Adaptations of *Atribacteria* to life in methane hydrates: hot traits for cold life

Jennifer B. Glass[1*], Piyush Ranjan[2#], Cecilia B. Kretz[1§], Brook L. Nunn[3], Abigail M. Johnson[1], James McManus[4], Frank J. Stewart[2]

[1]School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA, USA; [2]School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA, USA; [3]Department of Genome Sciences, University of Washington, Seattle, WA; [4]Bigelow Laboratory for Ocean Sciences, East Boothbay, ME, USA

*Corresponding Author: jennifer.glass@eas.gatech.edu

[#]Now at: Michigan Medicine, University of Michigan, Ann Arbor, Michigan, USA;

[§]Now at: Division of Scientific Education and Professional Development, Epidemiology Workforce Branch, Laboratory Leadership Service, Field Assignee New York City Public Health Laboratory, Centers for Disease Control and Prevention, Atlanta, Georgia, USA

Running Title: *Atribacteria* adaptions in methane hydrate ecosystem

Dedication: To Katrina Edwards

Originality-Significance Statement: This work provides insights into the metabolism and adaptations of elusive *Atribacteria* (JS-1 clade) that are ubiquitous and abundant in methane-rich ecosystems. We show that JS-1 (Genus 1) from methane hydrate stability zones contain metabolisms and stress survival strategies similar to hyperthermophilic archaea.

Summary: Gas hydrates harbor gigatons of natural gas, yet their microbiomes remain mysterious. We bioprospected methane hydrate-bearing sediments from under Hydrate Ridge (offshore Oregon, USA, ODP Site 1244) using 16S rRNA gene amplicon, metagenomic, and metaproteomic analysis. *Atribacteria* (JS-1 Genus 1) sequences rose in abundance with increasing sediment depth. We characterized the most complete JS-1 Genus 1 metagenome-assembled genomic bin (B2) from the deepest sample, 69 meters below the seafloor (E10-H5), within the gas hydrate stability zone. B2 harbors functions not previously reported for *Atribacteria*, including a primitive respiratory complex and myriad capabilities to survive extreme conditions (e.g. high salt brines, high pressure, and cold temperatures). Several *Atribacteria* traits, such as a hydrogenase-Na^+/H^+ antiporter supercomplex (Hun) and di-myo-inositol-phosphate (DIP) synthesis, were similar to those from hyperthermophilic archaea. Expressed *Atribacteria* proteins were involved in transport of branched chain amino acids and carboxylic acids. Transporter genes were downstream from a novel helix-turn-helix transcriptional regulator, AtiR, which was not present in *Atribacteria* from other sites. Overall, *Atribacteria* appear to be endowed with unique strategies that may contribute to its dominance in methane-hydrate bearing sediments. Active microbial transport of amino and carboxylic acids in the gas hydrate stability zone may influence gas hydrate stability.

Introduction

Gas hydrates, also known as clathrates, are cages of ice-like water crystals encasing gas molecules such as methane (CH_4). Because hydrates form under high pressure and low temperature, their distribution on Earth is limited to permafrost and continental margins (Hester and Brewer, 2009). These hydrates harbor gigatons of natural gas, which may serve as a potential energy source for the future (Chong et al., 2016). They are also susceptible to dissociation due to rising ocean temperatures, which could release massive methane reservoirs to the atmosphere and exacerbate global warming (Archer et al., 2009; Ruppel and Kessler, 2017).

Despite the global importance of gas hydrates, their microbiomes remain mysterious. Microbial cells are physically associated with hydrates (Lanoil et al., 2001), and the taxonomy of these hydrate-associated microbiomes is distinct from non-hydrate-bearing sites (Inagaki et al., 2006). Because salt ions are excluded during hydrate formation (Ussler III and Paull, 2001; Bohrmann and Torres, 2006), hydrate-associated microbes likely possess adaptations to survive high salinity and low water activity, as well as low temperatures and high pressures (Honkalas et al., 2016). However, knowledge of the genetic basis of such adaptations is incomplete, as genomic data for hydrate communities are sparse and most hydrate microbiomes have been characterized primarily through single-gene taxonomic surveys.

Global 16S rRNA gene surveys show that the JS-1 sub-clade of the uncultivated bacterial candidate phylum *Atribacteria* is the dominant taxon in gas hydrates (Reed et al., 2002; Inagaki et al., 2003; Kormas et al., 2003; Newberry et al., 2004; Webster et al., 2004; Inagaki et al., 2006; Webster et al., 2007; Fry et al., 2008; Kadnikov et al., 2012; Parkes et al., 2014; Chernitsyna et al., 2016) and in other deep sediment ecosystems with abundant methane (Gies et al., 2014; Carr et al., 2015; Hu et al., 2016). The other major *Atribacteria* lineage, OP-9, has only been found in hot springs (Dodsworth et al., 2013; Rinke et al., 2013) and thermal bioreactors (Nobu et al., 2015). Marine *Atribacteria* are dispersed through ejection from submarine mud volcanoes (Hoshino et al., 2017; Ruff et al., 2019), and environmental heterogeneity may select for locally adapted genotypes. Indeed, *Atribacteria* phylogeny is highly diverse, suggesting the potential for wide functional variation and niche specialization.

Genomic evidence for such *Atribacteria* specialization remains limited. To date, near-complete single-cell and metagenomic sequences from hot springs, wastewater, lake sediments, and non-hydrate bearing marine sediments have shown that *Atribacteria* lack respiratory pathways. The high-temperature OP-9 lineage likely ferments sugars (Dodsworth et al., 2013) whereas the low-temperature JS-1 lineage ferments propionate to hydrogen, acetate, and ethanol (Nobu et al., 2016). Both JS-1 and OP-9 lineages possess genes encoding bacterial microcompartment shell proteins that may sequester toxic aldehydes, enabling their condensation

to carbohydrates (Nobu et al., 2016). The available data on *Atribacteria* genomes suggest diversification linked to organic substrate utilization, although a range of other factors, including physical environmental conditions (e.g., temperature and pressure) undoubtedly also play a role.

Here we examined the distribution, phylogeny, and metabolic potential of uncultivated JS-1 *Atribacteria* in cold, salty, and high-pressure sediments beneath Hydrate Ridge, off the coast of Oregon, USA, using a combination of 16S rRNA gene amplicon, metagenomic, and metaproteomic analysis. We found that JS-1 Genus-1 are abundant in the gas hydrate stability zone (GHSZ) and that they harbor numerous strategies for tolerance of osmotic stress, including many biosynthesis pathways for unusual osmolytes similar to those of thermophiles.

Results and Discussion

Geochemical gradients. Sediment core samples spanned four geochemical zones from 0-69 meters below seafloor (mbsf) at the ODP Site 1244C,D,E at Hydrate Ridge, off the coast of Oregon, USA (Fig. S1; Tréhu et al., 2003): near surface (0-2 mbsf), sulfate-methane transition zone (SMTZ; 2-9 mbsf), metal reduction zone (18-36 mbsf), and GHSZ (45-124 mbsf; **Fig. 1**).

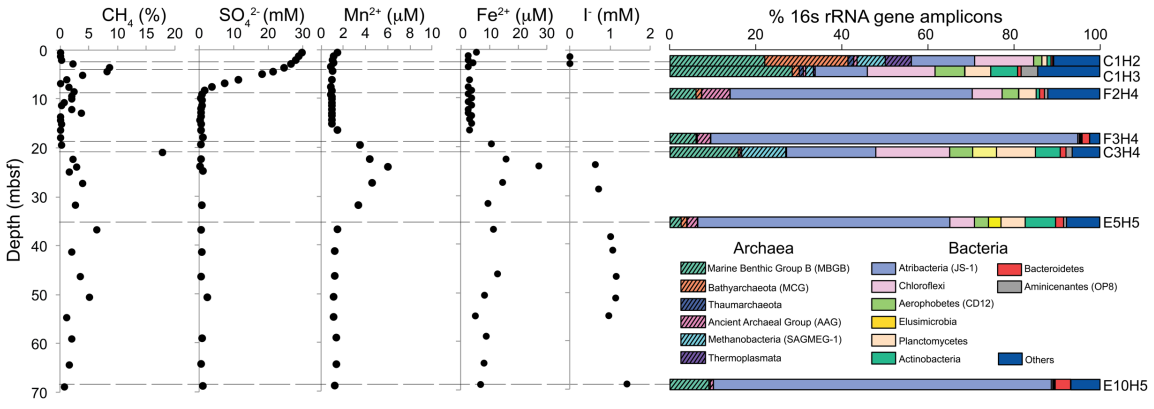


Figure 1. Porewater geochemistry (methane, sulfate, manganese, iron, and iodide) and 16S rRNA gene composition from sediment depth profiles at ODP 204 Site 1244, Hydrate Ridge, offshore Oregon, USA. Hatched and solid bars are archaeal and bacterial 16S rRNA genes, respectively. "Others" category represents bacterial and archaeal phyla with <2% of total sequences.

Sediment porewater methane concentrations rose from negligible at the seafloor to 8% by volume at 3-5 mbsf, and remained <5% below 5 mbsf, with the exception of one sample at 21 mbsf.

Sulfate rapidly dropped from 28 to <1 mM from 0-9 mbsf and remained <1 mM below 9 mbsf, with the exception of one sample at 50.7 mbsf (2.3 mM sulfate). Outside of the metal reduction zone, dissolved Mn was ~1 μM and dissolved Fe was 3-10 μM. Dissolved Mn and Fe peaked at 6 and 27 μM, respectively, coincident with a single layer of disseminated gas hydrate in the metal reduction zone. Dithionite-extractable Fe and Mn increased slightly from 2 to 21 mbsf (0.4 to 1.1% and 0.002 to 0.005%, respectively; **Table S1**). Iodide concentrations were highest in the GHSZ (1.4 mM), where liquid brines form as a result of methane hydrate formation. Estimated *in situ* salinity ranged from seawater salinity (35 g kg^{-1}) to >100 g kg^{-1} (Milkov et al., 2004). Total organic carbon concentrations in sediment varied between 1-2%. *In situ* temperature ranged from ~4°C at the seafloor to ~6-11°C in the GHSZ.

