Metabolic potential for reductive acetogenesis and a novel energy-converting [NiFe] hydrogenase in *Bathyarchaeia* from termite guts – a genome-centric analysis

Hui Qi Loh[1], Vincent Hervé[1], Andreas Brune[1*]

[1]Research Group Insect Microbiology and Symbiosis, Max Planck Institute for Terrestrial Microbiology, Marburg, Germany

*** Correspondence:**
Andreas Brune
brune@mpi-marburg.mpg.de

ORCID:

Hui Qi Loh: https://orcid.org/0000-0002-4613-2477

Vincent Hervé: http://orcid.org/0000-0002-3495-561X

Andreas Brune: http://orcid.org/0000-0002-2667-4391

Keywords: Bathyarchaeota, Homoacetogens, Wood–Ljungdahl pathway, Termites, Gut microbiota, Comparative genomics, Metagenome-assembled genomes.

Abstract

Symbiotic digestion of lignocellulose in the hindgut of higher termites is mediated by a diverse assemblage of bacteria and archaea. During a large-scale metagenomic study, we reconstructed 15 metagenome-assembled genomes (MAGs) of *Bathyarchaeia* that represent two distinct lineages in subgroup 6 (formerly MCG-6) unique to termite guts. One lineage (TB2; *Candidatus* Termitimicrobium) encodes all enzymes required for reductive acetogenesis from H_2 and CO_2 via an archaeal variant of the Wood–Ljungdahl pathway. This includes a novel 11-subunit hydrogenase, which possesses the genomic architecture of the respiratory Fpo-complex of other archaea but whose catalytic subunit is phylogenetically related to and shares the conserved [NiFe] cofactor-binding motif with [NiFe] hydrogenases of subgroup 4g. We propose that this novel Fpo-like hydrogenase provides the reduced ferredoxin required for CO_2 reduction and is driven by the electrochemical membrane potential generated from the ATP conserved by substrate-level phosphorylation. Members of the other lineage (TB1; *Candidatus* Termiticorpusculum) are not capable of lithotrophic acetogenesis because they consistently lack hydrogenases and/or methylene-tetrahydromethanopterin reductase, a key enzyme of the pathway. Both lineages have the genomic capacity to reduce ferredoxin by oxidizing amino acids and might conduct methylotrophic acetogenesis using unidentified methylated compound(s). Our results indicate that *Bathyarchaeia* of subgroup 6 contribute to acetate formation in the guts of higher termites and substantiate the genomic evidence for reductive acetogenesis from organic substrates, including methylated compounds, in other uncultured representatives of the phylum.

1 Introduction

Although *Bathyarchaeia* are widespread in anoxic environments, their physiology is only poorly understood. In the absence of any isolates and only few microscopic observations of their cells (Collins *et al.*, 2005, Kubo *et al.*, 2012), our knowledge about this deep-branching lineage is based almost exclusively on amplicon libraries of archaeal 16S rRNA genes and metagenomic studies (reviewed by Zhou *et al.* (2018)).

Ribosomal RNA genes affiliated with the "Miscellaneous Crenarchaeotal Group" (MCG) had already been recovered in early analyses of archaeal diversity in diverse anoxic habitats (e.g., Schleper *et al.*, 1997, Inagaki *et al.*, 2003, Ochsenreiter *et al.*, 2003), including the intestinal tract of termites (Friedrich *et al.*, 2001). Meanwhile, an enormous diversity of sequences from this group, which comprises numerous deep-branching lineages, has been recovered from a wide range of marine and freshwater habitats and terrestrial environments (e.g., Kubo *et al.*, 2012, Fillol *et al.*, 2016). A few years ago, the MCG was elevated to the phylum level (*Bathyarchaeota*; Meng *et al.*, 2014), but the most recent genome-based taxonomy demoted them again to the class level (*Bathyarchaeia*; Rinke *et al*., 2020). While the rank of the taxon is not relevant in the current context, we maintained the subgroup numbering used in previous studies (e.g., Kubo *et al.*, 2012; Lazar *et al.*, 2016) but replaced the prefix 'MCG-' with the prefix 'Bathy-' (Yu *et al.*, 2018).

The abundance of *Bathyarchaeia* in many anoxic habitats implies potentially important roles in biogeochemical cycles (Evans *et al.*, 2015; He *et al.*, 2016). Reconstruction of metagenome-assembled genomes (MAGs) provided information concerning the metabolic capacities of *Bathyarchaeia* and inspired predictions of their putative roles in anoxic sediments (reviewed by Zhou *et al.*, 2018). Several studies suggested that *Bathyarchaeia* are organotrophic and utilize a variety of organic substrates (e.g., Meng *et al.*, 2014; He *et al.*, 2016; Lazar *et al.*, 2016). The discovery of genes encoding a methyl-coenzyme M reductase (Mcr) complex and a complete Wood–Ljungdahl pathway in bathyarchaeon BA1 provided the first evidence of methanogenesis outside the Euryarchaeota (Evans *et al.*, 2015). Other studies detected key enzymes of the pathway in bathyarchaeal genomes of several subgroups and proposed that these lineages are involved in reductive acetogenesis from CO_2 (He *et al.*, 2016, Lazar *et al.*, 2016).

Considering the putative roles of *Bathyarchaeia* in methanogenesis and reductive acetogenesis and the evidence for the utilization of lignin-derived methoxy groups (Yu *et al.*, 2018), the presence of this group in termite guts is intriguing. Termites efficiently digest wood and other lignocellulosic substrates, either sound or in different stages of humification (Brune, 2014), in symbiosis with a specialized gut microbiota housed in their enlarged hindgut compartments (Brune and Dietrich, 2015). Hydrogen produced in microbial fermentation processes serves as electron donor for the reduction of CO_2, yielding acetate and methane as major products (Breznak and Switzer, 1986; Brauman et al., 1992). Methanogenesis in termite guts involves a diverse assemblage of hydrogenotrophic and methyl-reducing archaea (Brune, 2018), but reductive acetogenesis, which can contribute up to two-thirds of total acetate production, has so far been considered a bacterial activity.

In lower termites, reductive acetogenesis has been attributed to homoacetogenic members of the phylum *Spirochaetes*) (e.g., Leadbetter *et al.*, 1999, Ohkuma *et al.*, 2015) and a novel lineage of uncultured *Deltaproteobacteria* (Rosenthal *et al.,* 2013; Ikeda-Ohtsubo *et al.*, 2016). In higher termites (family Termitidae), which diverged from the lower termites about 50 million years ago (Bucek *et al.*, 2019), the situation is more complex. Particularly in the humus-feeding and soil-feeding groups, where the potential rates of reductive acetogenesis decrease in favor of methanogenesis (Brauman *et al.*, 1992;

Tholen and Brune, 1999), spirochetes are less abundant than in wood-feeding groups (Mikaelyan *et al.*, 2016). A study based on the formyltetrahydrofolate synthetase (FTHFS) gene, a key enzyme of the Wood–Ljungdahl pathway that has been used as a marker for reductive acetogenesis, indicated that the community of potential acetogens shifts from spirochetes in lower termites to clostridia in higher termites (Ottesen and Leadbetter, 2011).

In a large-scale metagenomic study of the gut microbiota of eight higher termites, we obtained 15 metagenome-assembled genomes (MAGs) assigned to *Bathyarchaeia* (Hervé *et al.*, 2020). Preliminary analysis revealed that they fell into a cluster comprising mainly termite gut MAGs, with members of Bathy-1 and Bathy-6 as next relatives. Here, we conducted detailed phylogenomic analyses of these MAGs and investigated their potential capacity for methanogenesis and reductive acetogenesis using a genome-centric approach.

2 Results and Discussion

2.1 Phylogeny of termite gut *Bathyarchaeia*

Bathyarchaeal MAGs were recovered from seven of the eight higher termites investigated, regardless of their feeding group (Hervé *et al.*, 2020; Table 1). Their absence from *Microcerotermes parvus* is most likely caused by the low total number of MAGs obtained from the metagenomes of this species. Based on average nucleotide identity, the MAGs were assigned to nine phylotypes (Table 1). MAGs of the same phylotype were always derived from different gut compartments of the same host species, indicating that they most likely represent bathyarchaeal populations distributed along the entire hindgut. Eleven of the 15 MAGs fulfill the criteria for high-quality MAGs (>90% complete and <5% contamination; Bowers *et al.*, 2017). Except for phylotype 5, each phylotype is represented by at least one high-quality MAG, which allows robust inference of metabolic potentials (Nelson *et al.*, 2020).

Phylogenomic analysis placed all phylotypes from termite guts within subgroup Bathy-6, an apical lineage of *Bathyarchaeia* that is well represented mostly in 16S rRNA gene libraries (He *et al.*, 2016) but comprises only a few MAGs from marine or estuarine sediments and the deep subsurface (Figure 1). The MAGs from termite guts form two distinct lineages, TB1 (phylotypes 1–7) and TB2 (phylotypes 8 and 9). TB2 is a sister group of bathyarchaeon SZUA-568 (hereafter denoted as Bathy-6-S), a MAG retrieved from marine hydrothermal vent sediments. Other MAGs in the radiation of Bathy-6 are bathyarchaea BE326-BA-RLH (hereafter denoted as Bathy-6-B) and AD8-1 (hereafter denoted as Bathy-6-A). They are all high-quality MAGs and were included in the subsequent analyses (Table 1). Only bathyarchaeon SG8-32-3 (previously assigned to Bathy-1) was omitted because the completeness of the assembly (50.4%; based on our CheckM analysis) was too low for a reliable assessment of its metabolic capacity.

Predicted genome sizes (1.0 – 2.5 Mbp), G+C contents (37.4 – 43.4 mol%), and coding densities (74.9 – 86.9%) of the MAGs from termite guts are in the same range as those of the other representatives of this subgroup (Table 1). While the average nucleotide identity (ANI) values among the phylotypes of TB1 and TB2 ranges between 78.1 and 81.6%, the ANI values between members of TB1, TB2, and the other phylotypes of Bathy-6 are below the cut-off of the fastANI tool (< 75%; Supplementary Figure S1), indicating that each lineage represents a separate genus-level taxon. This is confirmed by the results obtained with the GTDB toolkit, which classified members of TB1 and

TB2 as separate, genus-level lineages in the family UBA233 (order B26-1), a family that comprises also other members of Bathy-6. This indicates that TB1 and TB2 represent novel candidate genera in family UBA233, for which the names '*Candidatus* Termiticorpusculum' and '*Candidatus* Termitimicrobium' are proposed.

To identify the closest relatives of termite gut *Bathyarchaeia* and their respective habitats, we analyzed their phylogenetic position in the framework of rRNA genes available in public databases, which provides a much better coverage than the small number of MAGs of the Bathy-6 subgroup available to date (Figure 2). The 16S rRNA gene sequences encoded by the MAGs form a well-supported monophyletic group with all other sequences of *Bathyarchaeia* that were previously obtained from the hindguts of higher termites (Friedrich *et al.*, 2001; Shi *et al.*, 2015; Grieco *et al.*, 2019). Although each ribotype appears to be specific for a particular host species, the internal topology of the termite clade is not well resolved due to the large number of short sequences and the absence of 16S rRNA genes from many MAGs. The sequences in the termite clade are most closely related to clones obtained from a manure pit (EU662668; J. Ding, unpublished) and an anaerobic digestor fed with vinasses (U81774; Godon et al., 1997), and fall into the radiation of bathyarchaeal lineages in freshwater sediments, salt marshes, and anaerobic wastewater bioreactors (group 1.3b; Ochsenreiter *et al.*, 2003; Collins *et al.*, 2005).

2.2 Capacity for CO_2-reductive acetogenesis

We investigated the presence of all genes required for methanogenesis and reductive acetogenesis in all members of Bathy-6 with sufficiently complete genomes (Figure 3). All members of TB2 (phylotypes 8 and 9) encode the complete set of genes required for the reduction of CO_2 to acetyl-CoA via the archaeal version of the Wood–Ljungdahl pathway, using methanofuran (MFR) and tetrahydromethanopterin (H_4MPT) as C_1 carriers (Figure 4). Formyl-MFR dehydrogenase is molybdenum-dependent (FmdABCDF; Hochheimer *et al.*, 1996) and not the tungsten-dependent paralog. A homolog of *fmdE*, which occurs in methanogens, was not found in any of the MAGs, which suggests that the absence of subunit E is a characteristic feature of the bathyarchaeal complex. It has been shown that the Fmd complexes of *Methanobacterium thermoautotrophicum* and *Methanosarcina barkeri* are active also without this subunit (Hochheimer *et al.*, 1996, Vorholt *et al.*, 1996). The CO dehydrogenase/acetyl-CoA synthase complex (CdhABCDE) and the (ADP-forming) acetyl-CoA synthetase (Acd; Musfeldt *et al.*, 1999) are typical archaeal enzymes.

