathways. *Nat. Commun.* **9**, (2018).
32. Vanwonterghem, I. *et al.* Methylotrophic methanogenesis discovered in the archaeal phylum Verstraetearchaeota. *Nat. Microbiol.* **1**, 1–9 (2016).
33. Borrel, G. *et al.* Wide diversity of methane and short-chain alkane metabolisms in uncultured archaea. *Nat. Microbiol.* **4**, 603–613 (2019).
34. Rana, A. & Dey, A. Theoretical exploration of the mechanism of formylmethanofuran dehydrogenase: the first reductive step in CO2 fixation by methanogens. *J. Biol. Inorg. Chem.* **21**, 703–713 (2016).
35. Yan, Z. & Ferry, J. G. Electron bifurcation and confurcation in methanogenesis and reverse methanogenesis. *Front. Microbiol.* **9**, 1–10 (2018).

36. Gottschalk, G. & Thauer, R. K. The Na+-translocating methyltransferase complex from methanogenic archaea. *Biochim. Biophys. Acta - Bioenerg.* **1505**, 28–36 (2001).
37. Søndergaard, D., Pedersen, C. N. S. & Greening, C. HydDB: A web tool for hydrogenase classification and analysis. *Sci. Rep.* **6**, 1–8 (2016).
38. Moracci, M., Mose, B. C., Ii, F. & Universitario, C. Extremophiles Handbook. *Extrem. Handb.* (2011) doi:10.1007/978-4-431-53898-1.
39. Blumer-Schuette, S. E. *et al.* Complete genome sequences for the anaerobic, extremely thermophilic plant biomass-degrading bacteria Caldicellulosiruptor hydrothermalis, Caldicellulosiruptor kristjanssonii, Caldicellulosiruptor kronotskyensis, Caldicellulosiruptor owensensis, and Caldic. *J. Bacteriol.* **193**, 1483–1484 (2011).
40. Straub, C. T. *et al.* Quantitative fermentation of unpretreated transgenic poplar by Caldicellulosiruptor bescii. *Nat. Commun.* **10**, 1–6 (2019).
41. Yang, S. J. *et al.* Classification of 'Anaerocellum thermophilum' strain DSM 6725 as Caldicellulosiruptor bescii sp. nov. *Int. J. Syst. Evol. Microbiol.* **60**, 2011–2015 (2010).
42. Sista Kameshwar, A. K. & Qin, W. Understanding the structural and functional properties of carbohydrate esterases with a special focus on hemicellulose deacetylating acetyl xylan esterases. *Mycology* **9**, 273–295 (2018).
43. Radianingtyas, H. & Wright, P. C. Alcohol dehydrogenases from thermophilic and hyperthermophilic archaea and bacteria. *FEMS Microbiol. Rev.* **27**, 593–616 (2003).
44. Goorissen, H. P. Thermophilic methanol utilization by sulfate reducing bacteria. 92 (2002).
45. Dziewit, L. *et al.* Genome-guided insight into the methylotrophy of Paracoccus aminophilus JCM 7686. *Front. Microbiol.* **6**, 1–13 (2015).
46. Zhou, Z., Tran, P., Liu, Y., Kieft, K. & Anantharaman, K. METABOLIC: A scalable high-throughput metabolic and biogeochemical functional trait profiler based on microbial genomes. *bioRxiv* 761643 (2019) doi:10.1101/761643.
47. 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).