Phylogenetic diversity. Phylogenetic diversity and species richness in 16S rRNA gene amplicons were highest in the SMTZ and decreased with depth except in the metal reduction zone (**Fig. S2**). The relative abundance of *Atribacteria* (JS-1)-affiliated amplicons increased with depth, from 15% in the near surface to 86% in the GHSZ (**Table S1**). GHSZ sediment (sample E10-H5 from 69 mbsf) contained 230 *Atribacteria* OTUs (89-92% ANI) that spanned a wide diversity of clades within JS-1 Genus 1 (Yarza et al. 2014) **(Fig. 2)**. A single OTU matching GenBank AB804573.1, from an ocean drilling core from offshore Shimokita Peninsula, Japan, comprised 69% of *Atribacteria* 16S rRNA sequences in the GHSZ **(Table S2**). Other *Atribacteria* 16S rRNA sequences also matched marine samples from shallower Hydrate Ridge sediments (Marlow et al., 2014) and methane hydrate sediment off Taiwan (Lin et al., 2014) (**Table S2**). 16S rRNA sequences from amplicons and metagenomes generally showed consistent trends (**Fig. S3**). *Atribacteria* OTU abundance and composition varied significantly with sediment depth (**Fig. S4**).

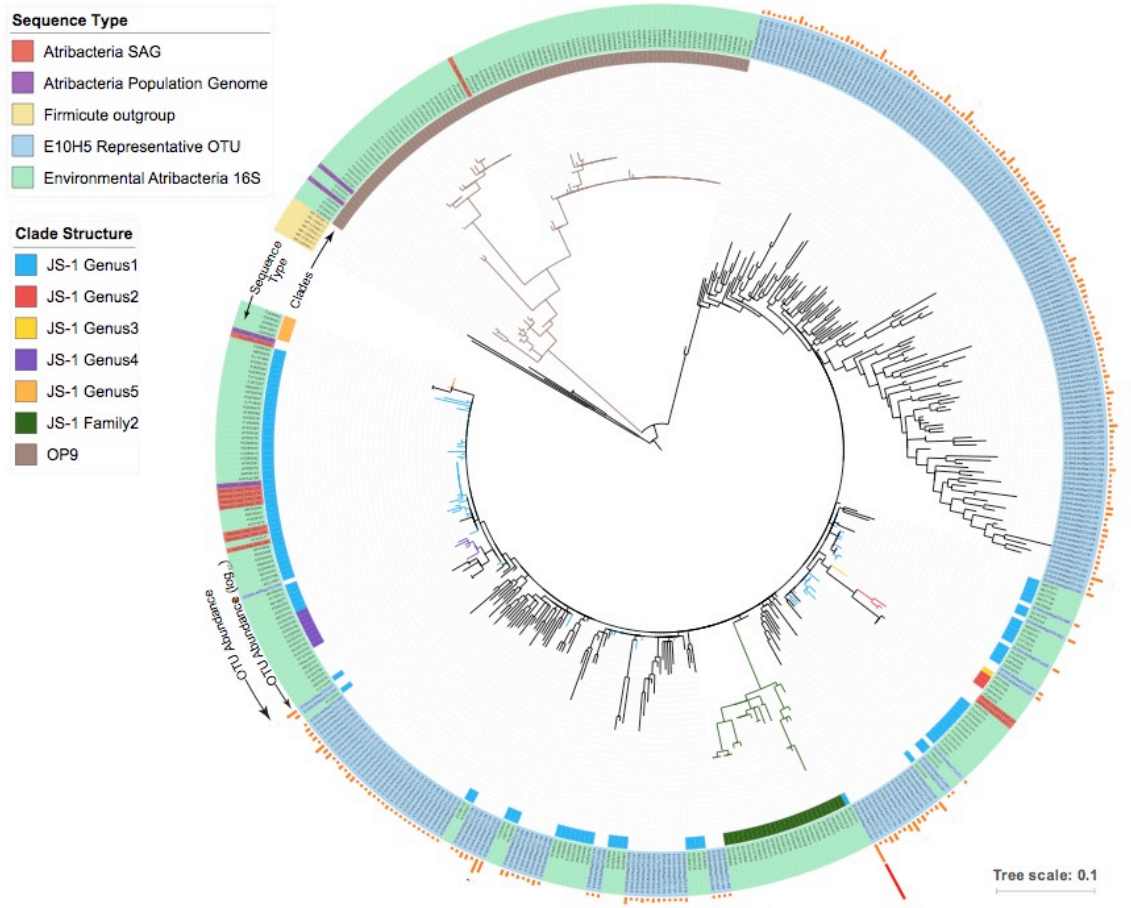


Figure 2: Phylogenetic reconstruction of *Atribacteria* 16S rRNA gene sequences from sample E10-H5 (69 mbsf). The tree includes the 230 *Atribacteria* OTUs with two or more sequences as well as reference sequences from environmental clones, SAGs, and MAGs, with *Firmicutes* as the outgroup. Reconstruction was performed in RAxML with 275 positions spanning the V3-V4 region of the 16S rRNA gene using a GAMMA model of rate heterogeneity, a GTR model of substitution, and 500 bootstraps followed by a thorough Maximum Likelihood search. The relative abundances of recovered amplicons from diverse lineages/OTUs is shown in the outermost circle. Additional information on the most abundant JS-1 OTUs from E10-H5 is provided in **Table S2**.

JS-1 Genus-1 partial genome. To gain insight into the function of JS-1 *Atribacteria* in the GHSZ, we analyzed a 4-Mbp metagenome-assembled genome (MAG) from sample E10-H5 (**Table S3**). This MAG, hereafter designated "B2", was chosen for its relatively high completeness (69%) and low contamination (2%). B2 lacked a 16S rRNA gene, but contained a *rpoB* gene with 94% similarity to *Atribacteria* bacterium 34_128 from an oil reservoir (Hu et al., 2016). B2 had 35% GC content, similar to other *Atribacteria* (Carr et al., 2015). Phylogenetic placement based on 69 concatenated single-copy genes confirmed that B2 belonged to JS1-Genus

1 and was most closely related to JS1-Genus 1 genomes from a sediment-hosted aquifer at Rifle, Colorado (RBG_COMBO_35; Anantharaman et al., 2016), cold CO_2-rich fluids at Crystal Geyser, Utah (CG2_30_33_13; Probst et al., 2017), and hydrothermal vent sediments at Guaymas Basin, Gulf of California (4572_76; Dombrowski et al., 2017) (**Fig. 3**).