Enzymes characteristic for the bacterial Wood–Ljungdahl pathway (FTHFS, methylene-THF cyclohydrolase/dehydrogenase, and methylene-THF reductase), which had been identified in MAGs of Bathy-3, -8 and -17 (Evans *et al.*, 2015, Zhou *et al.*, 2018), were not encoded by any member of Bathy-6. Also, phosphate acetyltransferase and acetate kinase, which are responsible for substrate-level phosphorylation in fermenting bacteria, were absent from all MAGs.

The same gene sets as in TB2 are encoded also by the more basal Bathy-6-S and Bathy-6-B (Figure 3), which indicates that the capacity to produce acetate from CO_2 might be a plesiomorphic trait of the Bathy-6 subgroup. The consistent absence of a key enzyme of the archaeal Wood–Ljungdahl pathway, methylene-H_4MPT reductase (Mer), from all seven phylotypes (11 MAGs) of the TB1 lineage and from the most basal member of the subgroup, Bathy-6-A, suggests that the capacity to reduce CO_2 to the methyl level was lost at least twice during the evolutionary radiation of Bathy-6.

Homologs of the methyl-coenzyme M reductase (Mcr) complex, which encodes the key enzyme of methanogenesis, were not detected in any of the MAGs. Our observation contrasts with the report of

Harris *et al.* (2018), who claimed that Bathy-6-B might represent an anaerobic methane oxidizer. However, their conclusion is based on the recovery of a 265-bp gene fragment classified as an *mcrA* gene in the original metagenome from which Bathy-6-B was assembled, i.e., not from the metagenomic bin. Considering also that the gene fragment in question shows highest similarity to a homolog from an uncultured euryarchaeal methanogen (GenBank: JX907770.1), it seems safe to conclude that members of the Bathy-6 subgroup are not methanogenic.

Although the capacity of *Bathyarchaeia* for reductive acetogenesis from CO_2 has been claimed repeatedly for several subgroups (He *et al*., 2016; Lazar *et al.*, 2016; Zhou *et al.*, 2018; Yu *et al.*, 2018), the evidence was never fully conclusive. Actually, the recent, comprehensive survey of all bathyarchaeal MAGs compiled by Zhou *et al.* (2018) lists only two MAGs that encode all genes required to operate the entire Wood–Ljungdahl pathway. One is the putatively methanogenic BA1 (Bathy-8) from a deep aquifer (Evans *et al.*, 2015); the other is bathyarchaeon ex4484_135 (Bathy-15) from marine hydrothermal sediment (Dombrowski *et al.*, 2017).

2.3 Capacity for methylotrophic acetogenesis

Since all members of Bathy-6 encode a complete CO dehydrogenase/acetyl-CoA synthase (Cdh) complex (Figure 3), they might still synthesize acetyl-CoA using methyl groups derived from external sources. In all homoacetogenic bacteria and methylotrophic methanogens studied to date, the methyl transferase systems consist of three components: (i) a set of substrate-specific methyl transferases (MT-I), (ii) their cognate methyl-accepting corrinoid proteins (CoP), and (iii) a second methyl transferase (MT-II) that transfers the methyl group of methyl-CoPs to THF (bacteria) or coenzyme M (archaea) (van der Meijden *et al.*,1983; Kreft and Schink, 1994; Kremp *et al.*, 2018; Supplementary Figure S4A). We found that all MAGs of Bathy-6 encode CoPs that fall into the radiation of homologs assigned to other uncultured Archaea, with the CoPs of the di- and trimethylamine-specific methyltransferase systems (MtbC and MttC) of *Methanomassiliicoccus luminyensis* (Kröninger *et al.*, 2017) and *Acetobacterium woodii* (Kremp *et al.*, 2018) as closest relatives with a reliable functional annotation (Supplementary Figure S4). However, unlike the situation in methylotrophic bacteria and euryarchaea, where the CoP gene is co-localized with the gene of the cognate substrate-specific MT-I homologs (MtbB or MttB), the CoP gene of Bathy-6 is flanked by a gene encoding subunit H of tetrahydromethanopterin *S*-methyltransferase (MtrH; Supplementary Figure S4B).

In many methanogenic archaea, MtrH is part of the energy-conserving MtrA–H complex and catalyzes the transfer of the (CO_2-derived) methyl group from methyl-tetrahydromethanopterin to the corrinoid prosthetic group of MtrA (Hippler and Thauer, 1999). However, in obligately methyl-reducing methanogens (Galagan *et al.*, 2002; Borrel *et al.*, 2014; Lang *et al.*, 2015), which methylate CoM via their diverse methyltransferase systems (see above), the Mtr complex is absent. The presence of an isolated *mtrH* gene colocalized with a CoP gene has been observed also in the putatively methanogenic BA1 and BA2 (*Bathyarchaeia*) and several MAGs related to '*Ca.* Methanomethylicus mesodigestum' (*Thermoproteota*). It was proposed that the encoded proteins represent methyltransferase systems, which prompted the hypothesis that these uncultured lineages are methylotrophic methanogens (Evans *et al.* 2015; Vanwonterghem *et al.* 2016).

It is tempting to assume that also the CoP–MtrH couple of Bathy-6 is involved in the transfer of methyl groups from so-far unidentified, substrate-specific methyltransferases to H_4MPT (Figure 4). However, a catabolic role of the CoP–MtrH couple is not the only possible interpretation. In '*Ca.* Methanomethylicus mesodigestum', the genes are colocalized with a homolog of *metE* encoding

methionine synthase (Supplementary Figure S4B), it is also possible that the CoP–MtrH couple of Bathy-6 is involved in anabolic reactions that transfer methyl groups (provided by the cleavage of acetyl-CoA) from H_4MPT to an unknown acceptor.

2.4 Hydrogen as electron donor

The operation of the Wood–Ljungdahl pathway requires electron donors in the form of reduced ferredoxin, NADH, and in the case of archaea, also reduced cofactor F_{420} ($F_{420}H_2$). The reduction of ferredoxin with H_2 is a critical step because it is endergonic at low hydrogen partial pressures and requires either an energy-converting hydrogenase or a flavin-based electron bifurcation system (Thauer *et al.*, 2008, Buckel and Thauer, 2013).

Hydrogenases are present only in TB2 and the basal lineages of TB1 (Figure 3). One is a cytosolic, bidirectional [NiFe] hydrogenase of Subgroup 3d, which use NAD as electron acceptor (Greening *et al.*, 2016). Phylogenetic analysis of the gene encoding the large subunit (*hoxH*) placed all homologs in a sister position to the Hox hydrogenases of phototrophic bacteria (Supplementary Figure S5). The gene order in the *hoxEFUYH* cluster is the same as in the gene clusters of other Hox complexes, which encode a prototypical heterodimeric [NiFe]-hydrogenase moiety (HoxHY) and a diaphorase moiety (HoxEFU); HoxEFU is homologous to the NuoEFG module of complex I and mediates the electron transport to NAD(P) (Eckert et al., 2012). Although members of Group 3 are called "bidirectional hydrogenases", hydrogen formation requires reduced ferredoxin or flavodoxin as electron donor (Gutekunst et al. 2014).

All MAGs that encode a Hox hydrogenase also possess a gene cluster that resembles that encoding the F_{420}:methanophenazine oxidoreductases (Fpo) of Euryarchaeota and the NADH:quinone oxidoreductases (Nuo) of Bacteria (complex I) (Figure 5). As in other basal lineages of complex I homologs, the FpoFO and NuoEFG modules, which provide substrate specificity for $F_{420}H_2$ or NADH, respectively, are absent (Moparthi and Hägerhäll, 2011).

However, six of the 11 subunits common to the Fpo and Nuo complexes are homologous to subunits of the energy-converting [NiFe] hydrogenases of Group 4, which are ancestral to the respiratory complex I (Friedrich and Scheide, 2000). Classification with HydDB placed the D subunit of the 11-subunit complex of the Bathy-6 MAGs among the catalytic subunits of [NiFe] hydrogenases in Subgroup 4g. The hydrogenases in Subgroup 4g are not only structurally heterogeneous and differ fundamentally both in the number of their subunits and the arrangement of their coding genes (Greening *et al.*, 2016; Schoelmerich and Müller, 2019; Figure 5), but their large subunits also fall into separate phylogenetic lineages (Groups 4g-1 to 4g-6; Figure 6). The genomic architecture of the Fpo-like hydrogenase complex of Bathy-6 (hereafter referred to as Hfo) closely resembles that of *Ca.* Methanomethylicus mesodigestum (*Thermoproteota*) and *Pyrodictium delaney* (*Crenarchaeota*) (Figure 5), whose large subunits represent phylogenetic sister groups (4g-5 and 4g-6) that are distinct from the other lineages (Figure 6). Interestingly, the complex of *Thermosphaera aggregans* (Subgroup 4g-5) and other members of *Desulfurococcales* (not shown) seems to deviate from the Hfo-like structure and contains homologs of the Mbh complex of *Pyrococcus furiosus* (Figure 5).

None of the hydrogenases of Subgroup 4g have been biochemically characterized, but they are presumed to couple the formation of H_2 from reduced ferredoxin to the formation of an electrochemical membrane potential (Greening *et al.*, 2016; Søndergaard et al. 2016; Schoelmerich and Müller, 2019). This is in agreement with biochemical data obtained for the Fpo-like 11-subunit complex of methanogenic *Euryarchaeota*, which generate an electrochemical membrane potential

during electron transport from reduced ferredoxin to methanophenazine (*Methanosaeta*; Welte and Deppenmeier, 2011) or a so far unidentified electron acceptor (*Methanomassiliicoccales*; Kröninger *et al.*, 2016).

The presence of an Fpo-like 11-subunit hydrogenase in *Bathyarchaeia* is most interesting from an evolutionary perspective, since it represents the first [NiFe] hydrogenase with the genomic architecture of complex I. The coordination sites of the [NiFe] cofactor on the large subunit of all [NiFe] hydrogenases (L1 and L2 motifs; Vignais and Billoud, 2007), which are no longer conserved in NuoD and FpoD, are present in all Bathy-6 homologs (Figure 7). This adds to the evidence that the 11-subunit complex of *Bathyarchaeia* is not a respiratory complex but is instead a novel energy-converting hydrogenase that catalyzes the reduction of ferredoxin with H_2 using the electrochemical membrane potential (Figure 4).

While Hox and Fpo-like hydrogenase in TB2 should provide the NADH and reduced ferredoxin required to operate the Wood–Ljungdahl pathway, the source of $F_{420}H_2$ as potential electron donor for methylene-H_4MPT reductase (Mer) remains unclear. A complete gene set encoding F_{420}-reducing [NiFe] hydrogenase (FrhABG, subgroup 3a; Supplementary Figure S5) is present only in Bathy-6-A. All members of TB2 and several phylotypes of TB1 encode a homolog of FrhB, an iron–sulfur flavoprotein with an F_{420}-binding site, but not the hydrogenase subunits (Figure 3). It is possible that FrhB is involved in the reduction of F_{420} via an interaction with HdrABC, as proposed for the methane-oxidizing *Ca.* Methanoperedens spp. (Arshad *et al.*, 2015).

The only member of subgroup Bathy-6 that encodes a complete FrhABG is Bathy-6-A. It is also the only MAG that encodes a methylviologen-dependent [NiFe] hydrogenase (MvhADG; Supplementary Figure S5, subgroup 3c), which forms an electron-bifurcating complex with the soluble heterodisulfide reductase (HdrABC) and catalyzes the hydrogen-dependent reduction of ferredoxin and the heterodisulfide of coenzyme M (CoM) and coenzyme B (CoB) in methanogens (Kaster *et al.*, 2011). The presence of genes encoding HdrABC, MvhADG, and a complete Wood–Ljungdahl pathway in the putatively methanogenic BA1 (Bathy-3) provides strong evidence that BA1 is capable of hydrogenotrophic methanogenesis (Evans *et al.,* 2015). In Bathy-6-A, however, the pathway is incomplete and the identity of the heterodisulfide reduced by Hdr remains unclear. Interestingly, the same constellation as in Bathy-6 has been recently reported for the bathyarchaeal MAG CR_14 from marine sediments, which represents another, novel subgroup of *Bathyarchaeia* (Farag *et al.*, 2020).