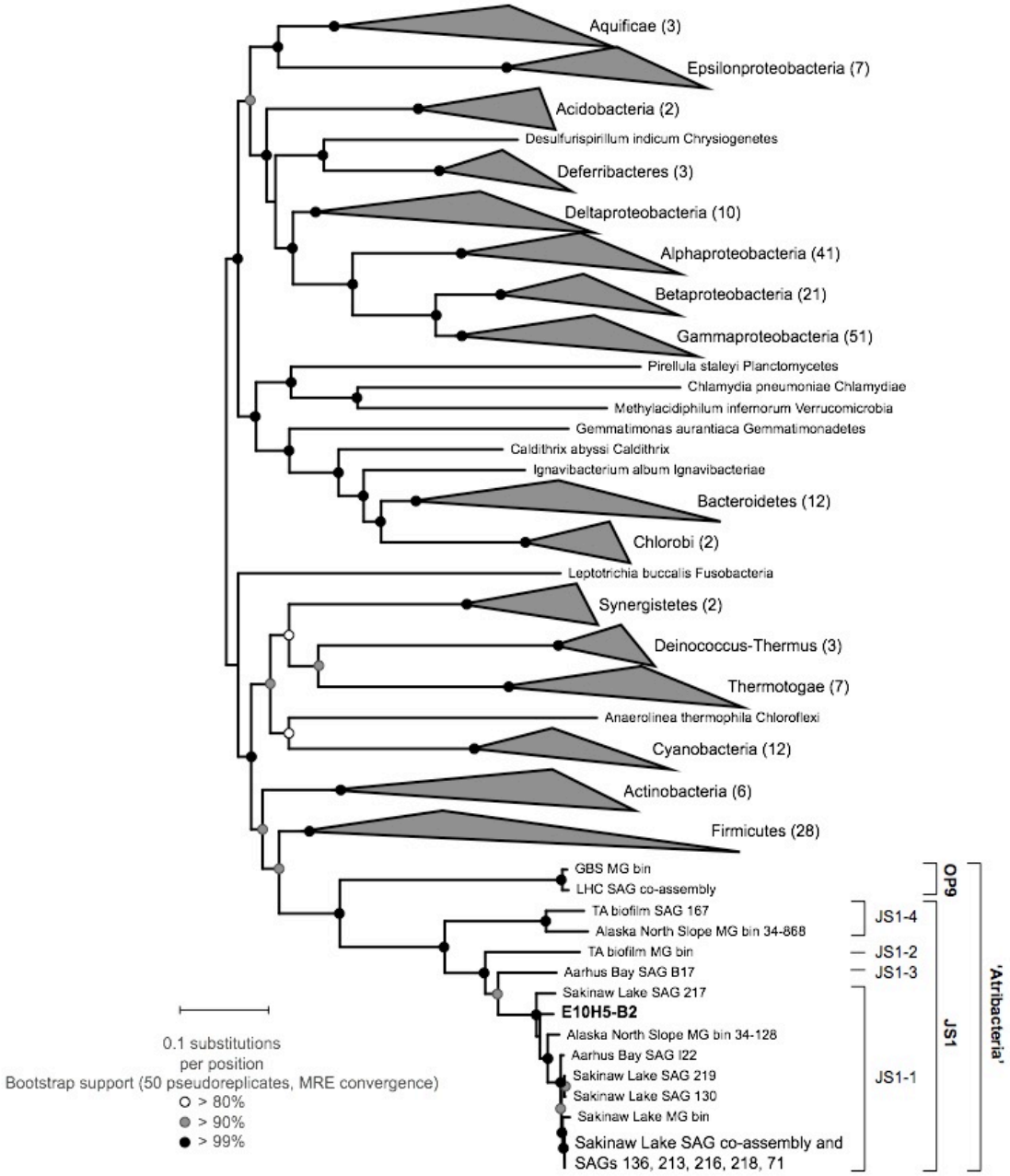


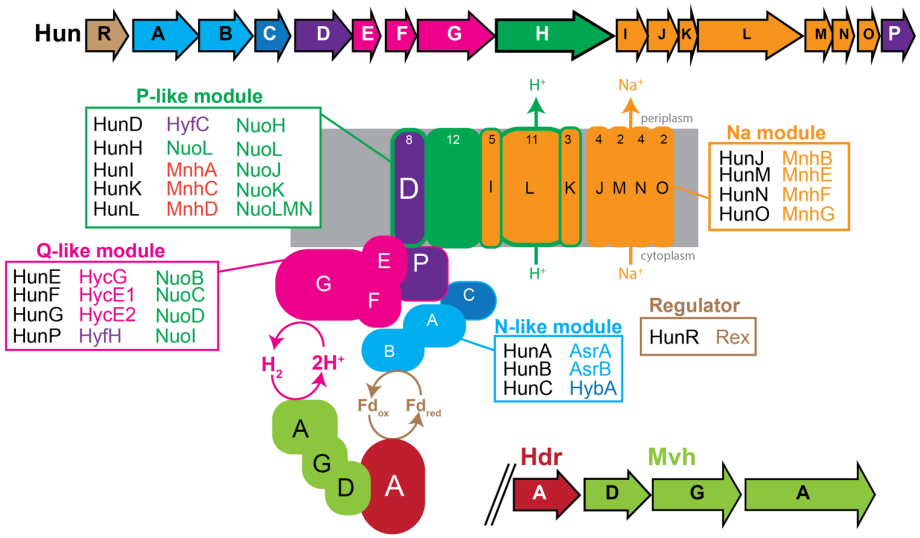
Figure 3: **Maximum likelihood phylogeny for B2 with 220 representative and 20 previously found *Atribacteria* SAGs and population genomes using multiple (minimum 6, maximum 69) core single copy genes.** Tree made in RAxML with GAMMA model, 1000 rapid bootstraps, MRE convergence bootstop (50 replicates) followed by a thorough ML search.

Despite the relatively cool *in situ* temperature of the E10-H5 sediment (7-8°C (ShipboardScientificParty, 2003)), the most closely related genomes from cultured isolates were thermophilic gram-positive Firmicutes: halophilic *Halothermothrix orenii* spp. (Mavromatis et al., 2009) and metal-respiring *Therminocola potens* strain JR (Byrne and Nicholas, 1986). Below we highlight features of the B2 genome and proteome potentially relevant to life in the unique environment of methane clathrates, with particular focus on a putative respiratory complex and genes involved in stress response and environmental homeostasis.

Predicted respiratory function of the Hun supercomplex. B2 contained genes for a putative operon encoding a 16-subunit respiratory complex, hereafter designated Hun. The *hun* operon was also present in two other MAGs from ODP Site 1244 (*Planctomycetes* C1H3-B36 and *Firmicutes* E5H5-B3) and in *Atribacteria, Actinobacteria*, and *Omnitrophica* MAGs from other deep subsurface ecosystems (Rinke et al., 2013; Baker et al., 2015; Anantharaman et al., 2016; Probst et al., 2017) (**Table S4**). The gene arrangement and predicted function of the putative Hun complex are similar to those of an ancient Mrp-Mbh-type membrane-bound [NiFe] hydrogenase-Na^+/H^+ antiporter respiratory complex in hyperthermophilic archaea (Yu et al., 2018), which is thought to be the ancestor of Complex I, also known as NADH:ubiquinone oxidoreductase (Nuo) (Friedrich and Scheide, 2000; Moparthi and Hägerhäll, 2011; Schut et al., 2013). Complex I's modules likely had separate origins: the ubiquinone-reducing subunits NuoBCD ("Q-module") evolved from an ancient membrane-bound [NiFe] hydrogenase, while its proton-pumping subunits NuoLMN ("P module") evolved from an ancient Na^+/H^+ antiporter (Mathiesen and Hägerhäll, 2002; Moparthi et al., 2014; Spero et al., 2015).

Atribacteria hun genes likely encode a complex of four protein modules that couple H^+ and Na^+ translocation to H_2 production, similar to Mrp-Mbh-type complexes in hyperthermophilic archaea **(Fig. 4**). Based on the similarity of HunAB to anaerobic sulfite reductase (Asr) subunits A and B, which transfer electrons from ferredoxin to the active site in AsrC (missing in the *hun*

176 operon), we inferred that N module-like subunits HunABC likely accept electrons from
177 ferredoxin and pass them through iron-sulfur clusters to Q-module-like subunit HunEFGP.
178 Instead of accepting electrons from NADPH and passing them to ubiquinone as in Complex I,
179 HunABC likely accepts electrons from ferredoxin and passes them to $2H^+$ for reduction to H_2 at
180 HunEFGP's Ni-Fe active site **(Table 1; Fig. 4**).



181

182 **Figure 4: Predicted structure and function of a multi-subunit respiratory complex,**
183 **hereafter "Hun", found in B2 and other deep subsurface genomes.** *Top:* conserved gene
184 cluster arrangement, with each color representing a different predicted protein. *Below:* predicted
185 cellular locations and functions based on homologs of the genes of the same colors encoded by
186 the putative *hun* operon, and predicted regeneration of substrates by the heterodisulfide reductase
187 (HdrA)-methyl viologen hydrogenase (MvhAGD) complex. Predicted functions of *hun* genes are
188 based on Mrp-Mbh complexes in thermophilic archaea (Schut et al., 2013; Yu et al., 2018). See
189 **Table S4** for accession numbers.
190

191 P-module-like subunits HunDHILK are predicted to be proton-pumping transmembrane proteins
192 and Na-module-like subunits HunIJKLMNO are homologs of the Na^+/H^+ antiporter
193 MnhABCDEFGH in Mrp-Mbh-type complexes. The presence of F0F1-type and V-type ATPases
194 suggest that H^+ and Na^+ ions pumped outward by HunIJKLMNO are pumped back in to make
195 ATP. Electrons from H_2 could be transferred back to ferredoxin by the activity of the
196 heterodisulfide reductase (HdrA)-methyl viologen hydrogenase (MvhAGD) complex (**Fig. 4**). A
197 redox-sensing transcriptional repressor gene (*hunR*) immediately upstream of the *hun* operon

198 suggests that the hydrogenase may not be used strictly for energy conservation, but could also be
199 for balancing reducing equivalent by disposing of extra electrons (McLaughlin et al., 2010).
200
201 **Osmotic stress survival.** Any life that can persist in brine pockets within methane hydrate must
202 contend with high salinity (up to ~3x that of seawater) and low water potential. B2 contained
203 numerous genes for the "salt out" survival strategy, in which osmotic pressure is maintained by
204 exporting cations (Wood, 2015). B2's cation export systems included efflux systems,
205 mechanosensitive ion channels, and Na^+-H^+ antiporters (**Table 1**).
206 A second salt survival strategy is import and/or biosynthesis of osmolytes, most often
207 polar, water-soluble, and uncharged organic compounds and/or extracellular polymers. For
208 example, glycine betaine is abundant in saline fluids from deep sediment basins (Daly et al.,
209 2016). B2 contained genes for transport of trehalose and biosynthesis of the common osmolytes
210 glutamine, glutamate, and poly-gamma-glutamate, all of which had homologs in other
211 *Atribacteria* MAGs (**Table 1**). B2 also encoded genes for glycine betaine and dihydroxyacetone
212 biosynthesis without homologs in other *Atribacteria*. Surprisingly, B2 also encoded biosynthetic
213 genes (myo-inositol-1 phosphate synthase (MIPS)/bifunctional IPC transferase and DIPP
214 synthase (IPCT-DIPPS)) for the unusual solute di-myo-inositol-phosphate (DIP) made by
215 hyperthermophiles (Santos and Da Costa, 2002). The MIPS gene had closest similarity to
216 halophilic and psychrophilic *Euryarchaeota*, without homologs in other *Atribacteria*. The IPCT-
217 DIPPS gene was also present in *Atribacteria* HGW-1 from subsurface Japan (Hernsdorf et al.,
218 2017) and *Atribacteria* 4572_76 from Guaymas Basin (Dombrowski et al., 2017).
219 Immediately upstream from B2's MIPS/IPCT-DIPPS genes was an acyl carrier protein
220 (*acpP*) gene, commonly involved in fatty acid and polyketide biosynthesis. Sixteen additional
221 *acpP* copies were present in B2, often flanked by transposon scars, suggestive of recent
222 horizontal gene transfer (**Table S5)**. Other *Atribacteria* MAGs had only 1-2 copies of *acpP*,
223 usually near fatty acid biosynthesis genes.

Table 1. Putative osmotic stress-related genes in B2. *Atribacteria* homologs all had >80% AAI. AAI to other taxa (56-76%) are provided. *indicates multiple copies.

Annotation	Gene	Accession	Top hit	Top hit
Na^+/H^+ antiporter	*mrpEFGB*	RXG65834.1-RXG65838.1	OQY40657.1-OQY40661.1	*Atribacteria* 4572_76
Na^+ efflux	*natB*	RXG65900.1	OGD31203.1	*Atribacteria* RBG....
Threonine efflux	*rhtB*	RXG66248.1	OGD15641.1	*Atribacteria* RBG....
Na^+ channel	DUF554	RXG63559.1	KUK55705.1	*Atribacteria* 34_128
Mechanosensitive ion channel	*mscS*	RXG63036.1	KUK56353.1	
Trehalose transporter	*sugAB*	RXG66833.1-RXG66834.1	KUK55397.1 KUK55398.1	
Glutamine synthetase	*glnA*	RXG65164.1	KUK55578.1	
K^+ transport	*trkAH**	RXG63511.1 RXG63512.1	PKP56013.1 PKP56012.1	*Atribacteria* HGW-1
Aromatic aa exporter	*yddG**	RXG63201.1	PKP55084.1	
Glutamate synthase	*gltD*	RXG66270.1	PKP56573.1	
Proline racemase	*prdF*	RXG63210.1	PKP58887.1	
Poly-gamma glutamate synthase	*pgsCBW*	RXG66317.1-RXG66319.1	PKP60458.1-PKP60460.1	
Glycerol uptake	*glpF*	RXG65629.1	OHV10031.1 (61%)	*Kushneria* YCWA18
Betaine-aldehyde dehydrogenase	*betB*	RXG62957.1	KUJ28189.1 (56%)	*Catabacter hongkongensis*
Dihydroxy-acetone kinase	*dhaKLM*	RXG65626.1-RXG65628.1	RLC64130.1- (67%) RLC64131.1 (61%)	*Chloroflexi* bacterium
DIPP synthesis pathway	MIPS/IPCT-DIPPS*	RXG66889.1 RXG66888.1	AAU82306.1 (76%) PKP58414.1	Archaeon GZfos13E1 *Atribacteria* HGW-1

Like other *Atribacteria*, B2 contained genes encoding a sugar phosphate-utilizing class of proteinaceous bacterial microcompartments that neighbored sugar isomerases, RnfC NADH dehydrogenase and an oxidoreductase (Axen et al., 2014; Nobu et al., 2016) (**Table S6**). Further exploration of sugar-related genes revealed that B2 and other *Atribacteria* encode the non-mevalonate pathway for isoprenoid biosynthesis (*ispDEFGH*), exopolysaccharide synthesis proteins, numerous glycosyltransferases for transferring UDP- and GDP-linked sugars to a variety of substrates, and several proteins related to N-linked glycosylation (**Table S7**). The capacity for glycosylation may be another adaptation for survival of salt stress (Kho and Meredith, 2018).

Expression of lipopolysaccharide and transport-related proteins. Metaproteomic analysis identified six expressed peptides affiliated with B2, all associated with assembly or transport (**Table 2**). One was an outer member lipopolysaccharide assembly protein (LptD), also known as

239 Imp/OstA (increased membrane permeability/organic solvent tolerance (Braun and Silhavy,
240 2002). Another was a capsular polysaccharide biosynthesis protein (YveK). The other expressed
241 peptides were predicted to be transporters of purines (BmpA), branched chain amino acids (LivH,
242 LivM), and C4-dicarboxylates (DctQ). All *liv* genes on the operon with the expressed *livH* had
243 homologs in other *Atribacteria* genomes (**Table S8**) with the exception of *livG*, which encodes a
244 protein related to the lipopolysaccharide export system ATP-binding protein LptB that may serve
245 a specific purpose in methane-hydrate *Atribacteria*. Upstream of *liv* genes we found a *ykkC-yxkD*
246 riboswitch implicated in detoxification and efflux control (Barrick et al., 2004), suggesting that
247 branched chain amino acids may be involved in environmental stress response, as seen in other
248 microbes (Liu et al., 2005).

249 **Table 2. Metaproteomic peptide hits for B2.**

Peptide	Protein	Contig	Gene	Top hit (% identity)	Top hit
EYKPKEDWKMNFSS SYNLNTK	LptD	C10125	33494	OQY39007.1 (90%)	*Atribacteria* 4572_76
GIIILIFLIAVITAVLV SYFVLSPTP	YveK	C456	RXG64813.1	PKP59499.1 (74%)	*Atribacteria* HGW-1
CSNLIIKALLVVLVL SLGITLGIAKAP	BmpA	C473	RXG64193.1	PKP58720.1 (94%)	*Atribacteria* HGW-1
KPFRKSPGLIILLSTV AVGFIIR	LivH	C8009	30420	OQY40503.1 (94%)	*Atribacteria* 4572_76
LIFLLLLAVAVVVPF LLGLLILRF	LivM	C2171	15004	RKY02958.1 (46%)	*Spirochaetes* bacterium
NKINLIFSILIIIFLIVL TYEGIILVKVGLNA	DctQ	C95	RXG62936.1	AEG13811.1 (34%)	*Desulfofundulus kuznetsovii*

250

251 In addition to numerous transporters for branched chain amino acids, B2 encoded abundant
252 TRAP (tripartite ATP-independent periplasmic) transporters of dicarboxylic (DctPQM) and
253 tricarboxylic (TctCBA) acids (**Table S7; Fig. 5**). TRAP transporters use an electrochemical
254 gradient (H^+ or Na^+) and a substrate-binding protein to transport solutes across the membrane
255 (Fischer et al., 2010). A conserved arginine residue in the DctP substrate-binding protein confers
256 specificity for carboxylate groups (Lecher et al., 2009; Fischer et al., 2015).

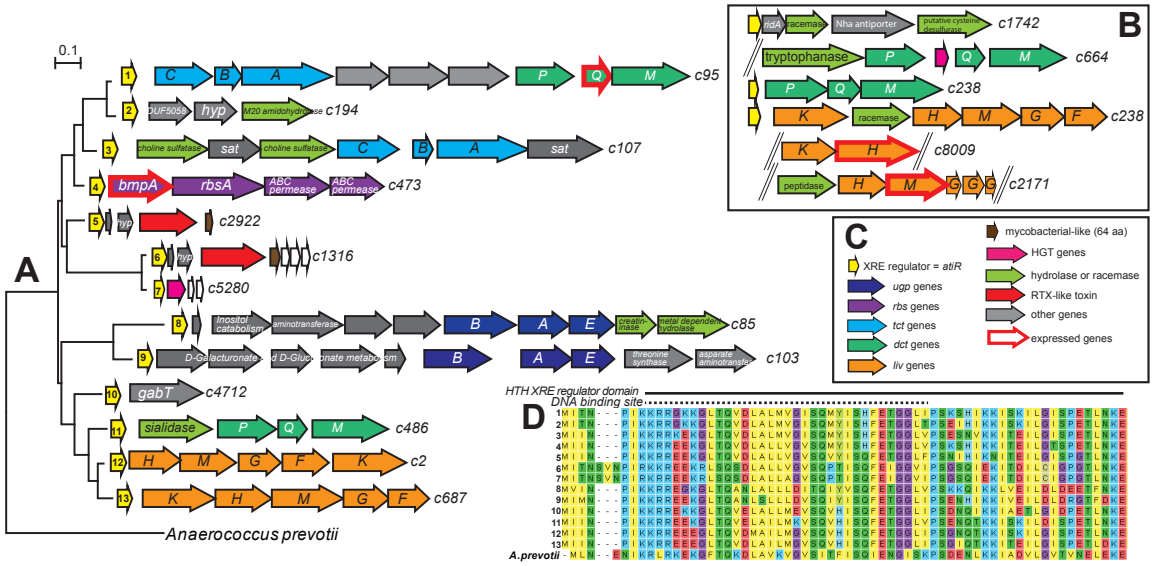


Figure 5: Phylogeny of HTH-XRE regulators/antitoxins (yellow), hereafter "AtiR", from B2 and synteny of downstream genes. Genes highlighted in thick red lines were expressed in the metaproteome. **A)** AtiR maximum likelihood phylogeny based on contigs (labeled on the right) from E10-H5 B2, with *Anaerococcus prevotii* as the outgroup. **B)** Additional putative operons from B2 likely regulated by *atiR*, which is truncated partially or completely on these contigs. **C)** Legend for panels A and B; **D)** AtiR amino acid alignment for the 13 AtiR sequences from *Atribacteria* E10-H5-B2 shown in panel A. Abbreviations: *bmpA*: basic membrane protein A; *dctPQM*: C4-dicarboxylate transporter; *gabT*: 4-aminobutyrate aminotransferase; *livHMGF*: branched chain amino acid transporter; *rbs*: ribose transporter; *sat:* sulfate adenylyltransferase; *tctCBA*: tricarboxylate transporter; *ugpBAE*: sn-glycerol-3-phosphate transporter. See **Table S7** for accession numbers and % identity to closest gene hits in other genomes.

A novel regulator. Three out of six of the expressed transporter proteins were encoded by genes located downstream from a novel gene predicted to encode a helix-turn-helix xenobiotic response element transcriptional regulator, which we named "AtiR" (**Table S8; Fig. 5**). AtiR was not found in *Atribacteria* MAGs (the top BLAST hit was the skin firmicute *Anaerococcus prevotii* (41-49% AAI)), suggesting that it may serve a specific purpose in methane-hydrate *Atribacteria*. Genes downstream of *atiR* were dominated by transporters for organic solutes (*tct, dct, ugp*), branched chain amino acids (*liv*), hydrolases (choline sulfatase, sialidase, tryptophanase, cysteine desulfurase), peptidases, and racemases (**Table S8; Fig. 5**). In two instances, genes encoding RTX-toxin repeats were located on *atiR* contigs (**Table S8**). B2 also contained numerous MazEF toxin-antitoxin systems (**Table S9**), which trigger programmed cell death in response to stress (Engelberg-Kulka et al., 2005). *Atribacteria* may use AtiR to regulate cellular degradation of