2.5 Organic substances as electron donors

Most members of TB1 and all basal lineages of Bathy-6 lack Hox and Fpo-like hydrogenase (Figure 3), which means that they cannot grow lithotrophically with H_2 as electron donor. However, the reduced Fd required to operate acetogenesis, either via the Wood–Ljungdahl pathway (TB2) or by methylotrophy (all phylotypes), could be provided also by the oxidation of organic substrates (Figure 4). Such organotrophic acetogenesis is common among bacterial homoacetogens (Schink, 1994; Drake, 1994). All Bathy-6 genomes (except Bathy-6-A) encode pyruvate:ferredoxin oxidoreductase (Por) and indolepyruvate:ferredoxin oxidoreductase (Ior), and some also encode 2-oxoglutarate:ferredoxin oxidoreductase (Oor), all of which catalyze the oxidative decarboxylation of 2-oxo acids to their corresponding acyl-CoA esters (Figure 3). The 2-oxo-acids would result from the transamination of amino acids via numerous aminotransferases encoded by all genomes; a putative amino acid permease, however is encoded only in TB2. ATP would be formed via the ADP-dependent acetyl-CoA synthetase, which accepts also other acyl substrates in *Pyrococcus furiosus* (Mai and Adams, 1996). Such pathways have been shown to operate in other archaea (*Pyrococcus*

297 *furiosus, Thermococcus* spp.; Kengen and Stams, 1993, Heider *et al.*, 1996) and in the insect gut-
298 associated bacterium *Elusimicrobium minutum* (Herlemann, *et al.,* 2009) during growth on glucose,
299 where they result in a net formation of alanine.

300 The data compiled by Zhou *et al.* (2018) suggest that several lineages of *Bathyarchaeia*, including
301 Bathy-6-A; Lazar *et al.*, 2016), have the capacity to ferment various organic carbon compounds.
302 However, genes encoding extracellular peptidases, which are numerous in other *Bathyarchaeia*, seem
303 to be less prevalent in the MAGs of Bathy-6 and Bathy-1 (Feng *et al.*, 2019), which suggests that
304 members of these subgroups are limited to the utilization of amino acids or oligopeptides that are
305 small enough to be transported across the cytoplasmic membrane.

306 A capacity of members of Bathy-6 to utilize sugars is not as apparent. Like Bathy-6-A (Lazar *et al.*,
307 2016), all MAGs of TB1 and TB2 encode many genes of the classical Embden-Meyerhof-Parnas
308 (EMP) pathway, including glyceraldehyde-3-phosphate dehydrogenase and phosphoglycerate kinase.
309 However, all MAGs lack hexokinase, the alternative archaeal glycolytic enzymes (Bräsen et al.,
310 2014), and most MAGs lack phosphofructokinase and pyruvate kinase. It is possible that EMP
311 pathway functions only in gluconeogenesis; fructose bisphosphatase is present in all MAGs. Sugar
312 transporters were not detected; the role of the lipooligosaccharide ABC transporter encoded by
313 almost all phylotypes from termite guts (except phylotype 9) is not clear (Supplementary Table S3).
314 The identification of a cellulolytic system in Bathy-6-A (Lazar et al., 2016) requires verification.

315 **2.6 Energy conservation**

316 In acetogenic bacteria growing on hydrogen and CO_2, all ATP synthesized by substrate-level
317 phosphorylation is consumed in the activation of formate. Therefore, energy conservation involves
318 electron-transport phosphorylation, which is driven by the oxidation of reduced ferredoxin via
319 membrane-bound electron-transport complexes (Schuchmann and Müller, 2014; Basen and Müller,
320 2017). By contrast, the activation of formate (i.e., the formation of formylmethanofuran) in the
321 archaeal variant of the Wood–Ljungdahl pathway is not ATP-dependent but is instead driven by the
322 reducing power of ferredoxin, yielding a full ATP per acetate produced via substrate-level
323 phosphorylation. However, thermodynamics dictates that a fraction of this ATP has to be reinvested,
324 because a metabolism where the net ATP yield exceeds the free-energy change of the reaction would
325 become endergonic (Thauer et al., 2008).

326 Fermenting bacteria that lack respiratory chains energize their membrane by operating their ATP
327 synthase in the reverse direction (Buckel and Thauer, 2013). Likewise, members of Bathy-6 that have
328 the capacity to grow lithotrophically on hydrogen and CO_2 (i.e., the phylotypes in TB2) might use
329 part of the ATP gained by substrate-level phosphorylation to generate an electrochemical membrane
330 potential that drives the reduction of ferredoxin via the Fpo-like hydrogenase complex (see above).
331 Other energy-converting complexes that would allow generation of reduced ferredoxin, such as the
332 Group-4 [NiFe] hydrogenases in acetogenic bacteria and methanogenic archaea (Ech, Künkel *et al.*,
333 2001; Eha and Ehb, Tersteegen and Hedderich, 2001) or an NADH:Fd oxidoreductase complex
334 (RnfABCDEG, Westphal *et al.*, 2018) were not detected in any member of Bathy-6.

335 During fermentative growth, the Fpo-like hydrogenase complex (if present) might operate in the
336 reverse direction, using reduced ferredoxin provided by the oxidation of organic substrates to
337 produce H_2 and generate an electrochemical membrane potential, like many fermenting bacteria with
338 an energy-converting hydrogenase. In that context, it is intriguing that several phylotypes of TB1 and
339 TB2 (Figure 3), and also bathyarchaeal MAGs from other subgroups (Evans *et al.*, 2015; Zhou *et al.*,

2018), do not encode an ATP synthase (neither the genes for the archaeal V-type ATP synthase nor for the bacterial equivalent were detected). If one disregards the possibility of incomplete genome assemblies, these organisms must generate their membrane potential (vital for any organism) by other means.

In principle, the Wood–Ljungdahl pathway is reversible and can also oxidize acetate to H_2 and CO_2 given the appropriate thermodynamic framework. This has been demonstrated in syntrophic cultures of "Reversibacter"-like microorganisms with hydrogenotrophic partners (Lee and Zinder, 1988; Schnürer et al. 1997) and has been suggested to occur also in *Bathyarchaeia* (Evans *et al.*, 2015; Xiang *et al.*, 2017). However, at least in the termite hindgut, where the hydrogen partial pressure is much higher than in sediments (Ebert and Brune, 1997; Schmitt-Wagner et al., 1999) and reductive acetogenesis often prevails over methanogenesis as electron sink (Brauman et al., 1992; Tholen et al. 1999, Tholen and Brune, 2000), an anaerobic oxidation of acetate is an unlikely scenario.

2.7 Ecological aspects

Although the proportion of archaeal rRNA in termite hindguts is relatively small (0.9–2.3% of all prokaryotic rRNA; Brauman *et al.*, 2001), methanogenesis represents a substantial hydrogen sink (Brune, 2019). Considering that the proportion of reads assigned to bathyarchaeal MAGs in the hindgut metagenomes of higher termites (0.03–2.5%; avg. 0.69%) is four times higher than that assigned to euryarchaeal MAGs (0.02–0.79%; avg. 0.16%; Table S2 in Hervé *et al.*, 2020), the population sizes of *Bathyarchaeia* might be sufficient to contribute significantly to acetogenesis, particularly in soil-feeding species.

However, the substrates of termite gut *Bathyarchaeia* remain open to speculation. While only members of TB2 have the genomic capacity for lithotrophic acetogenesis, almost all members of Bathy-6 have the capacity to ferment amino acids and might employ organotrophic acetogenesis from methylated substrates as an electron sink. Stable-isotope probing of salt marsh sediments indicated that members of Bathy-8 and Bathy-6 assimilate organic substrates, notably excluding proteins and inorganic carbon (Seyler *et al.*, 2014). Yu *et al.* (2018), however, reported that addition of lignin to an estuarine sediment sample selectively stimulated the growth of Bathy-8 and the incorporation of carbon from ^{13}C-bicarbonate into archaeal tetraether lipids, which suggests that members of Bathy-8 are methylotrophs that use lignin-derived methyl groups. Together with the potential capacity for methyl group utilization in many bathyarchaeotal MAGs (Seyler *et al.*, 2014; Yu *et al.,* 2018; this study), these results explain the observations of Lever *et al.* (2010), who found that porewater acetate in deep-subseafloor sediments was depleted in ^{13}C relative to sedimentary organic matter and postulated that a substantial fraction of the acetate produced in marine sediments might stem from reductive acetogenesis, fueled by microbial fermentation products, molecular hydrogen, and the methoxy groups of lignin monomers.

The utilization of the methoxy groups of lignin-derived aromatic compounds is a common trait of many acetogenic bacteria (Schink *et al.*, 1992, Drake, 1994). Methoxylated aromatic compounds are demethylated by the hindgut microbiota of termites (Brune *et al.*, 1995), but the organisms responsible for this activity have not been identified. It is tempting to speculate that termite gut *Bathyarchaeia* are organoheterotrophic (TB1) or lithoheterotrophic (TB2) acetogens that utilize methylated compounds such as lignin derivatives as methyl group donors and reduce CO_2 either with molecular hydrogen or with reducing equivalents derived from the oxidation of organic substrates.

It has been speculated that acetogenic archaea might have an energetic advantage over acetogenic bacteria, as they do not have to invest ATP to activate formate (He *et al.* 2016). However, the net synthesis of ATP is limited by the free-energy change of an acetogenic metabolism, which is independent of its reaction path and requires part of the ATP gained by substrate-level phosphorylation to be reinvested (e.g., for ferredoxin reduction; see above). Rather, it is feasible that the capacity for methylotrophic acetogenesis, which is less sensitive to low hydrogen partial pressures than hydrogenotrophic acetogenesis, provides an energetic advantage, analogous to the situation in methyl-reducing methanogens (Feldwert *et al.*, 2020). Moreover, it has been argued that long generation times contribute to the difficulties surrounding the enrichment and isolation of *Bathyarchaeia* in the laboratory (Yu *et al.* 2018). In view of the relatively short residence time of organic matter in termite guts (24 – 48 h; Kovoor, 1967; Bignell *et al.*, 1980), the growth rates of termite gut *Bathyarchaeia* must be high enough to avoid washout – unless they are attached to the intestinal surface.

2.8 Taxonomy

***Candidatus* Termiticorpusculum**
Etymology: L. n. *termes -itis*, a worm that eats wood, a termite; ; L. neut. n. *corpusculum*, a little body, a particle; N.L. neut. n. *Termiticorpusculum*, a little body associated with termites
Uncultured. Unclassified genus-level lineage in the Bathy-6 subgroup of *Bathyarchaeia* (Fig. 1; TB1 lineage). Comprises the phylotypes 1–7 (Table 1).
Habitat: The hindgut of higher termites

***Candidatus* Termitimicrobium**
Etymology: L. n. *termes -itis*, a worm that eats wood, a termite; N.L. neut. n. *microbium*, microbe; from Gr. masc. adj. *mikros*, small; from Gr. masc. n. *bios*, life; N.L. neut. n. *Termitimicrobium*, small life(-form) associated with termites
Uncultured. Unclassified genus-level lineage in the Bathy-6 subgroup of *Bathyarchaeia* (Fig. 1; TB2 lineage). Comprises the phylotypes 8–9 (Table 1).
Habitat: The hindgut of higher termites

3 Conclusions

To date, the non-methanogenic archaea in termite guts and their potential role in symbiotic digestion have received little attention. Our study provides strong evidence that termite gut *Bathyarchaeia* and other members of the Bathy-6 subgroup are archaeal acetogens: they possess the genomic potential to conserve energy by the production of acetyl-CoA from CO_2 (*Ca.* Termitimicrobium; TB2) and/or possibly methyl groups (almost all members of Bathy-6, including *Ca.* Termiticorpusculum; TB1). As in bacterial acetogens, their energy metabolism is likely mixotrophic. We identified a complete gene set encoding a novel Fpo-like 11-subunit hydrogenase, which closes the evolutionary gap between the ancestral [NiFe] hydrogenases and the respiratory complex I and would enable members of TB2 to grow lithotrophically on H_2. All members of Bathy-6 probably derive reducing equivalents from the oxidation of organic substrates (*viz*., amino acids) and might use reductive acetogenesis as an electron sink.

These findings agree with previous claims concerning the capacity for reductive acetogenesis in other subgroups of *Bathyarchaeia*. However, this is the first time that all genes encoding the Wood–Ljungdahl pathway and the components required for the provision of reducing equivalents and

energy conservation are conclusively documented. Although eight of the nine closely related phylotypes of termite gut *Bathyarchaeia* were represented by high-quality MAGs, a complete pathway was detected only in members of TB2 and two more basal lineages from other environments. This underscores the long-standing caution that the mere presence of marker genes of the Wood–Ljungdahl pathway do not qualify an organism as an acetogen since many of its enzymes are are found also in non-acetogenic organisms, where they are involved in the assimilation and interconversion of C_1 metabolites (Drake, 1994).

4 Experimental procedures

4.1 Metagenome-assembled genomes (MAGs)

Data on the MAGs from termite guts are from Hervé *et al.* (2020). All other MAGs were retrieved from the NCBI Assembly database (https://www.ncbi.nlm.nih.gov); accession numbers are listed in Table 1. Assembly coverage was determined as described by Hervé *et al.* (2020). Average nucleotide acid identities (ANI) were calculated with fastANI (Jain *et al.*, 2018). Protein-coding genes were predicted with Prodigal v2.6.3 (Hyatt *et al.*, 2010).

4.2 Genome phylogeny

A concatenated gene tree of bathyarchaeotal MAGs was constructed using the deduced amino acid sequences of 43 marker genes extracted with CheckM v1.0.8 (Parks *et al.*, 2015). The sequences were aligned using MAFFT v7.305b with the FFT-NS-2 method, and the resulting alignment was filtered using trimAL v1.2 with the gappyout method (Capella-Gutiérrez *et al.*, 2009, Katoh and Standley, 2013). Tree topology was inferred with IQ-TREE (multicore v1.6.11; Nguyen *et al.*, 2015) using the best-fit evolutionary model suggested by ModelFinder under the Bayesian Information Criterion (Kalyaanamoorthy *et al.*, 2017); node support was assessed using the Shimodaira–Hasegawa approximate-likelihood-ratio test (SH-aLRT) with 1,000 resamplings (Lemoine *et al.*, 2018).

Taxonomic classification was done with the GTDB-tk version 0.3.2 using the Genome Taxonomy Database (GTDB) release 04-RS89 (https://gtdb.ecogenomic.org/; Chaumeil *et al.*, 2018).

4.3 16S rRNA gene phylogeny

SSU rRNA gene sequences in the MAGs and other bathyarchaeotal bins obtained from the original metagenomes (Hervé *et al.*, 2020) were identified using the *ssu_finder* function implemented in CheckM. Sequences were imported into the alignment of rRNA gene sequences in the SILVA SSURef NR database release 132 (https://www.arb-silva.de; Quast et al., 2013) using Arb v6.0.6 (Ludwig *et al.*, 2004). After automatic alignment of the imported sequences using the *PT server* and the *Fast Aligner* tool implemented in Arb, the alignment was manually refined using the Arb editor, considering secondary structure information to identify homologous base positions. After removing sites with more than 50% gaps, the alignment consisted of 1,424 sites with unambiguously aligned base positions. Phylogenetic trees were reconstructed by maximum-likelihood analysis with IQ-TREE using the best-fit evolutionary model (GTR+F+R4) suggested by ModelFinder; node support was assessed using SH-aLRT with 1,000 resamplings. Gene fragments (<1,300 bp) were inserted into the core tree using the *parsimony* tool implemented in Arb.

4.4 Gene discovery and annotation

For an initial exploration of the genes potentially involved in energy metabolism, bathyarchaeotal MAGs were analyzed using the annotation provided in the IMG/Mer database (https://img.jgi.doe.gov/mer/; Chen *et al.*, 2019). Annotation results were verified, and missing functions were identified with Hidden Markov Model (HMM) searches, using HMMER v3.1b2 (Eddy, 2011) with a threshold E-value of 1E–5; the respective models are listed in Table S3. The identity of all genes of interest was confirmed using the NCBI Conserved Domain search (Marchler-Bauer and Bryant, 2004) and BLASTp (Altschul *et al.*, 1990). Additionally, Bathy-6-S, and Bathy-6-B were annotated with BlastKOALA (Kanehisa *et al.*, 2016). When indicated, closest neighbors were identified by BLAST, aligned using MAFFT v7.305b with the L-INS-i method (Katoh and Standley, 2013). Phylogenetic trees were reconstructed by maximum-likelihood analysis with IQ-TREE (Nguyen et al. 2015) using the best-fit evolutionary model (LG+G+I) suggested by ModelFinder (Kalyaanamoorthy et al., 2017). Node support was assessed using SH-aLRT with 1,000 resamplings (Lemoine et al. 2018).

4.5 Analysis of [NiFe] hydrogenases

Putative [NiFe] hydrogenase genes were identified by HMM searches (see above), using the highly resolved models provided by Anantharaman *et al.* (2016). Search results were confirmed with HydDB, a web-based tool for hydrogenase classification and analysis (https://services.birc.au.dk/hyddb/; Søndergaard *et al.*, 2016).

The deduced amino acid sequences of the large subunit (LSU) of [NiFe] hydrogenases recovered from the MAGs and their top BLAST hits on the IMG/Mer database were imported into an alignment of NuoD and FpoD homologs (Lang *et al.*, 2015), which was completed with representative members of other hydrogenase classes extracted from HydDB. The alignment was manually refined in the Arb editor. Phylogenetic trees were reconstructed by maximum-likelihood analysis with IQ-TREE (Nguyen et al. 2015) using the best-fit evolutionary model (LG+G+I) suggested by ModelFinder (Kalyaanamoorthy et al., 2017). Node support was assessed using SH-aLRT with 1,000 resamplings (Lemoine et al. 2018).

5 Author Contributions

HQL and AB designed the study. HQL analysed data and wrote the first draft of the manuscript. VH contributed to the analyses. AB analysed data and revised the manuscript. All authors edited and approved the final version of the manuscript.

6 Acknowledgments

This study was funded by the Deutsche Forschungsgemeinschaft (DFG) in the Collaborative Research Center SFB 987 and by the Max Planck Society. HQL was supported by a doctoral fellowship of the International Max Planck Research School for Environmental, Cellular and Molecular Microbiology (IMPRS-Mic), Marburg, Germany.

The authors thank the Joint Genome Institute for their metagenome sequencing service and for providing the IMG/ER platform.

7 References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990). Basic local alignment search tool. J Mol Bio 215: 403-410.

Anantharaman, K., Brown, C.T., Hug, L.A., Sharon, I., Castelle, C.J., Probst, A.J. *et al.* (2016). Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. Nat Comm 7: 13219.

Arshad, A., Speth, D.R., de Graaf, R.M., Op den Camp, H.J.M., Jetten, M.S.M. and Welte, C.U. (2015). A metagenomics-based metabolic model of nitrate-dependent anaerobic oxidation of methane by *Methanoperedens*-like archaea. Front Microbiol 6: 1423.

Basen, M. and Müller, V. (2017). "Hot" acetogenesis. Extremophiles 21: 15–26.

Bignell, D.E., Oskarsson, H. and Anderson, J.M. (1980). Distribution and abundance of bacteria in the gut of a soil-feeding termite *Procubitermes aburiensis* (Termitidae, Termitinae). J Gen Microbiol 117:393–403

Borrel, G., Parisot, N., Harris, H.M., Peyretaillade, E., Gaci, N., Tottey, W. *et al.* (2014). Comparative genomics highlights the unique biology of Methanomassiliicoccales, a Thermoplasmatales-related seventh order of methanogenic archaea that encodes pyrrolysine. BMC Genomics 15:679.

Bowers, R.M., Kyrpides, N.C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T.B.K. *et al.* (2017). Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. Nat Biotech 35: 725–731.

Brauman, A.; Kane, M.D.; Labat, M.; Breznak, J.A. (1992). Genesis of acetate and methane by gut bacteria of nutritionally diverse termites. Science 257: 1384-1387.

Brauman, A., Dore, J., Eggleton, P., Bignell, D., Breznak, J.A. and Kane, M.D. (2001). Molecular phylogenetic profiling of prokaryotic communities in guts of termites with different feeding habits. FEMS Microbiol Ecol 35: 27–36.

Breznak, J.A. and Switzer, J.M. (1986). Acetate synthesis from H_2 plus CO_2 by termite gut microbes. Appl Environ Microbiol 52: 623-630.

Brune, A. (2014). Symbiotic digestion of lignocellulose in termite guts. Nat Rev Microbiol 12: 168–180.

Brune, A. (2018). Methanogens in the digestive tract of termites. In: (Endo)symbiotic methanogenic archaea. J.H.P. Hackstein (ed). Cham: Springer International Publishing, 81-101.

538 Brune, A. (2019) Methanogenesis in the digestive tracts of insects and other arthropods. In:
539 *Biogenesis of Hydrocarbons (Book series: Handbook of Hydrocarbon and Lipid*
540 *Microbiology)* (Stams, A.J.M., Sousa, D., eds.). Springer, Cham, pp. 229–260.

541 Brune, A. and Dietrich, C. (2015). The gut microbiota of termites: Digesting the diversity in the light
542 of ecology and evolution. Annu Rev Microbiol 69: 145-166.

543 Brune, A., Miambi, E., Breznak, J.A. (1995). Roles of oxygen and the intestinal microflora in the
544 metabolism of lignin-derived phenylpropanoids and other monoaromatic compounds by
545 termites. Appl Environ Microbiol 61:7 2688-2695.

546 Bucek, A., Šobotník, J., He, S., Shi, M., McMahon, D.P., Holmes, E.C. *et al.* (2019). Evolution of
547 termite symbiosis informed by transcriptome-based phylogenies Curr Biol 29:21.

548 Buckel, W. and Thauer, R.K. (2013) Energy conservation via electron bifurcating ferredoxin
549 reduction and proton/Na^+ translocating ferredoxin oxidation. Biochim Biophys Acta 1827:
550 94-113.

551 Capella-Gutiérrez, S., Silla-Martínez, J.M. and Gabaldón, T. (2009). trimAl: a tool for automated
552 alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25: 1972-1973.

553 Chaumeil, P. A., Mussig, A. J., Hugenholtz, P. and Parks, D. H. (2018). GTDB-Tk: a toolkit to
554 classify genomes with the Genome Taxonomy Database. Bioinformatics btz848.

555 Chen, I.A., K. Chu, K. Palaniappan, M. Pillay, A. Ratner, J. Huang, M., *et al.* (2019) IMG/M v.5.0:
556 an integrated data management and comparative analysis system for microbial genomes and
557 microbiomes. Nucleic Acids Res 47: D666-d677.

558 Collins, G., O'Connor, L., Mahony, T., Gieseke, A., de Beer, D. and O'Flaherty, V. (2005).
559 Distribution, localization, and phylogeny of abundant populations of *Crenarchaeota* in
560 anaerobic granular sludge. Appl Environ Microbiol 71: 7523-7527.

561 Dombrowski, N., Seitz, K.W., Teske, A.P. and Baker, B.J. (2017). Genomic insights into potential
562 interdependencies in microbial hydrocarbon and nutrient cycling in hydrothermal sediments.
563 Microbiome 5, 106 (2017). https://doi.org/10.1186/s40168-017-0322-2

564 Drake, H.L. (1994). Acetogenesis, acetogenic bacteria, and the acetyl-CoA "Wood/Ljungdahl"
565 Pathway: Past and current perspectives. In: Drake, H.L. (eds) Acetogenesis. Chapman & Hall
566 Microbiology Series (Physiology / Ecology / Molecular Biology / Biotechnology). Springer,
567 Boston, MA

568 Eckert, C., Boehm, M., Carrieri, D., Yu, J., Dubini, A., Nixon, P.J. and Maness, P.-C. (2012) Genetic
569 analysis of the Hox hydrogenase in the cyanobacterium *Synechocystis sp.* PCC 6803 reveals
570 subunit roles in association, assembly, maturation, and function. J Biol Chem 287: 43502-
571 43515.

572 Eddy, S.R. (2011). Accelerated profile HMM searches. PLoS Comput Biol 7: e1002195.

573 Evans, P.N., Parks, D.H., Chadwick, G.L., Robbins, S.J., Orphan, V.J., Golding, S.D. *et al.* (2015).
574 Methane metabolism in the archaeal phylum Bathyarchaeota revealed by genome-centric
575 metagenomics. Science 350: 434-438.

576 Farag, I.F., Biddle, J.F., Zhao, R., Martino, A.J., House, C.H. and León-Zayas, R.I. (2020). Metabolic
577 potentials of archaeal lineages resolved from metagenomes of deep Costa Rica sediments.
578 ISME J 14: 1345–1358.

579 Feldewert, C., Lang, K. and Brune, A. (2020). The hydrogen threshold of obligately methyl-reducing
580 methanogens. FEMS Microbiol. Lett. 367: fnaa137.

581 Feng, X, Wang, Y. Zubin, R. and Wang, F. (2019). Core Metabolic Features and Hot Origin of
582 Bathyarchaeota. Engineering 5: 498-504.

583 Fillol, M., Auguet, J., Casamayor, E.O. and Borrego, C.M. (2016). Insights in the ecology and
584 evolutionary history of the Miscellaneous Crenarchaeotic Group lineage. ISME J 10: 665–
585 677.

586 Friedrich, M.W., Schmitt-Wagner, D., Lueders, T. and Brune, A. (2001). Axial differences in
587 community structure of Crenarchaeota and Euryarchaeota in the highly compartmentalized
588 gut of the soil-feeding termite *Cubitermes orthognathus*. Appl Environ Microbiol 67: 4880-
589 4890.

590 Friedrich, T. and Scheide, D. (2000). The respiratory complex I of bacteria, archaea and eukarya and
591 its module common with membrane-bound multisubunit hydrogenases. FEBS Lett 479: 1-5.