281 peptides and proteins to amino acids, either for nutrients acquisition or for survival under
282 environmental stress (Bergkessel et al., 2016).
283
284 **Adaptations to life in methane hydrates.** The GHSZ in deep subsurface sediments is dominated
285 by *Atribacteria* that appear to contain unique adaptations for survival in an extreme system with
286 high salinity, high pressure, low water activity, and low temperatures. Our analysis of the B2
287 *Atribacteria* MAG from the GHSZ (69 mbsf at Hydrate Ridge, offshore Oregon, *in situ* sediment
288 temperature ~6-11°C) revealed multiple survival strategies with similarity to hyperthermophiles.
289 In B2, these "hot traits in cold life" included genes for an ancient respiratory system (Hun) and an
290 unusual osmolyte (DIP). Other probable environmental stress adaptations include glycosylation,
291 membrane modifications, and a novel regulatory mechanism (AtiR) for transport of carboxylic
292 acids and branched chain amino acids.
293 Our findings suggest that *Atribacteria* may actively modulate the composition and
294 concentration of organic compounds in methane hydrate sediments. Active cellular transport of
295 organics would change environmental concentrations, which in turn could influence hydrate
296 stability. The hydrophobicity of branched chain amino acids has been shown to influence hydrate
297 stability; less hydrophobic amino acids like glycine and alanine inhibit hydrate formation by
298 disrupting the hydrogen bond network, while more hydrophobic amino acids, such as leucine,
299 valine and isoleucine, promote hydrate growth by strengthening the local water structure (Sa et
300 al., 2013; Liu et al., 2015; Veluswamy et al., 2017). Gas hydrate growth is also promoted by
301 anionic surfactants (Kumar et al., 2015), which include carboxylic acids. Thus, we surmise that
302 bacterial transport of organic compounds may influence hydrate stability. Our results motivate
303 future studies of methane stability that account for the influence of microbial processes, in
304 particular those of abundant *Atribacteria*.
305

Acknowledgments. We thank Vinayak Agarwal, Jennifer Biddle, Jordan Bird, Frederick Colwell, Sheng Dai, Konstantinos Konstantinidis, Peter Girguis, Julie Huber, Raquel Lieberman, Karen Lloyd, Katie Marshall, Alejandra Prieto Davo, Brandi Reese, Claudia Remes, Despina Tsementzi, Paula Welander, Loren Williams, Jieying Wu, and Jenny Yang for helpful discussions; Phil Rumford and curatorial staff at the ODP Gulf Coast Repository for providing samples; and Shweta Biliya, Annie Hartwell, and Janet Hatt for technical assistance. This research was funded by Center for Dark Energy Biosphere Investigations (C-DEBI) Small Research Grant to J.B.G. and C.B.K. (NSF OCE-0939564), NASA Exobiology grant to J.B.G. and F.J.S. (NNX14AJ87G), NSF Biology Oceanography grant to F.J.S and J.B.G. (NSF OCE-1558916), and a Georgia Tech Earth and Atmospheric Sciences Frontiers Postdoctoral Fellowship to C.B.K. Metaproteomic analysis by B.L.N. was partially supported by the University of Washington's Proteomic Resource (UWPR95794). This is C-DEBI contribution [provided upon paper acceptance].

Experimental Procedures

Sample collection. Sediments were cored at ODP site 1244 (44°35.1784´N; 125°7.1902´W; 895 m water depth; **Fig. S1**) on the eastern flank of Hydrate Ridge ~3 km northeast of the southern summit on ODP Leg 204 in 2002 (Tréhu et al., 2003) and stored at -80°C at the ODP Gulf Coast Repository.

Geochemistry. Data for dissolved methane, sulfate, manganese, iron, and iodide in sediment porewaters were obtained from (Tréhu et al., 2003). Reactive iron and manganese were extracted from frozen sediments using the citrate-dithionite method (Roy et al., 2013) and measured by inductively coupled plasma optical emission spectrometer (Agilent Technologies 700 Series). Total carbon, total nitrogen and total sulfur were determined by CNS analyzer (Perkin Elmer 2400). Total inorganic carbon was measured by CO_2 coulometer (CM5130) with a CM5130 acidification module. Geochemical metadata are given in **Table S1** and archived in **BCO-DMO project 626690**.

DNA extraction. DNA was extracted, in duplicate, from 8-20 g of sediment from the following depths in meters below seafloor (mbsf): 1.95-2.25 (C1-H2); 3.45-3.75 (C1-H3); 8.60 (F2-H4); 18.10 (F3-H4); 20.69 (C3-H4); 35.65 (E5-H5); 68.55 (E10-H5); 138.89 (core E19-H5) using a MO-BIO PowerSoil total RNA Isolation Kit with the DNA Elution Accessory Kit, following the manufacturer protocol without beads. Approximately 2 grams of sediments were used per extraction, and DNA pellets from the two replicates from each depth were pooled together. DNA concentrations were measured using a Qubit 2.0 fluorometer with dsDNA High Sensitivity reagents (Invitrogen, Grand Island, NY, USA). DNA yields ranged from 4-15 ng per gram of sediments. Core E19-H5 (139 mbsf) yielded only 2 ng DNA per gram of sediment and yielded unreliable data due to contamination with sequences from the enzymes used in the library preparations. Therefore, this core segment was excluded from further analysis.

16S rRNA gene amplicon sequencing. Microbial community composition was assessed by Illumina sequencing of the V3-V4 region of the 16S rRNA gene. The V3-V4 region was PCR-amplified using primers F515 and R806 (Caporaso et al., 2011), each appended with barcodes and Illumina-specific adapters according to (Kozich et al., 2013). Reactions consisted of 1-2 µL DNA template (2 ng), 5 µL of 10x Taq Mutant reaction buffer, 0.4 µL of Klentaq LA Taq Polymerase (DNA Polymerase Technology, St. Louis, MO, USA), 2 µL of 10 mM dNTP mix (Sigma Aldrich, St. Louis, MO, USA), 2 µL of reverse and forward primers (total concentration 0.4 µM), and the remainder DNA-free water to 50 µL (Ambion, Grand Island, NY, USA). PCR conditions were an initial 5-min denaturation at 94ºC, followed by 35 cycles of denaturation at 94ºC (40 sec), primer annealing at 55ºC (40 sec), and primer extension at 68ºC (30 sec). Amplicon libraries were purified using a QIAquick PCR Purification Kit (Qiagen, Germantown, MD, USA), quantified by Qubit (Life Technologies), and pooled in equimolar concentration. Amplicons were sequenced on an Illumina MiSeq across two different runs using the V2 500-cycle kit with 5% PhiX to increase read diversity. 16S rRNA sequences were deposited into NCBI SAMN04214977-04214990 (**PRJNA295201**).

16S rRNA gene amplicon analysis. Sequences were trimmed using Trim Galore (criteria: length >100 bp length, Phred score >25), and paired reads were merged using FLASH (Magoč & Salzberg, 2011) with the criteria of a minimum length of 250 bp per input read, minimum length of 300 bp for merged fragments, and maximum fragment standard deviation of 30 bp. Merged reads were imported into QIIME1 (Caporaso et al., 2010) and chimeric sequences were detected by searches using 'identity_chimeric_seqs' and then removed. Sequences sharing 97% nucleotide similarity were clustered into operational taxonomic units (OTUs) using 'pick_open_reference_otus' with taxonomy assigned to OTUs by comparison to the greengenes

database (DeSantis et al., 2006). The datasets were rarefied to a uniform depth of 14,391 sequences, and the rarefied OTU table was used for all downstream analyses. A core set of QIIME diversity analyses was performed using 'core_diversity_analyses'. The phylogenetic diversity (PD) metric (Faith, 1992) was used to quantify alpha diversity across samples.

***Atribacteria* OTU phylogenetic analysis.** We generated a reference alignment of *Atribacteria* full length 16S rRNA sequences to use as a scaffold for mapping OTU sequences generated in this study. The reference alignment included *Atribacteria* 16S rRNA gene sequences from environmental clones (from Nobu et al. (2016), Carr et al. (2015) and Yarza et al. (2014)) and published SAGs and MAGs available in Prokka at the time of analysis (spring 2018), as well as 8 sequences from *Firmicutes* bacteria for use as an outgroup. The sequences were aligned in MAFFT with the linsi option, alignment reordering, and reverse complement matching enabled. We then extracted representative sequences from 230 OTU clusters identified as *Atribacteria* OP-9 and JS-1 in the E10-H5 amplicon dataset; OTUs represented by only a single sequence were excluded. These sequences were recruited to the reference alignment via MAFFT using previously described parameters, without modifying base pair positions in the reference alignment. The alignment was manually inspected and trimmed to include only the V3-V4 region spanned by the *Atribacteria* OTU sequences, resulting in a final alignment with 275 bases.

This alignment was used for phylogeny reconstruction in RAxML with a GTR model of base substitution and GAMMA model of rate heterogeneity, and 500 rapid bootstraps followed by a thorough ML search. The resulting phylogenetic tree was edited for viewing in iTOL. The relative abundance of each OTU (from which a representative sequence was extracted) was mapped onto the resulting phylogeny and shown as a proportion of total sequences in the amplicon dataset.

Pairwise distances between all *Atribacteria* sequences in the alignment were calculated using the p-distance method in MEGA7 and summarized in R as: min 0.0, 1st quartile 0.5, median 0.09, mean 0.11, 3rd quartile 0.18 and max 0.27. Pairwise distances between only the OTUs generated in this study were summarized in R as: min 0.004, 1st quartile 0.056, median 0.075, mean 0.076, 3rd quartile 0.095 and max 0.194.

***Atribacteria* community structure.** OTU abundance from the rarified *Atribacteria* OTU table (previously generated during diversity analysis) was used for NMDS analysis after square root transformation and calculation of Bray-Curtis dissimilarity metrics, all processed via the metaMDS function from Vegan package in R. After examination of the Shepard plot for scatter around the regression line, the NMDS plot was created showing individual OTUs and the midpoint for whole communities. A hierarchical clustering dendrogram was generated using Bray-Curtis dissimilarities.