592 Galagan, J.E., Nusbaum, C., Roy, A., Endrizzi, M.G., Macdonald, P., FitzHugh, W., et al. (2002).
593 The genome of *M. acetivorans* reveals extensive metabolic and physiological diversity.
594 Genome Res. 12: 532-542. doi:10.1101/gr.223902

595 Greening, C., Biswas, A., Carere, C., Jackson, C.J., Taylor, M.C., Stott, M.B., Cook, G.M. and
596 Morales, S.E. (2016). Genomic and metagenomic surveys of hydrogenase distribution
597 indicate H2 is a widely utilised energy source for microbial growth and survival. ISME J 10:
598 761–777.

599 Godon, J.-J., Zumstein, E., Dabert, P., Habouzit, F. and Moletta, R. (1997). Molecular microbial
600 diversity of an anaerobic digestor as determined by small-subunit rDNA sequence analysis.
601 Appl Environ Microbiol 63: 2802-2813

602 Grieco, M.B., Lopes, F.A.C., Oliveira, L.S., Tschoeke, D.A., Popov, C.C., Thompson, C.C. *et al.*
603 (2019). Metagenomic analysis of the whole gut microbiota in Brazilian Termitidae termites
604 *Cornitermes cumulans, Cyrilliotermes strictinasus, Syntermes dirus, Nasutitermes jaraguae,*
605 *Nasutitermes aquilinus, Grigiotermes bequaerti, and Orthognathotermes mirim*. Curr
606 Microbiol 76: 687-697.

607 Gutekunst, K., Chen, X., Schreiber, K., Kaspar, U., Makam, S., andAppel, J. (2014). The
608 bidirectional NiFe-hydrogenase in *Synechocystis sp.* PCC 6803 is reduced by flavodoxin and
609 ferredoxin and is essential under mixotrophic, nitrate-limiting conditions. J Biol Chem 289:
610 1930-1937.

611 Harris, R.L., Lau, M.C.Y., Cadar, A., Bartlett, D.H., Cason, E., van Heerden, E. and Onstott, T.C.
612 (2018). Draft genome sequence of "*Candidatus* Bathyarchaeota" archaeon BE326-BA-RLH,
613 an uncultured denitrifier and putative anaerobic methanotroph from South Africa's deep
614 continental biosphere. Microbiol Resour Announc 7: e01295-01218.

615 He, Y., Li, M., Perumal, V., Feng, X., Fang, J., Xie, J., Sievert, S.M. and Wang F. (2016). Genomic
616 and enzymatic evidence for acetogenesis among multiple lineages of the archaeal phylum
617 Bathyarchaeota widespread in marine sediments. Nat Microbiol 1: 16035.

618 Heider, J., Mai, X., and Adams M.W. (1996). Characterization of 2-ketoisovalerate ferredoxin
619 oxidoreductase, a new and reversible coenzyme A-dependent enzyme involved in peptide
620 fermentation by hyperthermophilic archaea. J Bacteriol 178: 780-787.

621 Herlemann, D.P.R., Geissinger, O., Ikeda-Ohtsubo, W., Kunin, V., Sun, H., Lapidus, A., Hugenholtz,
622 P. and Brune, A. (2009). Genomic analysis of "*Elusimicrobium minutum*," the first cultivated
623 representative of the phylum "*Elusimicrobia*" (formerly termite group 1). Appl Environ
624 Microbiol 75: 2841-2849.

625 Hervé, V., Liu, P., Dietrich, C., Sillam-Dussès, D., Stiblik, P., Šobotník, J. *et al.* (2020).
626 Phylogenomic analysis of 589 metagenome-assembled genomes encompassing all major
627 prokaryotic lineages from the gut of higher termites. PeerJ 8: e8614.

628 Hippler, B. and Thauer, R.K. (1999). The energy conserving
629 methyltetrahydromethanopterin:coenzyme M methyltransferase complex from methanogenic
630 archaea: function of the subunit MtrH. FEBS Lett. 449: 165-168.

631 Hochheimer, A., Linder, D., Thauer, R.K. and Hedderich, R. (1996). The molybdenum
632 formylmethanofuran dehydrogenase operon and the tungsten formylmethanofuran
633 dehydrogenase operon from *Methanobacterium thermoautotrophicum*. Structures and
634 transcriptional regulation. Eur J Biochem 242: 156-162.

635 Hyatt, D., Chen, G., Locascio, P.F., Land, M.L., Larimer, F.W. and Hauser, L.J. (2010). Prodigal:
636 prokaryotic gene recognition and translation initiation site identification. BMC Bioinfo 11:
637 119-119.

638 Ikeda-Ohtsubo, W., Strassert, J.F.H., Köhler, T., Mikaelyan, A., Gregor, I., McHardy, A.C., Tringe,
639 S.G., Hugenholtz, P., Radek, R. and Brune, A. (2016). '*Candidatus* Adiutrix intracellularis',
640 an endosymbiont of termite gut flagellates, is the first representative of a deep-branching
641 clade of *Deltaproteobacteria* and a putative homoacetogen. Environ. Microbiol. 18: 2548-
642 2564

643 Inagaki, F., Suzuki, M., Takai, K., Oida, H., Sakamoto, T., Aoki, K. *et al.* (2003). Microbial
644 communities associated with geological horizons in coastal subseafloor sediments from the
645 sea of Okhotsk. Appl Environ Microbiol 69: 7224-7235.

646 Jain, C., Rodriguez-R, L.M., Phillippy, A.M., Konstantinidis, K.T. and Aluru, S. (2018). High
647 throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat
648 Comm 9:5114.

649 Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A. and Jermiin, L.S. (2017).
650 ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods 14: 587-
651 589.

652 Kanehisa, M., Sato, Y. and Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG tools for
653 functional characterization of genome and metagenome sequences. J Mol Biol 428: 726-731.

654 Kaster, A., Moll, J., Parey, K. and Thauer, R.K. (2011). Coupling of ferredoxin and heterodisulfide
655 reduction via electron bifurcation in hydrogenotrophic methanogenic archaea Proc Natl Acad
656 Sci 108: 2981-2986.

657 Katoh, K. and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7:
658 improvements in performance and usability. Mol Biol Evol 30: 772-780.

659 Kengen, S.W.M. and Stams, A.J.M. (1993). Formation of L-alanine as a reduced end product in
660 carbohydrate fermentation by the hyperthermophilic archaeon *Pyrococcus furiosus*. Arch
661 Microbiol 161: 168-175.

662 Kovoor, J. (1967). Etude radiographique du transit intestinal chez un Termite supérieur. Experientia
663 23, 820–821 https://doi.org/10.1007/BF02146863

664 Kreft, J.-U. and Schink, B. (1994). O-demethylation by the homoacetogenic anaerobe *Holophaga*
665 *foetida* studied by a new photometric methylation assay using electrochemically produced
666 cob(I)alamin. Eur. J. Biochem. 226: 945-951

667 Kremp, F., Poehlein, A., Daniel, R. and Müller, V. (2018). Methanol metabolism in the acetogenic
668 bacterium *Acetobacterium woodii*. Environ. Microbiol. 20: 4369-4384

669 Kröninger, L., Berger, S., Welte, C. and Deppenmeier, U. (2016). Evidence for the involvement of
670 two heterodisulfide reductases in the energy-conserving system of *Methanomassiliicoccus*
671 *luminyensis*. FEBS J 283: 472-483.

672 Kubo, K., Lloyd, K.G., Biddle, J.F., Amann, R., Teske, A. and Knittel, K. (2012). Archaea of the
673 Miscellaneous Crenarchaeotal Group are abundant, diverse and widespread in marine
674 sediments. ISME J 6: 1949-1965.

675 Künkel, A., Vorholt, J.A., Thauer, R.K. and Hedderich, R. (1998). An *Escherichia coli* hydrogenase-
676 3-type hydrogenase in methanogenic archaea. Eur J Biochem 252: 467-476.

677 Lang, K., Schuldes, J., Klingl, A., Poehlein, A., Daniel, R., and Brune, A. (2015). New mode of
678 energy metabolism in the seventh order of methanogens as revealed by comparative genome
679 analysis of "*Candidatus* Methanoplasma termitum". Appl Environ Microbiol 81: 1338-1352.

680 Lazar, C.S., Baker, B.J., Seitz, K., Hyde, A.S., Dick, G.J., Hinrichs K. and Teske, A.P. (2016).
681 Genomic evidence for distinct carbon substrate preferences and ecological niches of
682 Bathyarchaeota in estuarine sediments. Environ Microbiol 18: 1200-1211.

683 Leadbetter, J.R., Schmidt, T.M., Graber, J.R. and Breznak, J.A. (1999) Acetogenesis from H_2 Plus
684 CO_2 by Spirochetes from Termite Guts. Science 283: 686-689.

Lee, M.J. and Zinder, S.H. (1988). Isolation and characterization of a thermophilic bacterium which oxidizes acetate in syntrophic association with a methanogen and which grows acetogenically on H_2-CO_2. Appl Environ Microbiol 54:1 124-129.

Lemoine F, Domelevo Entfellner JB, Wilkinson E, Correia D, Dávila Felipe M, De Oliveira T, Gascuel O. Renewing Felsenstein's phylogenetic bootstrap in the era of big data. Nature. 2018;556(7702):452–456. doi: 10.1038/s41586-018-0043-0.

Lever, M.A., Heuer, V.B., Morono, Y., Masui, N., Schmidt, F., Alperin, M.J., et al. (2010). Acetogenesis in deep subseafloor sediments of the Juan de Fuca ridge flank: a synthesis of geochemical, thermodynamic, and gene-based evidence. Geomicrobiol J 27: 183-211.

Ludwig, W., O. Strunk, R. Westram, L. Richter, H. Meier, Yadhukumar *et al.* (2004). ARB: a software environment for sequence data. Nucleic Acids Res 32: 1363-1371.

Mai, X, and Adams, M.W. (1996). Purification and characterization of two reversible and ADP-dependent acetyl coenzyme A synthetases from the hyperthermophilic archaeon *Pyrococcus furiosus*. J Bacteriol 178: 5897-5903

Marchler-Bauer, A. and Bryant, S.H. (2004). CD-Search: protein domain annotations on the fly. Nucleic Acids Res 32: W327-W331.

Meng, J., Xu, J., Qin, D., He, Y., Xiao, X. and Wang, F. (2014). Genetic and functional properties of uncultivated MCG archaea assessed by metagenome and gene expression analyses. ISME J 8: 650-659.

Mikaelyan, A., Meuser, K. and Brune, A. (2016). Microenvironmental heterogeneity of gut compartments drives bacterial community structure in wood- and humus-feeding higher termites. FEMS Microbiol Ecol 93: fiw210.

Moparthi, V.K. and Hägerhäll, C. (2011). The evolution of respiratory chain complex I from a smaller last common ancestor consisting of 11 protein subunits. J Mol Evol 72:5-6 484-497.

Musfeldt, M., Selig, M. and Schönheit, P. (1999). Acetyl coenzyme A synthetase (ADP forming) from the hyperthermophilic archaeon *Pyrococcus furiosus*: identification, cloning, separate expression of the encoding genes, *acdAI* and *acdBI*, in *Escherichia coli*, and in vitro reconstitution of the active heterotetrameric enzyme from its recombinant subunits. J Bacteriol 181: 5885-5888.

Nelson, W. C., Tully, B. J., and Mobberley, J. M. (2020). Biases in genome reconstruction from metagenomic data. PeerJ, *8*, e10119

Nguyen, L., Schmidt, H.A., von Haeseler, A. and Minh, B.Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32: 268-274.

Ochsenreiter, T., Selezi, D., Quaiser, A., Bonch-Osmolovskaya, L. and Schleper, C. (2003). Diversity and abundance of Crenarchaeota in terrestrial habitats studied by 16S RNA surveys and real time PCR. Environ Microbiol 5: 787-797.

722 Ohkuma, M., Noda, S., Hattori, S., Iida, T., Yuki, M., Starns, D. *et al.* (2015). Acetogenesis from H_2
723 plus CO_2 and nitrogen fixation by an endosymbiotic spirochete of a termite-gut cellulolytic
724 protist. Proc Natl Acad Sci 112: 10224-10230.

725 Ottesen, E.A. and Leadbetter J.R. (2011). Formyltetrahydrofolate synthetase gene diversity in the
726 guts of higher termites with different diets and lifestyles. Appl Environ Microbiol 77: 3461-
727 3467.

728 Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P. and Tyson, G.W. (2015). CheckM:
729 assessing the quality of microbial genomes recovered from isolates, single cells, and
730 metagenomes. Genome Res 25: 1043-1055.

731 Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P. *et al.* (2013). The SILVA
732 ribosomal RNA gene database project: improved data processing and web-based tools.
733 Nucleic Acids Res 41: D590-D596.

734 Rinke, C, Chuvochina, M., Mussig, A.J., Chaumeil, P., Waite, D.W., Whitman,W.B, Parks, D.H.
735 and Hugenholtz, P. (2020). A rank-normalized archaeal taxonomy based on genome
736 phylogeny resolves widespread incomplete and uneven classifications. bioRxiv [Preprint].
737 Available at: https://doi.org/10.1101/2020.03.01.972265 (Accessed March 10, 2020).

738 Rosenthal, A.Z., Zhang, X., Lucey, K.S., Ottesen, E.A., Trivedi, V., Choi, H.M.T., Pierce, N.A. and
739 Leadbetter, J.R. (2013) Localizing transcripts to single cells suggests an important role of
740 uncultured deltaproteobacteria in the termite gut hydrogen economy. Proc. Natl. Acad. Sci.
741 USA 110: 16163-16168

742 Schleper, C., Holben W. and Klenk, H.P. (1997). Recovery of crenarchaeotal ribosomal DNA
743 sequences from freshwater-lake sediments. Appl Environ Microbiol 63: 321-323.

744 Schnürer, A., Svensson, B.H. and Schink, B. (1997). Enzyme activities in and energetics of acetate
745 metabolism by the mesophilic syntrophically acetate-oxidizing anaerobe *Clostridium*
746 *ultunense*. FEMS Microbiol. Lett. 154: 331-336.

747 Shi, Y., Huang, Z., Han, S., Fan, S. and Yang, H. (2015). Phylogenetic diversity of Archaea in the
748 intestinal tract of termites from different lineages. J Basic Microb 55: 1021-1028.

749 Schink, B., Brune, A. and Schnell, S. (1992). Anaerobic degradation of aromatic compounds. In:
750 Microbial Degradation of Natural Products (G. Winkelmann, ed.). VCH Verlagsgesellschaft,
751 Weinheim, s219-242.

752 Schink, B. (1994). Diversity, ecology, and isolation of acetogenic bacteria. In: Drake H.L. (eds)
753 Acetogenesis. Chapman & Hall Microbiology Series (Physiology / Ecology / Molecular
754 Biology / Biotechnology). Springer, Boston, MA

755 Schoelmerich, M.C. and Müller, V. (2019). Energy-converting hydrogenases: the link between H_2
756 metabolism and energy conservation. Cell Mol Life Sci 77: 1461-1481.

757 Schuchmann, K. and Müller, V. (2014) Autotrophy at the thermodynamic limit of life: a model for
758 energy conservation in acetogenic bacteria. Nat Rev Microbiol 12: 809-821.

759 Seyler, L.M., McGuinness, L.M., Kerkhof, L.J. (2014). Crenarchaeal heterotrophy in salt marsh
760 sediments. ISME J 8: 1534-1543.

761 Søndergaard, D., Pedersen, C.N.S. and Greening, C. (2016). HydDB: A web tool for hydrogenase
762 classification and analysis. Sci Rep 6: 34212.

763 Tersteegen, A. and Hedderich, R. (1999). *Methanobacterium thermoautotrophicum* encodes two
764 multisubunit membrane-bound [NiFe] hydrogenases. Eur J Biochem 264: 930-943.

765 Thauer, R.K., Kaster, A.-K., Seedorf, H., Buckel, W. and Hedderich, R. (2008) Methanogenic
766 archaea: ecologically relevant differences in energy conservation. Nat Rev Microbiol 6: 579-
767 591.

768 Tholen, A. and Brune A. (1999). Localization and in situ activities of homoacetogenic bacteria in the
769 highly compartmentalized hindgut of soil-feeding higher termites (*Cubitermes* spp.). Appl
770 Environ Microbiol 65: 4497-4505.

771 van der Meijden, P., Heythuysen, H.J., Pouwels, A. et al. (1983) Methyltransferases involved in
772 methanol conversion by *Methanosarcina barkeri*. Arch. Microbiol. 134: 238-242

773 Vorholt, J.A., Vaupel, M. and Thauer, R.K. (1996). A polyferredoxin with eight [4Fe–4S] clusters as
774 a subunit of molybdenum formylmethanofuran dehydrogenase from *Methanosarcina barkeri*.
775 Eur J Biochem 236: 309-317.

776 Welte, C. and Deppenmeier U. (2011). Membrane-bound electron transport in *Methanosaeta*
777 *thermophila*. J Bacteriol 193: 2868-2870.

778 Westphal, L., Wiechmann, A., Baker, J., Minton, N.P. and Müller, V. (2018). The Rnf complex is an
779 energy-coupled transhydrogenase essential to reversibly link cellular NADH and ferredoxin
780 pools in the acetogen *Acetobacterium woodii*. J Bacteriol 200: e00357-18.

781 Xiang, X., Wang, R., Wang, H., Gong, L., Man, B. and Xu, Y. (2017). Distribution of
782 Bathyarchaeota communities across different terrestrial settings and their potential ecological
783 functions. Sci Rep 7: 45028.

784 Yu, T., Wu, W., Liang, W., Lever, M.A., Hinrichs, K. and Wang F. (2018). Growth of sedimentary
785 Bathyarchaeota on lignin as an energy source. Proc Natl Acad Sci 115: 6022-6027.

786 Zhou, Z., Pan, J., Wang, F., Gu, J. and Li, M. (2018). Bathyarchaeota: globally distributed metabolic
787 generalists in anoxic environments. FEMS Microbiol Rev 42: 639-655.

788

789 8 Data Availability Statement

790 The MAGs can be accessed at the NCBI GenBank database (Table 1) and through the IMG platform
791 (https://ncbi.nlm.nih.gov/genbank/ and https://img.jgi.doe.gov/ respectively). Genome IDs and
792 accession numbers for the NCBI SRA database are given in Table S1.

9 Tables

Table 1. Characteristics of the MAGs of *Bathyarchaeia* from termite guts and other members of Bathy-6 included in the analyses.

Phylo-type[a]	MAG[b]	Compartment	Relative abundance (%)[c]	Complete-ness (%)[d]	Contami-nation (%)[d]	Assembly size (bp)	Number of contigs	G+C content (mol%)	Coding density (%)	Predicted genes	Accession number[e]
1	Co191P1_bin46	P1	0.36	95.8	5.7	1762101	230	37.8	80.2	1772	WQRU00000000
	Co191P3_bin4	P3	0.09	99.1	4.2	1808297	159	37.8	79.6	1717	WQSY00000000
	Co191P4_bin18	P4	2.46	99.2	4.2	1994150	212	37.9	80.0	1899	WQTO00000000
2	Emb289P3_bin80	P3	0.13	96.3	6.3	2128005	163	39.0	82.5	2062	WQYG00000000
3	Lab288P3_bin115	P3	0.20	91.5	3.3	1167853	190	38.2	86.9	1242	WRCG00000000
	Lab288P4_bin25	P4	0.13	96.3	3.3	1375305	225	38.1	85.1	1455	WREZ00000000
4	Th196P4_bin19	P4	1.76	99.2	3.7	2287482	173	35.6	74. 9	2201	WRNB00000000
5	Cu122P1_bin20	P1	0.07	90.0	8.9	1504932	227	37.4	84.2	1628	WQTR00000000
6	Nc150P3_bin14	P3	0.02	63.8	2.3	656967	123	38.5	84.5	772	WRGI00000000
	Nc150P4_bin1	P4	0.28	98.1	4.7	1587817	173	38.9	82.6	1621	WRGM00000000
7	Nt197P4_bin22	P4	0.76	99.1	4.7	2179374	105	39.3	82.8	2153	WRJX00000000
8	Emb289P1_bin127	P1	0.08	99.1	1.9	2139595	140	43.4	82.9	2055	WQVG00000000
	Emb289P3_bin109	P3	0.23	96.3	2.8	2080780	121	43.4	83.1	2162	WQWQ00000000
9	Lab288P3_bin169	P3	1.20	98.6	2.8	2243011	107	43.3	83.3	2269	WRCX00000000
	Lab288P4_bin61	P4	0.52	99.1	3.7	2504117	128	43.0	83.2	2483	WRFL00000000
S	SZUA-568[f]	NA	NA	90.7	8.4	1641847	207	41.1	86.5	1810	QKIA00000000
B	Be326-BA-RLH[f]	NA	NA	89.8	3.7	2076091	227	44.9	86.1	2394	QYYE00000000
A	AD8-1[f]	NA	NA	95.8	4.2	1583813	83	32.4	84.5	1735	LFWW00000000

[a] Average nucleotide identity (ANI) ≥ 99%; for details, see Figure S1.

796 [b] The first letters of the MAG names indicate the host species (Co, *Cornitermes* sp.; Emb, *Embiratermes neotenicus*; Lab, *Labiotermes*
797 *labralis*; Th, *Termes hospes*; Cu, *Cubitermes ugandensis*; Nc, *Nasutitermes corniger*; Nt, *Neocapritermes taracua*).

798 [c] Relative abundance of the reads assigned to each MAG among the total number of reads in the corresponding metagenome (Hervé *et al.*,
799 2020).

800 [d] Completeness and contamination were estimated with CheckM using 104 single-copy marker genes (Parks *et al*., 2015). For detailed
801 results of the CheckM analysis, see Table S1.

802 [e] For NCBI Nucleotide database; IMG genome IDs are given in Table S1.

803 [f] Referred to as phylotypes Bathy-6-S (J. Pan and Z. Zhou, unpublished), Bathy-6-B (Harris *et al.*, 2018), and Bathy-6-A (Lazar *et al.*, 2016).

10 Figure legends

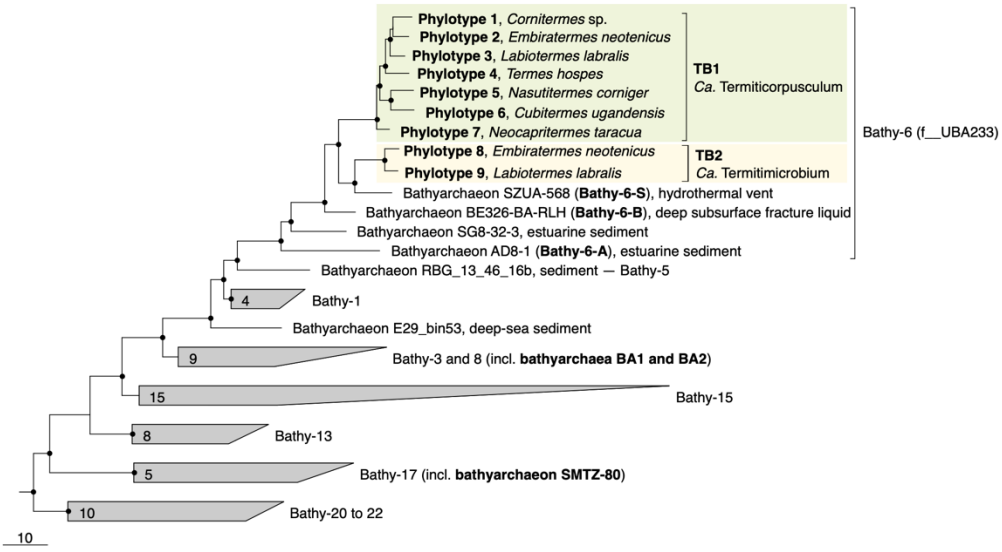


Figure 1. Genome-based phylogeny of termite gut *Bathyarchaeia*, illustrating the relationship of lineages TB1 and TB2 to other MAGs in the Bathy-6 subgroup (f__UBA233 in the GTDB taxonomy). MAGs of other subgroups that are mentioned in the text are marked in bold. The maximum-likelihood tree was inferred from a concatenated alignment of 43 marker genes using the LG+F+I+G4 model and rooted with selected Crenarchaeota and Euryarchaeota as outgroup. A fully expanded tree with the accession numbers for all genomes is shown in the Supplementary Material (Supplementary Figure S2). The scale bar indicates 10 amino acid substitutions per site. Highly supported nodes (SH-aLRT, ● ≥ 95%, 1,000 replications) are indicated.