Multiple displacement amplification, library preparation, and sequencing. Genomic DNA was amplified using a REPLI-g Single Cell Kit (Qiagen, Germantown, MD, USA) using UV-treated sterile plasticware and reverse transcription-PCR grade water (Ambion, Grand Island, NY, USA). Quantitative PCR showed that the negative control began amplifying after 5 hr of incubation at 30°C, and therefore, the 30°C incubation step was shortened to 5 hr using a Bio-Rad C1000 Touch thermal cycler (Bio-Rad, Hercules, CA, USA). DNA concentrations were measured by Qubit. Two micrograms of MDA-amplified DNA were used to generate genome libraries using a TruSeq DNA PCR-Free Kit following the manufacturer's protocol (Illumina, San Diego, CA, USA). The resulting libraries were sequenced using a Rapid-Run on an Illumina HiSeq 2500 to obtain 100 bp paired-end reads. Sequencing statistics are provided in **Table S3**. Metagenomic sequences were deposited into NCBI SAMN07256342-07256348 (**PRJNA390944**).

Metagenome assembly, binning, and annotation. Demultiplexed Illumina reads were mapped to known adapters using Bowtie2 in local mode to remove any reads with adapter contamination.

Demultiplexed Illumina read pairs were quality trimmed with Trim Galore (Babraham Bioinformatics) using a base Phred33 score threshold of Q25 and a minimum length cutoff of 80 bp. Paired-end reads were then assembled into contigs using SPAdes assembler with --meta option for assembling metagenomes, iterating over a range of k-mer values (21,27,33,37,43,47,51,55,61,65,71,75,81,85,91,95). Assemblies were assessed with reports generated with QUAST. Features on contigs were predicted through the Prokka pipeline with RNAmmer for rRNA, Aragorn for tRNA, Infernal and Rfam for other non-coding RNA and Prodigal for protein coding genes. Metagenomic 16S rRNA sequences were analyzed by BLASTN analysis against the Greengenes reference database. Matches with a bit score above 50 and reads matching multiple reference genes with the highest bit score were retained for comparison with 16S rRNA amplicons (**Fig. S3).** Annotation of protein-coding genes was performed as follows: 1) BLASTP search against the default set of core genomes, followed by HMM search against a set of default core HMM profiles available in Prokka, 2) use of the BLAST Descriptor Annotator algorithm in BLAST2GO, which conducts BLAST against the NCBI nr database, 3) KEGG orthology assignment using GhostKoala and 4) InterProScan analysis, which involves cross-reference HMM searches across multiple databases to find Pfam families with close homology.

Metagenome contigs were partitioned through MetaBAT (Kang et al., 2015) into metagenome-assembled genomes (MAGs) using tetranucleotide frequency and sequencing depth. Sequencing depth was estimated by mapping reads on to assembled contigs using Bowtie2 and Samtools. Completeness, contamination and strain level heterogeneity were assessed using single copy marker genes in CheckM (Parks et al., 2015). Gene features and their functional annotations for genome bins were extracted from the metagenome for the contigs that belong to the bins. Initial taxonomic affiliation for bins was inferred via the least common ancestor (LCA) algorithm in MEGAN6 and by the top BLAST matches to the marker gene *rpoB*. The B2 MAG was deposited into Genbank as "Candidatus *Atribacteria* bacterium 1244-E10-H5-B2" (SAMN07342547; **NMQN00000000.1**).

Phylogeny reconstruction for MAGS. Coding sequences from whole genomes were downloaded from the NCBI representative genomes collection using NCBI e-utilities, comprising 405 genomes in total, spanning all bacterial lineages. Only one candidate per genus with more than 1000 genes and maximum isolate information available was selected for this purpose. Sequence duplication (100% identity, unlikely to be biological duplication) within genomes was removed using CD-HIT. Available reference *Atribacteria* genomes, 24 in total, as either single-cell amplified genomes (SAGs) or MAGs, were downloaded and annotated using the Prokka pipeline. A list of 139 core single copy genes (CSCG) as HMM profiles was obtained from Rinke et al. (2013). B2 and representative reference *Atribacteria* genomes were then scanned for the presence of these HMM profiles using HMMer with the recommended score threshold for each profile as provided in Rinke et al. (2013). In a series of manual subsampling steps, 69 CSCG clusters were selected in 220 representative genomes and 20 *Atribacteria* genomes where 1) 69 clusters were present in only a single copy, 2) all 69 clusters were present in 220 representative genomes and 3) the minimum number of clusters present in any *Atribacteria* genome was 6. All 69 CSCG clusters were aligned individually using the L-INS-i mode in MAFFT. Alignments were then concatenated using a custom script Aln.cat.rb from the Enveomics collection (link) with invariable sites removed. Phylogeny reconstruction was performed in RAxML using a GAMMA model of rate heterogeneity, iterating over all models of protein substitution to choose the one with best log likelihood. The analysis was performed with 1000 rapid bootstraps with the MRE convergence bootstrap criterion (50 bootstrap replicates performed), followed by a thorough ML search. The resulting phylogenetic tree was modified for optimal viewing in iTOL with a full view including all lineages and a pruned view confirming placement of MAG B2 in the *Atribacteria* phylogeny. *Atribacteria* taxonomic classifications were based on Yarza et al. (2014). To examine gene orthology between B2 and other reference *Atribacteria*, 23 reference

Atribacteria (MAGs and SAGs) genomes were annotated using Prokka. The predicted genes were analyzed by BLAST best hit (BBH) clustering for orthologous group identification through Proteinortho5. In B2, 55% of genes (2333/4254) lacked orthologs in other *Atribacteria* genomes.

Metaproteomic sample preparation, mass spectrometry, and data analyses. Proteins from E10-H5 were extracted from a 10 g of frozen sediment using a protocol adapted from Nicora et al. (2013). Briefly, 2.5 mL of desorption buffer (0.5 M NaCl, 0.1 M glycerol, 0.2% SDS, 6 M urea, 1 mM EDTA, 100 mM ammonium bicarbonate) and 2 mL of a pH-buffered amino acid solution (containing equimolar histidine, lysine, and arginine, all 83 g 1 L^{-1} in ultra-pure water, pH 7.0) was added to the sample on ice. The goal of the pH-buffered amino acid solution is to fill the electronegative mineral sites in the sample with positively charged amino acids to reduce absorption of proteins to the particles. Samples were vortexed 4x, alternating 5 minutes vortexing and 5 min ice. The sediment slurry was then sonicated with Bronson probe sonicator (4 x 30 s) to lyse cells and heated at 95°C for 5 min. The sediment was pelleted by centrifugation (10,000 x g, 30 min, 4°C), and the supernatant was collected and stored on ice. The sediment pellet was washed 2 more times with 3 mL desorption buffer and supernatants were combined. In order to remove the SDS prior to protein digestion and mass spectrometry analysis, the filter aided sample preparation (FASP) method was used (Ostasiewicz et al., 2010). Millipore Amicon 10 kDa filter units were used and cleaned following manufacturer's directions. Samples were loaded on top of filters (~9 mL) and centrifuged (3000 rpm, 90 min, 4°C). In order to remove all SDS, proteins retained on the filter were rinsed 3 times by adding 5 mL of 8 M urea in 50 mM ammonium bicarbonate and repeating the prior centrifugation step. Iodoacetamide (3 mL, 15 mM) was added to samples, incubated in the dark at room temperature for 30 minutes, and then centrifuged (3000 rpm, 90 min, 4°C). Proteins were then rinsed two times with 10 mL of 100 mM ammonium bicarbonate and centrifuged to remove liquid (3000 rpm, 90 min, 4°C). To digest protein on the filter, 0.5 µg of trypsin (modified, sequencing grade, Promega) was added to the filter, topped with 2.5 mL of 25 mM ammonium bicarbonate, vortexed, and incubated 12 hr at room temperature. Filtrate was collected by centrifugation (3000 rpm, 90 min, 4°C), and SpeedVaced to near dryness at 4°C. Peptides were then resuspended in 50 µL of 2% acetonitrile and 0.1% formic acid and desalted using Nest Group C18 Proto centrifugal macro columns following manufacturer's instructions. Each 10 µL sample was separated on a NanoAquity UPLC with a 60 min gradient (2-35% acetonitrile) and analyzed on a Thermo Scientific Orbitrap Fusion Tribrid Mass Spectrometer operated in top20 data dependent acquisition mode.

A protein database for identifying the collected fragmentation spectra was generated from *Atribacteria* MAGs (C1H2_C3H4ab_E10H5_contam.fasta).These databases were concatenated with 50 common contaminants, yielding a protein database of 10,325 proteins. To assign spectra to peptide sequences, correlative database searches were completed using Comet v. 2015.01 rev. 2 (Eng et al., 2013; Eng et al., 2015). Comet parameters included: trypsin enzyme specificity, semi-digested, allowance of 1 missed cleavage, 10 ppm mass tolerance, cysteine modification of 57 Da (resulting from the iodoacetamide) and modifications on methionine of 15.999 Da (oxidation). Minimum protein and peptide thresholds were set at $P > 0.95$ on Protein and Peptide Prophet (Nesvizhskii et al., 2003). Protein inferences from the whole-cell lysates were accepted by ProteinProphet if the thresholds noted above were passed, two or more peptides were identified, and at least one terminus was tryptic (Keller et al., 2002; Nesvizhskii et al., 2003; Pedrioli, 2010). For each peptide discussed in the manuscript, manual inspection of the spectral identification was completed. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (Vizcaíno et al., 2015) with the dataset identifier **PXD01247** (https://www.ebi.ac.uk/pride/archive/ Login: reviewer08969@ebi.ac.uk Password: BP2V3yGA).

525 **References**

526 Anantharaman, K., Brown, C.T., Hug, L.A., Sharon, I., Castelle, C.J., Probst, A.J. et al.
527 (2016) Thousands of microbial genomes shed light on interconnected biogeochemical
528 processes in an aquifer system. *Nat Commun* **7**: 13219.

529 Archer, D., Buffett, B., and Brovkin, V. (2009) Ocean methane hydrates as a slow tipping
530 point in the global carbon cycle. *Proc Natl Acad Sci* **106**: 20596-20601.

531 Axen, S.D., Erbilgin, O., and Kerfeld, C.A. (2014) A taxonomy of bacterial
532 microcompartment loci constructed by a novel scoring method. *PLoS Comput Biol* **10**:
533 e1003898.