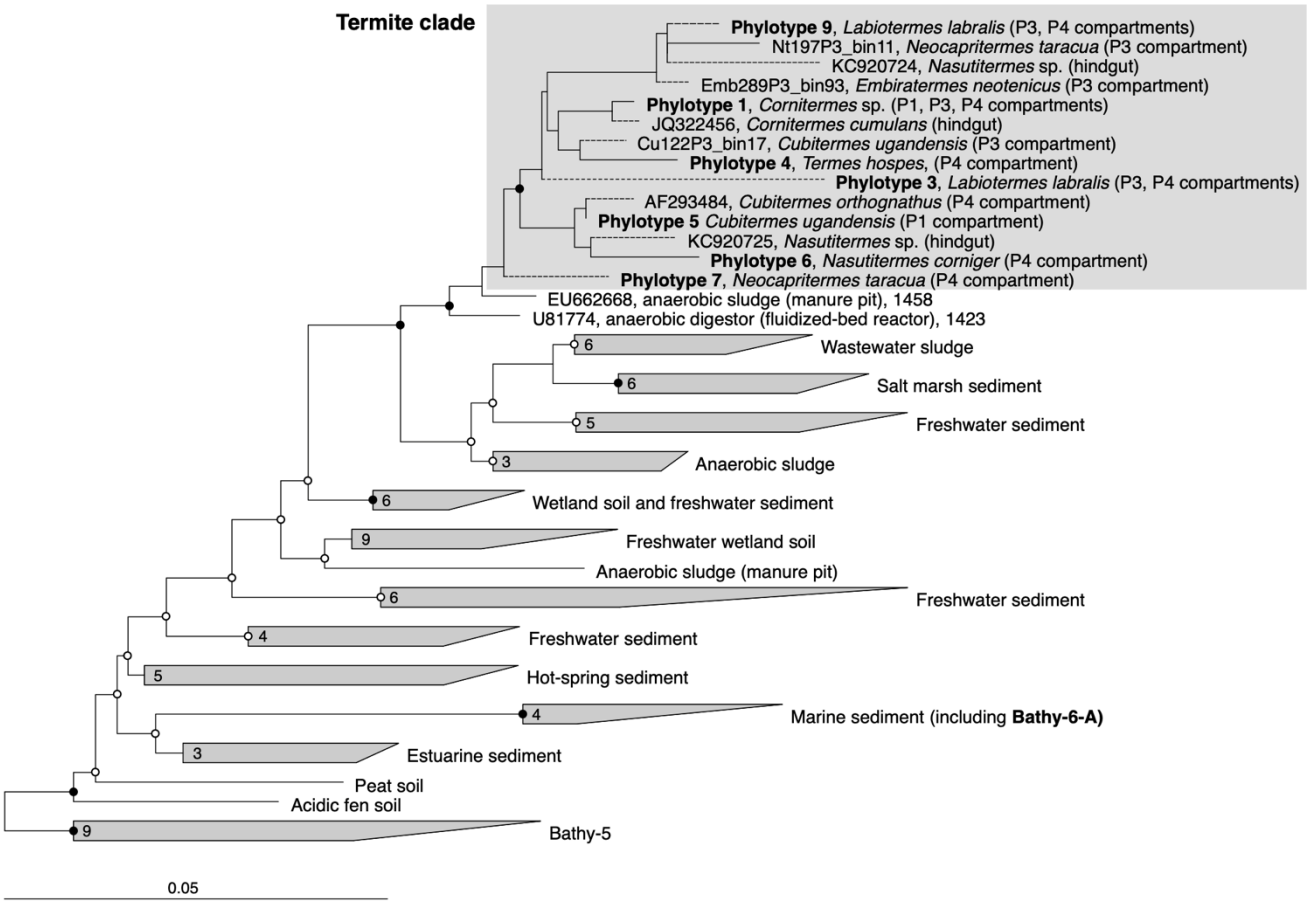


Figure 2. 16S rRNA-based phylogeny of subgroup Bathy-6, indicating the placement of the termite clade among *Bathyarchaeia* from other environments. The maximum-likelihood tree is based on a curated alignment (1,424 positions) of all sequences in the SILVA database and their homologs retrieved from the bathyarchaeal MAGs and the low-quality bins obtained from the termite gut metagenomes (Hervé *et al.*, 2020). The tree was rooted with members of Bathy-5 as outgroup. The scale bars indicate 0.05 nucleotide substitutions per site. SH-aLRT values (● ≥ 95%; ○ ≥ 80%, 1,000 replications) indicate node support. Branches marked with dashed lines indicate shorter sequences that were added using the parsimony tool. A fully expanded tree with the accession numbers of all sequences is shown in the Supplementary Material (Supplementary Figure S3).

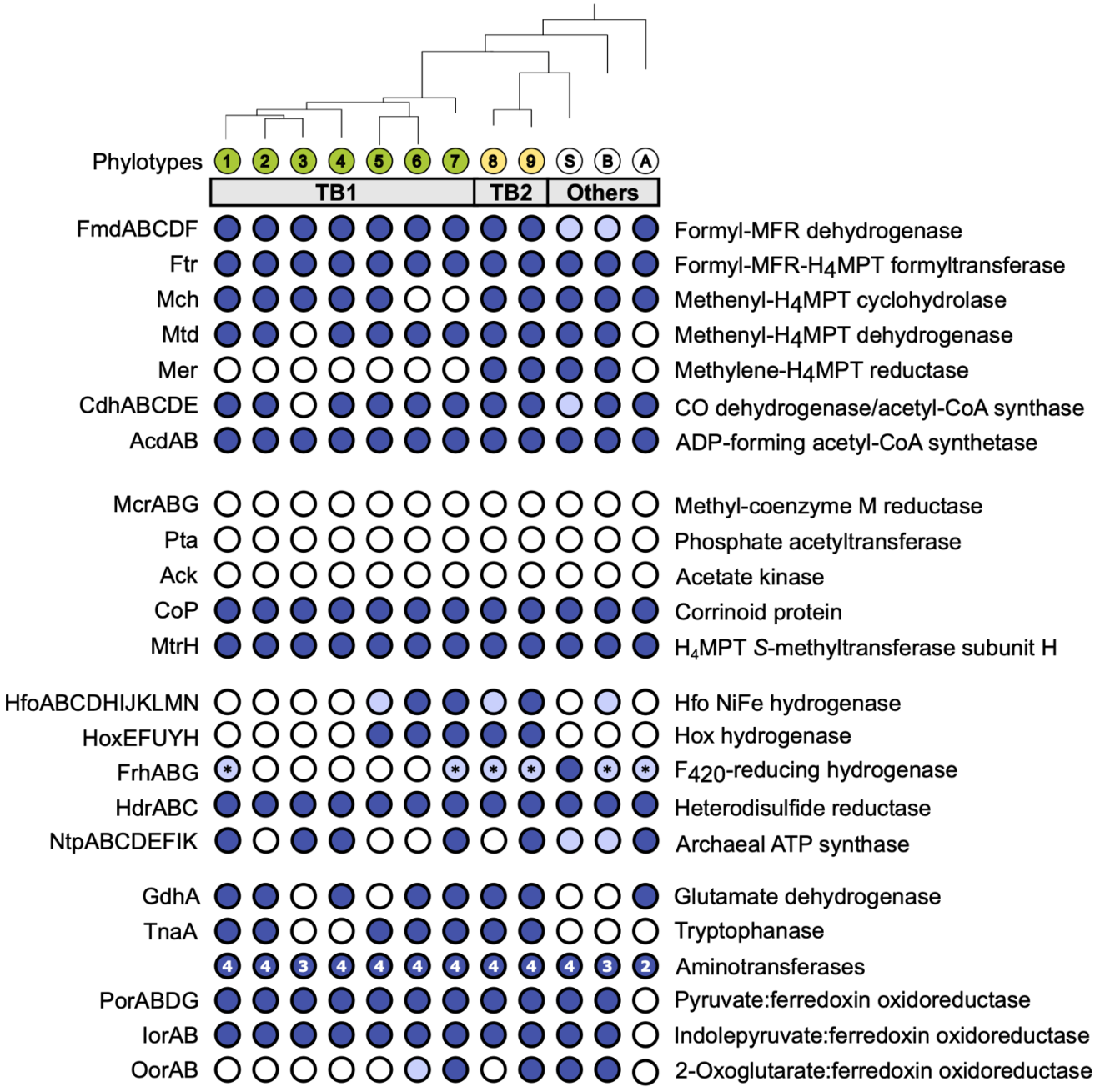
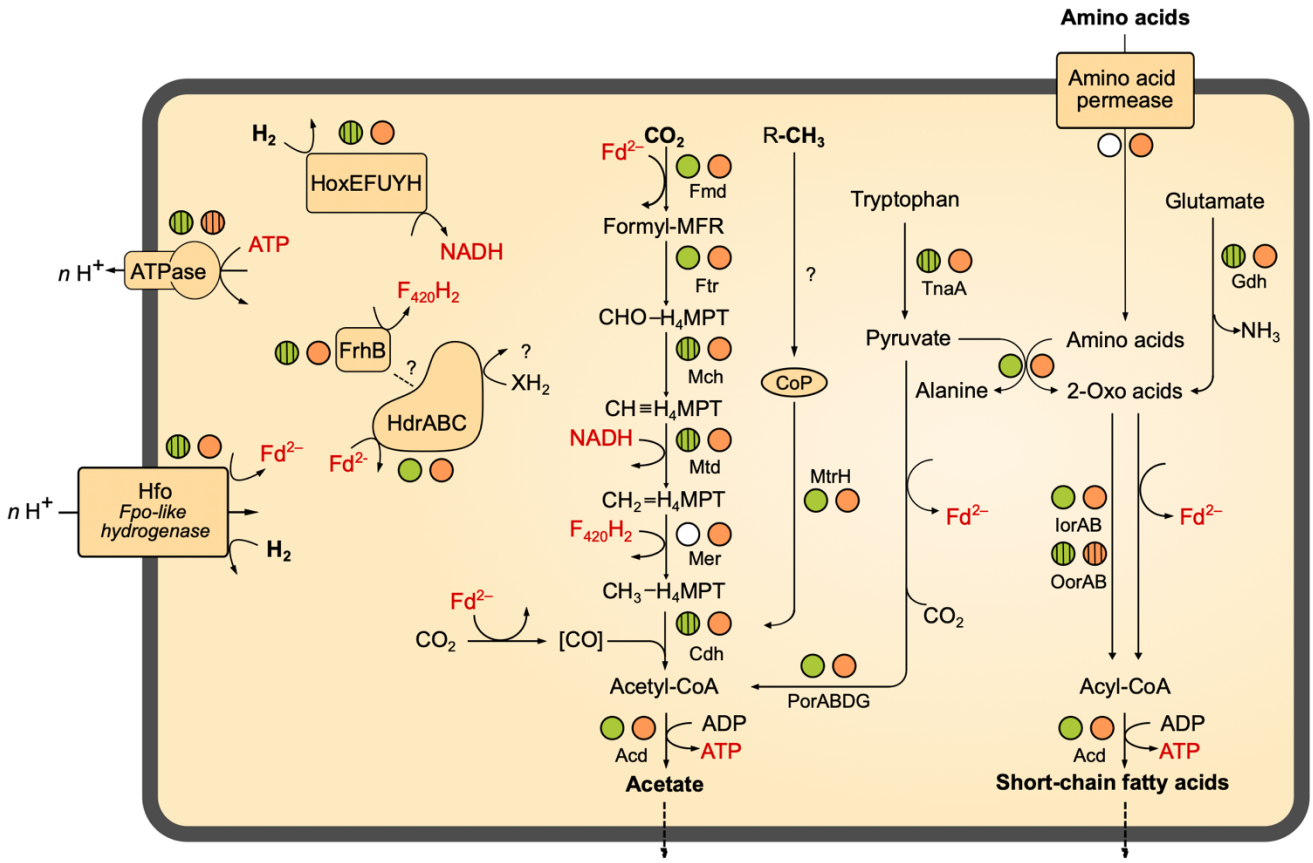


Figure 3. Gene functions encoded by termite gut bathyarchaea (TB1 and TB2) and other representatives of the Bathy-6 subgroup. All phylotypes with sufficiently complete genomes were included; their phylogenetic relationship was taken from Figure 1 (for strain designations, see Table 1). Colored circles indicate presence and open circles indicate absence of the respective function; light blue indicates that a gene set is incomplete. The number of aminotransferases encoded by each phylotype are indicated in the circle. The asterisk (*) in FrhABG indicates that only FrhB is present. If a phylotype is represented by more than one MAG, the annotation results were combined; details can be found in the Supplementary Material (Table S2). H_4MPT: tetrahydromethanopterin, MFR: methanofuran, Fpo: F_{420}:methanophenazine oxidoreductase.



835

836 **Figure 4.** Metabolic map of termite gut *Bathyarchaeia*. The circles next to each enzyme indicate the
837 presence of the coding genes in TB1 (green) and TB2 (orange), respectively. Striped circles indicate
838 that a function is not encoded by all genomes of the respective group; a white filling indicates
839 absence from all MAGs of the group (see also Figure 3). The directionality of Fpo-like hydrogenase
840 (Hfo) and ATP (synth)ase is explained in the text. A detailed list of genes present in the respective
841 MAGs is provided as Supplementary Material (Supplementary Table S2).

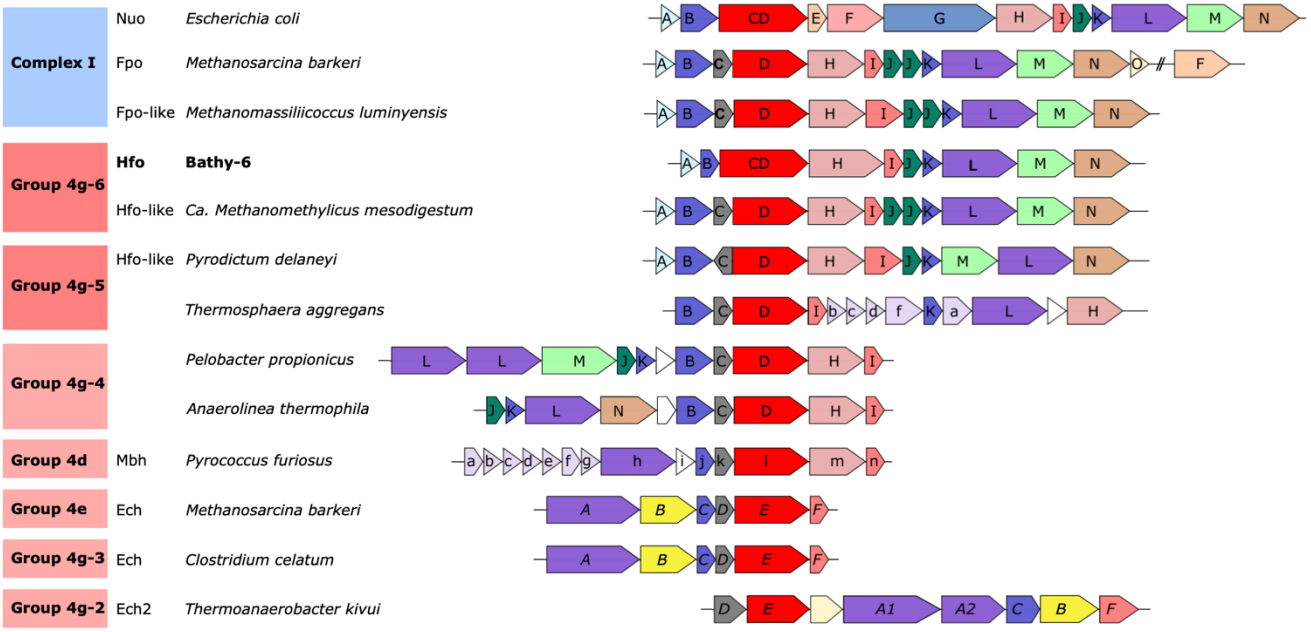


Figure 5. Genomic architecture of the gene clusters encoding the respiratory complex I (Nuo and Fpo) and the ancestral [NiFe] hydrogenases in Group 4. Colors indicate homologous genes; The phylogenetic analysis of the catalytic subunit of [NiFe] hydrogenases and its homologs (labeled in red) is shown in Figure 6. The font style of the gene labels indicates differences in the subunit nomenclature of Nuo/Fpo (uppercase), Mbh (lowercase), and Ech (italics).

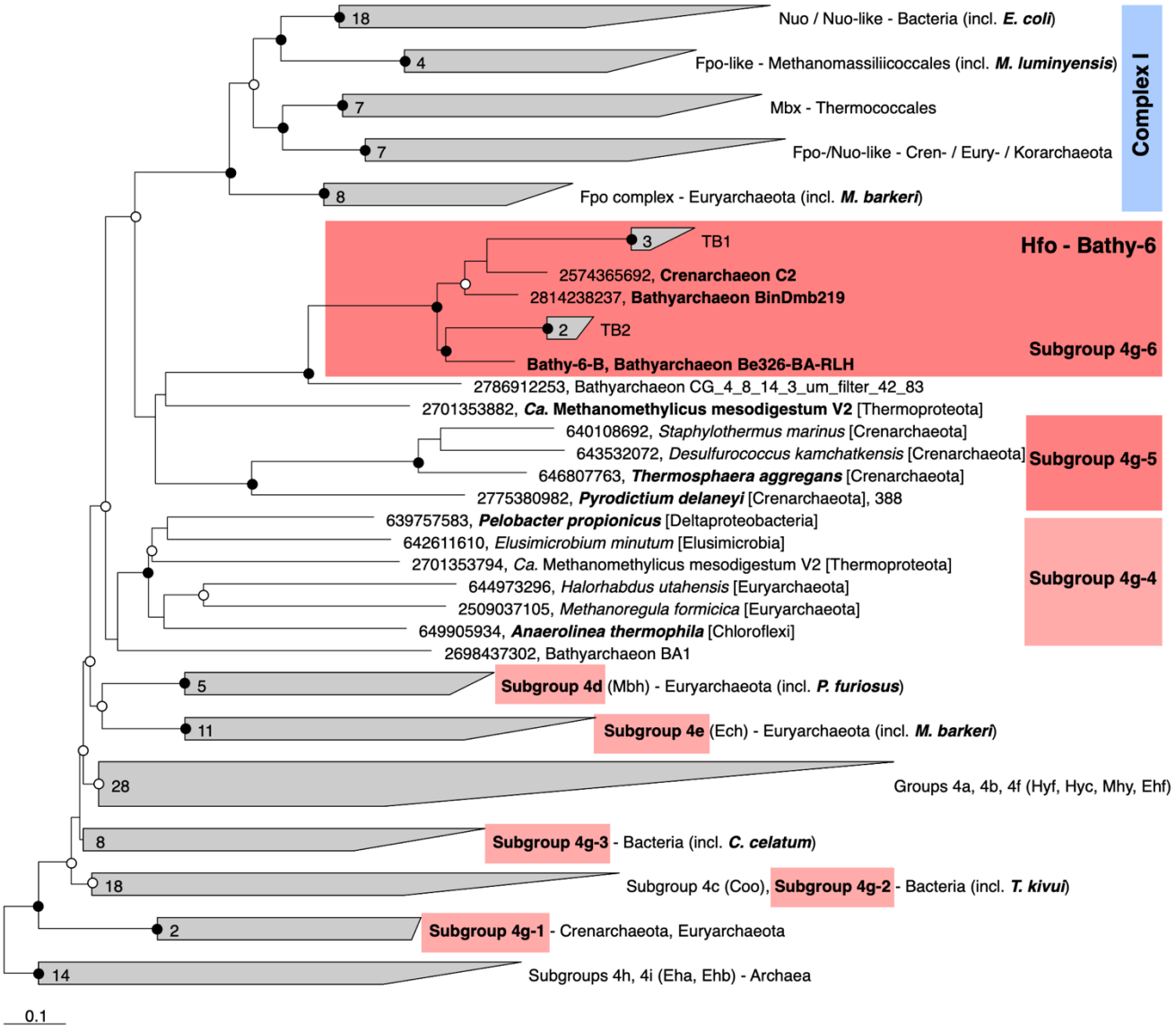


Figure 6. Phylogeny of the catalytic subunits of Group 4 [NiFe] hydrogenases and their homologs in respiratory complex I (FpoD and NuoD). The maximum-likelihood tree is based on a curated alignment of the deduced amino acid sequences; the scale bar indicates 0.1 amino acid substitutions per site. SH-aLRT values (● ≥ 95%; ○ ≥ 80%, 1,000 replications) indicate node support. The genomic context of the highlighted genes is shown in Figure 5. Gene numbers indicate IMG/Mer gene IDs.

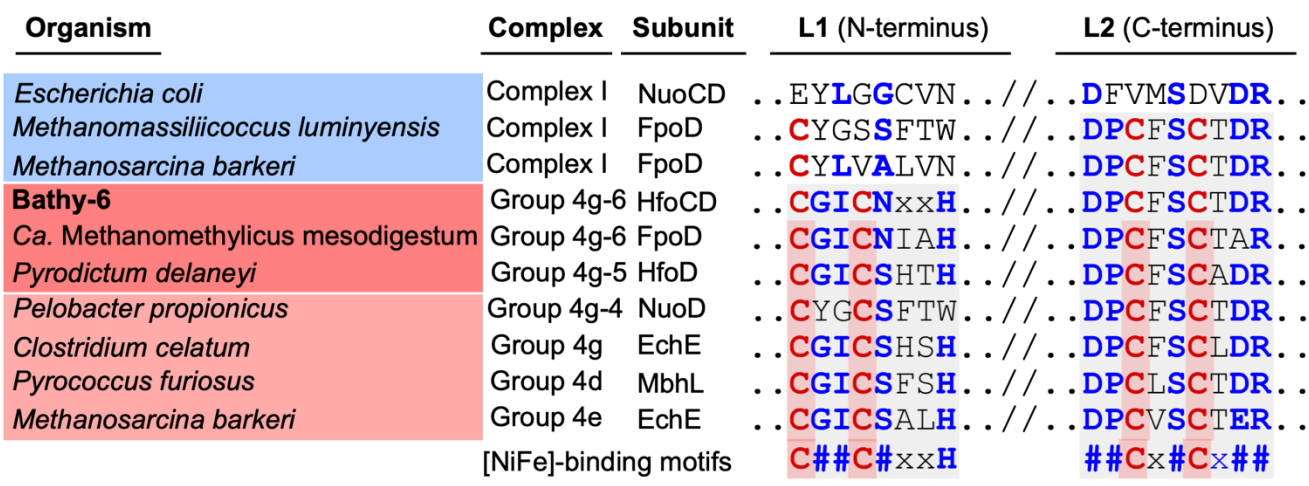


Figure 7. Comparison of the [NiFe]-binding motifs (L1 and L2) in the large subunits of selected Group 4 [NiFe] hydrogenases with the corresponding amino acid residues (IUPAC code) of their homologs in the respiratory complex I (Nuo and Fpo). Gray shading indicates that the typical motifs of [NiFe] hydrogenases are present (L1 motif: C[GS][ILV]C[AGNS]xxH; L2 motif: [DE][PL]Cx[AGST]Cx[DE][RL]; Vignais and Billoud, 2007). The four cysteine residues that coordinate the [NiFe] cluster are in red; other conserved residues are in blue.

11 Supporting Information

Supplementary Figure S1. Average nucleotide identity (ANI) of the MAGs in subgroup Bathy-6. The termite gut *Bathyarchaeia* were assigned to phylotypes based on ANI > 99%. NA indicates ANI values <75%, which are not returned by the fastANI program.

Supplementary Figure S2. Genome-based phylogeny of termite gut *Bathyarchaeia* illustrating the relationship of lineages TB1 and TB2 to other MAGs in the Bathy-6 subgroup. MAGs mentioned in the text are marked in bold. The maximum-likelihood tree was inferred from a concatenated alignment of 43 proteins using the LG+F+I+G4 model and rooted with selected Crenarchaeota and Euryarchaeota as outgroup. The numbers in circles indicate the phylotypes discussed in the text (Table 1). MAGs included in the comparative analysis (Figure 3) are shown in bold. The tree was rooted other archaeal genomes as outgroup. The scale bar indicates 10 amino acid substitutions per site. Node support values (SH-aLRT) are shown in blue. A simplified version of the tree is shown in Figure 1.

Supplementary Figure S3. 16S rRNA-based phylogeny of subgroup Bathy-6, indicating the placement of the sequences from termite guts among those obtained from other environments. The maximum-likelihood tree is based on a curated alignment (1,424 positions) of all sequences in the SILVA database and their homologs retrieved from the bathyarchaeal MAGs (in bold) and the low-quality bins obtained from the termite gut metagenomes (Hervé *et al.*, 2020). The tree was rooted using members of Bathy-5 as outgroup. The scale bars indicate 0.05 nucleotide substitutions per site. Node support values (SH-aLRT) are shown in blue. Branches marked with dashed lines indicate shorter sequences that were added using the ARB parsimony tool. A simplified version of the tree is shown in Figure 2.

885 **Supplementary Figure S4.** The methyltransferase-associated corrinoid protein (CoP) of Bathy-6 and
886 its homologs. (**A**) The canonical methyl transferase system of bacteria and archaea. (**B**) Gene
887 neighborhood of the CoP gene of Bathy-6 and selected homologs (for accession numbers, see panel
888 **C**). Colors indicate the presumed functions of the respective gene products (see panel **A**). Unrooted
889 phylogenetic trees of the methyltransferase-associated CoP genes (**C**) and the associated *mtrH* genes
890 (**D**) of Bathy-6 and their closest relatives (deduced amino acid sequences). Genes that appear in panel
891 **D** are shown in bold. Numbers are IMG/Mer gene IDs. The scale bar indicates 1.0 amino acid
892 substitution per site. Node support values (SH-aLRT) are shown in blue.

893 **Supplementary Figure S5.** Phylogenetic tree of the catalytic subunit of the Hox hydrogenase of
894 Bathy-6 and its homologs among Group 3 [NiFe] hydrogenases. The maximum-likelihood tree is
895 based on deduced amino acid sequences and was rooted [NiFe] hydrogenase sequences of Groups 1
896 and 2. The scale bar indicates 0.5 nucleotide substitutions per site. Node support values (SH-aLRT)
897 are shown in blue.

898 **Supplementary Table S1.** Taxonomic assignment and characteristics of the bathyarchaeotal MAGs
899 from termite guts (from Hervé *et al*., 2020).

900 **Supplementary Table S2.** Annotation details of the genes that encode the metabolic pathways and
901 other functional markers in the 15 bathyarchaeotal MAGs from termite guts, as discussed in the text
902 (see Figures 3 and 4).

903