534 Baker, B.J., Lazar, C.S., Teske, A.P., and Dick, G.J. (2015) Genomic resolution of linkages
535 in carbon, nitrogen, and sulfur cycling among widespread estuary sediment bacteria.
536 *Microbiome* **3**: 14.

537 Barrick, J.E., Corbino, K.A., Winkler, W.C., Nahvi, A., Mandal, M., Collins, J. et al. (2004)
538 New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control.
539 *Proc Natl Acad Sci* **101**: 6421-6426.

540 Bergkessel, M., Basta, D.W., and Newman, D.K. (2016) The physiology of growth arrest:
541 uniting molecular and environmental microbiology. *Nature Reviews Microbiology* **14**: 549.

542 Bohrmann, G., and Torres, M.E. (2006) Gas Hydrates in Marine Sediments. In *Marine*
543 *Geochemistry*. Schulz, H.D., and Zabel, M. (eds). Berlin, Heidelberg: Springer, pp. 481-
544 512.

545 Braun, M., and Silhavy, T.J. (2002) Imp/OstA is required for cell envelope biogenesis in
546 *Escherichia coli*. *Mol Microbiol* **45**: 1289-1302.

547 Byrne, M.D., and Nicholas, D. (1986) Multiple-phase equilibration headspace analysis for
548 the determination of N2O and N2 during bacterial denitrification. *Anal Biochem* **154**: 470-
549 475.

550 Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh,
551 P.J. et al. (2011) Global patterns of 16S rRNA diversity at a depth of millions of sequences
552 per sample. *Proc Natl Acad Sci* **108**: 4516-4522.

553 Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K.
554 et al. (2010) QIIME allows analysis of high-throughput community sequencing data.
555 *Nature Methods* **7**: 335-336.

556 Carr, S.A., Orcutt, B.N., Mandernack, K.W., and Spear, J.R. (2015) Abundant Atribacteria
557 in deep marine sediment from the Adelie Basin, Antarctica. *Front Microbiol* **6**: 872.

558 Chernitsyna, S., Mamaeva, E., Lomakina, A., Pogodaeva, T., Galach'yants, Y.P., Bukin,
559 S. et al. (2016) Phylogenetic diversity of microbial communities of the Posolsk Bank
560 bottom sediments, Lake Baikal. *Microbiology* **85**: 672-680.

561 Chong, Z.R., Yang, S.H.B., Babu, P., Linga, P., and Li, X.-S. (2016) Review of natural gas
562 hydrates as an energy resource: Prospects and challenges. *Appl Energ* **162**: 1633-1652.

563 Daly, R.A., Borton, M.A., Wilkins, M.J., Hoyt, D.W., Kountz, D.J., Wolfe, R.A. et al.
564 (2016) Microbial metabolisms in a 2.5-km-deep ecosystem created by hydraulic fracturing
565 in shales. *Nat Microbiol* **1**: 16146.

566 Dodsworth, J.A., Blainey, P.C., Murugapiran, S.K., Swingley, W.D., Ross, C.A., Tringe,
567 S.G. et al. (2013) Single-cell and metagenomic analyses indicate a fermentative and
568 saccharolytic lifestyle for members of the OP9 lineage. *Nat Commun* **4**: 1854.

569 Dombrowski, N., Seitz, K.W., Teske, A.P., and Baker, B.J. (2017) Genomic insights into
570 potential interdependencies in microbial hydrocarbon and nutrient cycling in hydrothermal
571 sediments. *Microbiome* **5**: 106.

572 Eng, J.K., Jahan, T.A., and Hoopmann, M.R. (2013) Comet: an open‐source MS/MS
573 sequence database search tool. *Proteomics* **13**: 22-24.

574 Eng, J.K., Hoopmann, M.R., Jahan, T.A., Egertson, J.D., Noble, W.S., and MacCoss, M.J.
575 (2015) A deeper look into Comet—implementation and features. *J Am Soc Mass Spec* **26**:
576 1865-1874.

577 Engelberg-Kulka, H., Hazan, R., and Amitai, S. (2005) mazEF: a chromosomal toxin-
578 antitoxin module that triggers programmed cell death in bacteria. *J Cell Sci* **118**: 4327-
579 4332.

580 Faith, D.P. (1992) Conservation evaluation and phylogenetic diversity. *Biol Conserv* **61**:
581 1-10.

582 Fischer, M., Zhang, Q.Y., Hubbard, R.E., and Thomas, G.H. (2010) Caught in a TRAP:
583 substrate-binding proteins in secondary transport. *Trends Microbiol* **18**: 471-478.

584 Fischer, M., Hopkins, A.P., Severi, E., Hawkhead, J., Bawdon, D., Watts, A.G. et al. (2015)
585 Tripartite ATP-independent periplasmic (TRAP) transporters use an arginine-mediated
586 selectivity filter for high affinity substrate binding. *J Biol Chem* **290**: 27113-27123.

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

590 Fry, J.C., Parkes, R.J., Cragg, B.A., Weightman, A.J., and Webster, G. (2008) Prokaryotic
591 biodiversity and activity in the deep subseafloor biosphere. *FEMS Microbiol Ecol* **66**: 181-
592 196.

593 Gies, E.A., Konwar, K.M., Beatty, J.T., and Hallam, S.J. (2014) Illuminating microbial
594 dark matter in meromictic Sakinaw Lake. *Appl Environ Microbiol* **80**: 6807-6818.

595 Hernsdorf, A.W., Amano, Y., Miyakawa, K., Ise, K., Suzuki, Y., Anantharaman, K. et al.
596 (2017) Potential for microbial H2 and metal transformations associated with novel bacteria
597 and archaea in deep terrestrial subsurface sediments. *ISME J* **11**: 1915-1929.

598 Hester, K.C., and Brewer, P.G. (2009) Clathrate hydrates in nature. *Ann Rev Mar Sci* **1**:
599 303-327.

600 Honkalas, V., Dabir, A., and Dhakephalkar, P.K. (2016) Life in the Anoxic Sub-Seafloor
601 Environment: Linking Microbial Metabolism and Mega Reserves of Methane Hydrate. In
602 *Anaerobes in Biotechnology*. Hatti-Kaul, R., Mamo, G., and Mattiasson, B. (eds). Cham:
603 Springer, pp. 235-262.

604 Hoshino, T., Toki, T., Ijiri, A., Morono, Y., Machiyama, H., Ashi, J. et al. (2017)
605 Atribacteria from the subseafloor sedimentary biosphere disperse to the hydrosphere
606 through submarine mud volcanoes. *Front Microbiol* **8**: 1135.

607 Hu, P., Tom, L., Singh, A., Thomas, B.C., Baker, B.J., Piceno, Y.M. et al. (2016) Genome-
608 resolved metagenomic analysis reveals roles for candidate phyla and other microbial
609 community members in biogeochemical transformations in oil reservoirs. *mBio* **7**: e01669-
610 01615.

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

614 Inagaki, F., Nunoura, T., Nakagawa, T., Teske, A., Lever, M.A., Lauer, A. et al. (2006)
615 Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep
616 marine sediments on the Pacific Ocean Margin. *Proc Natl Acad Sci* **103**: 2815–2820.

617 Kadnikov, V.V., Mardanov, A.V., Beletsky, A.V., Shubenkova, O.V., Pogodaeva, T.V.,
618 Zemskaya, T.I. et al. (2012) Microbial community structure in methane hydrate-bearing
619 sediments of freshwater Lake Baikal. *FEMS Microbiol Ecol* **79**: 348-358.

620 Kang, D.D., Froula, J., Egan, R., and Wang, Z. (2015) MetaBAT, an efficient tool for
621 accurately reconstructing single genomes from complex microbial communities. *PeerJ* **3**:
622 e1165.

623 Keller, A., Purvine, S., Nesvizhskii, A.I., Stolyar, S., Goodlett, D.R., and Kolker, E. (2002)
624 Experimental protein mixture for validating tandem mass spectral analysis. *OMICS* **6**: 207-
625 212.

626 Kho, K., and Meredith, T.C. (2018) Salt-induced stress stimulates a lipoteichoic acid-
627 specific three component glycosylation system in *Staphylococcus aureus*. *J Bacteriol* **200**:
628 e00017-00018.

629 Kormas, K.A., Smith, D.C., Edgcomb, V., and Teske, A. (2003) Molecular analysis of deep
630 subsurface microbial communities in Nankai Trough sediments (ODP Leg 190, Site 1176).
631 *FEMS Microbiol Ecol* **45**: 115-125.

632 Kozich, J.J., Westcott, S.L., Baxter, N.T., Highlander, S.K., and Schloss, P.D. (2013)
633 Development of a dual-index sequencing strategy and curation pipeline for analyzing
634 amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl Environ*
635 *Microbiol* **79**: 5112-5120.

636 Kumar, A., Bhattacharjee, G., Kulkarni, B., and Kumar, R. (2015) Role of surfactants in
637 promoting gas hydrate formation. *Ind Eng Chem Res* **54**: 12217-12232.

638 Lanoil, B.D., Sassen, R., La Duc, M.T., Sweet, S.T., and Nealson, K.H. (2001) Bacteria
639 and Archaea physically associated with Gulf of Mexico gas hydrates. *Appl Environ*
640 *Microbiol* **67**: 5143-5153.

641 Lecher, J., Pittelkow, M., Zobel, S., Bursy, J., Bönig, T., Smits, S.H. et al. (2009) The
642 crystal structure of UehA in complex with ectoine—a comparison with other TRAP-T
643 binding proteins. *J Mol Biol* **389**: 58-73.

644 Lin, L.-H., Wu, L.-W., Cheng, T.-W., Tu, W.-X., Lin, J.-R., Yang, T.F. et al. (2014)
645 Distributions and assemblages of microbial communities along a sediment core retrieved
646 from a potential hydrate-bearing region offshore southwestern Taiwan. *Journal of Asian*
647 *Earth Sciences* **92**: 276-292.

648 Liu, Y., Chen, B., Chen, Y., Zhang, S., Guo, W., Cai, Y. et al. (2015) Methane storage in
649 a hydrated form as promoted by leucines for possible application to natural gas
650 transportation and storage. *Energy Technol* **3**: 815-819.

651 Liu, Y., Gao, W., Wang, Y., Wu, L., Liu, X., Yan, T. et al. (2005) Transcriptome analysis
652 of *Shewanella oneidensis* MR-1 in response to elevated salt conditions. *J Bacteriol* **187**:
653 2501-2507.

654 Marlow, J.J., Steele, J.A., Case, D.H., Connon, S.A., Levin, L.A., and Orphan, V.J. (2014)
655 Microbial abundance and diversity patterns associated with sediments and carbonates from
656 the methane seep environments of Hydrate Ridge, OR. *Front Mar Sci* **1**: 44.

657 Mathiesen, C., and Hägerhäll, C. (2002) Transmembrane topology of the NuoL, M and N
658 subunits of NADH: quinone oxidoreductase and their homologues among membrane-
659 bound hydrogenases and bona fide antiporters. *BBA-Bioenergetics* **1556**: 121-132.

660 Mavromatis, K., Ivanova, N., Anderson, I., Lykidis, A., Hooper, S.D., Sun, H. et al. (2009)
661 Genome analysis of the anaerobic thermohalophilic bacterium *Halothermothrix orenii*.
662 *PLoS One* **4**: e4192.

663 McLaughlin, K.J., Strain-Damerell, C.M., Xie, K., Brekasis, D., Soares, A.S., Paget, M.S.,
664 and Kielkopf, C.L. (2010) Structural basis for NADH/NAD+ redox sensing by a Rex
665 family repressor. *Mol Cell* **38**: 563-575.

666 Milkov, A.V., Dickens, G.R., Claypool, G.E., Lee, Y.-J., Borowski, W.S., Torres, M.E. et
667 al. (2004) Co-existence of gas hydrate, free gas, and brine within the regional gas hydrate
668 stability zone at Hydrate Ridge (Oregon margin): evidence from prolonged degassing of a
669 pressurized core. *Earth Planet Sci Lett* **222**: 829-843.

670 Moparthi, V.K., and Hägerhäll, C. (2011) The evolution of respiratory chain complex I
671 from a smaller last common ancestor consisting of 11 protein subunits. *J Mol Evol* **72**: 484-
672 497.

673 Moparthi, V.K., Kumar, B., Al-Eryani, Y., Sperling, E., Górecki, K., Drakenberg, T., and
674 Hägerhäll, C. (2014) Functional role of the MrpA-and MrpD-homologous protein subunits
675 in enzyme complexes evolutionary related to respiratory chain complex I. *BBA-*
676 *Bioenergetics* **1837**: 178-185.

677 Nesvizhskii, A.I., Keller, A., Kolker, E., and Aebersold, R. (2003) A statistical model for
678 identifying proteins by tandem mass spectrometry. *Anal Chem* **75**: 4646-4658.

679 Newberry, C.J., Webster, G., Cragg, B.A., Parkes, R.J., Weightman, A.J., and Fry, J.C.
680 (2004) Diversity of prokaryotes and methanogenesis in deep subsurface sediments from
681 the Nankai Trough, Ocean Drilling Program Leg 190. *Environ Microbiol* **6**: 274-287.

682 Nicora, C.D., Anderson, B.J., Callister, S.J., Norbeck, A.D., Purvine, S.O., Jansson, J.K.
683 et al. (2013) Amino acid treatment enhances protein recovery from sediment and soils for
684 metaproteomic studies. *Proteomics* **13**: 2776-2785.

685 Nobu, M.K., Narihiro, T., Rinke, C., Kamagata, Y., Tringe, S.G., Woyke, T., and Liu, W.-
686 T. (2015) Microbial dark matter ecogenomics reveals complex synergistic networks in a
687 methanogenic bioreactor. *ISME J* **9**: 1710-1722.

688 Nobu, M.K., Dodsworth, J.A., Murugapiran, S.K., Rinke, C., Gies, E.A., Webster, G. et al.
689 (2016) Phylogeny and physiology of candidate phylum 'Atribacteria' (OP9/JS1) inferred
690 from cultivation-independent genomics. *ISME J* **10**: 273-286.

691 Ostasiewicz, P., Zielinska, D.F., Mann, M., and Wiśniewski, J.R. (2010) Proteome,
692 phosphoproteome, and N-glycoproteome are quantitatively preserved in formalin-fixed
693 paraffin-embedded tissue and analyzable by high-resolution mass spectrometry. *J Prot Res*
694 **9**: 3688-3700.

695 Parkes, R.J., Cragg, B., Roussel, E., Webster, G., Weightman, A., and Sass, H. (2014) A
696 review of prokaryotic populations and processes in sub-seafloor sediments, including
697 biosphere: geosphere interactions. *Mar Geol* **352**: 409-425.

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

701 Pedrioli, P.G. (2010) Trans-Proteomic Pipeline: A Pipeline for Proteomic Analysis. In
702 *Proteome Bioinformatics*. Hubbard, S., and Jones, A. (eds): Humana Press, pp. 213-238.

703 Probst, A.J., Castelle, C.J., Singh, A., Brown, C.T., Anantharaman, K., Sharon, I. et al.
704 (2017) Genomic resolution of a cold subsurface aquifer community provides metabolic
705 insights for novel microbes adapted to high CO2 concentrations. *Environ Microbiol* **19**:
706 459-474.

Reed, D.W., Fujita, Y., Delwiche, M.E., Blackwelder, D.B., Sheridan, P.P., Uchida, T., and Colwell, F.S. (2002) Microbial communities from methane hydrate-bearing deep marine sediments in a forearc basin. *Appl Environ Microbiol* **68**: 3759-3770.

Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N.N., Anderson, I.J., Cheng, J.-F. et al. (2013) Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **499**: 431-437.

Roy, M., McManus, J., Goni, M.A., Chase, Z., Borgeld, J.C., Wheatcroft, R.A. et al. (2013) Reactive iron and manganese distributions in seabed sediments near small mountainous rivers off Oregon and California (USA). *Cont Shelf Res* **54**: 67-79.

Ruff, S., Felden, J., Gruber-Vodicka, H., Marcon, Y., Knittel, K., Ramette, A., and Boetius, A. (2019) In situ development of a methanotrophic microbiome in deep-sea sediments. *The ISME journal* **13**: 197.

Ruppel, C.D., and Kessler, J.D. (2017) The interaction of climate change and methane hydrates. *Rev Geophys* **55**: 126-168.

Sa, J.-H., Kwak, G.-H., Lee, B.R., Park, D.-H., Han, K., and Lee, K.-H. (2013) Hydrophobic amino acids as a new class of kinetic inhibitors for gas hydrate formation. *Sci Rep* **3**: 2428.

Santos, H., and Da Costa, M.S. (2002) Compatible solutes of organisms that live in hot saline environments. *Environ Microbiol* **4**: 501-509.

Schut, G.J., Boyd, E.S., Peters, J.W., and Adams, M.W. (2013) The modular respiratory complexes involved in hydrogen and sulfur metabolism by heterotrophic hyperthermophilic archaea and their evolutionary implications. *FEMS Microbiol Rev* **37**: 182-203.

ShipboardScientificParty (2003) Site 1244. In *Proc ODP, Init Repts, 204*. Tréhu, A., Bohrmann, G., Rack, F., Torres, M., and al., e. (eds). College Station, TX: Ocean Drilling Program, pp. 1–132 doi:110.2973/odp.proc.ir.2204.2103.2003.

Spero, M.A., Aylward, F.O., Currie, C.R., and Donohue, T.J. (2015) Phylogenomic analysis and predicted physiological role of the proton-translocating NADH: quinone oxidoreductase (complex I) across bacteria. *mBio* **6**: e00389-00315.

Tréhu, A., Bohrmann, G., Rack, F., and Torres, M. (2003) Volume 204 Initial Reports. In *Proc ODP, Initial Reports*, pp. 77845-79547.

Ussler III, W., and Paull, C.K. (2001) Ion exclusion associated with marine gas hydrate deposits. In *Natural Gas Hydrates: Occurrence, Distribution, and Detection*. Paull, C.K., and Dillion, W.P. (eds): American Geophysical Union.

Veluswamy, H.P., Lee, P.Y., Premasinghe, K., and Linga, P. (2017) Effect of biofriendly amino acids on the kinetics of methane hydrate formation and dissociation. *Ind Eng Chem Res* **56**: 6145-6154.

744 Vizcaíno, J.A., Csordas, A., Del-Toro, N., Dianes, J.A., Griss, J., Lavidas, I. et al. (2015)
745 2016 update of the PRIDE database and its related tools. *Nuc Acid Res* **44**: D447-D456.

746 Webster, G., Parkes, R.J., Fry, J.C., and Weightman, A.J. (2004) Widespread occurrence
747 of a novel division of bacteria identified by 16S rRNA gene sequences originally found in
748 deep marine sediments. *Appl Environ Microbiol* **70**: 5708-5713.

749 Webster, G., Yarram, L., Freese, E., Koster, J., Sass, H., Parkes, R.J., and Weightman, A.J.
750 (2007) Distribution of candidate division JS1 and other Bacteria in tidal sediments of the
751 German Wadden Sea using targeted 16S rRNA gene PCR-DGGE. *FEMS Microbiol Ecol*
752 **62**: 78-89.

753 Wood, J.M. (2015) Bacterial responses to osmotic challenges. *J Gen Physiol* **145**: 381-388.

754 Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F.O., Ludwig, W., Schleifer, K.-H. et al.
755 (2014) Uniting the classification of cultured and uncultured bacteria and archaea using 16S
756 rRNA gene sequences. *Nat Rev Microbiol* **12**: 635-645.

757 Yu, H., Wu, C.-H., Schut, G.J., Haja, D.K., Zhao, G., Peters, J.W. et al. (2018) Structure
758 of an ancient respiratory system. *Cell* **173**: 1636-1649.
759