Tracking active heterotrophic microbial communities in the Guaymas Basin deep biosphere with BONCAT-FACS

Andrew Montgomery[1,2,3,*,#], Sylvia Nupp[1,2,3,*], Chase R. Gray[1,2,3], Zackary J. Jay[1,2,3], Virginia Edgcomb[4], Roland Hatzenpichler[1,2,3,5,#]

1. Department of Chemistry and Biochemistry, Montana State University, Bozeman, MT, USA
2. Thermal Biology Institute, Montana State University, Bozeman, MT, USA
3. Center for Biofilm Engineering, Montana State University, Bozeman, MT, USA
4. Woods Hole Oceanographic Institution, Woods Hole, MA, USA
5. Department of Microbiology and Cell Biology, Montana State University, Bozeman, MT, USA

*These authors contributed to this work equally.

#correspondence to: R.H., rolandhatzenpichler@gmail.com; A.M., andymontgomery2009@gmail.com

ORCIDs: A.M., 0000-0003-2798-1390; S.N., 0009-0008-1814-1211; R.H., 0000-0002-5489-3444; Z.J.J., 0000-0003-3062-4933; C.R.G., 0009-0005-0296-1491; V.E., https://orcid.org/0000-0001-6805-381X.

Abstract

The marine deep biosphere harbors microbial communities that drive organic matter transformations and biogeochemical cycles. Previous work on these communities has focused either on genomic characterization or metabolic activity measurements. However, to understand microbial ecophysiology in the deep biosphere taxonomic identity and metabolic function must be connected on both single-cell and ecosystem scales. In this work, we optimized a bioorthogonal non-canonical amino acid tagging fluorescence-activated cell sorting (BONCAT-FACS) workflow for low-biomass deep-biosphere sediments obtained during International Ocean Discovery Program Expedition 385 (IODP 385). BONCAT-FACS with 16S rRNA gene amplicon sequencing as well as metagenomics of sediment communities was applied to characterize translationally active communities in hydrothermally altered subsurface sediments of the Guaymas Basin. Our results revealed a heterotrophic microbial population throughout all sediments examined, with taxa translationally active down to our deepest sampling point, 154 meters below the seafloor. Based on 16S rRNA gene identities, the translationally active microbial community was dominated by heterotrophic members of the *Gammaproteobacteria*, *Bacilli*, *Deinococci*, and *Alphaproteobacteria*. These taxa are likely key contributors to cycling the large quantities of hydrothermally altered organic matter in Guaymas Basin sediments. To further elucidate the metabolic capacity of active taxa, we mapped 16S rRNA gene amplicons to metagenome assembled genomes (MAGs) previously obtained from IODP 385. These MAGs contained genes associated with C1 metabolism, carbohydrate degradation, and fermentation, indicating that active taxa leverage these metabolisms for energy conservation. Our results demonstrate that BONCAT-FACS provides high-throughput and single-cell insights into the metabolic activity of microbes in the low-biomass marine subsurface.

43 **Importance**

44 The marine deep biosphere spans the entire global ocean and offers a wide range of environmental
45 conditions for microbes to flourish. However, the identities and metabolic functions of these
46 microbial populations are not well understood. In this study, we investigated translationally active
47 microbial communities and elucidated their metabolic activity in hydrothermally impacted low-
48 biomass subsurface marine sediments. By combining activity measurements, amplicon
49 sequencing, and metagenomics data, we improve our understanding of carbon transformations and
50 microbial ecology of the deep biosphere. Our findings emphasize the ecological significance of
51 heterotrophic microbes and the carbon-degrading metabolisms they use to survive extreme
52 conditions. This work advances the applications of next-generation physiology techniques to
53 document single-cell metabolic activity in a high-throughput manner in the marine deep biosphere.

54 **Introduction**

55 The marine deep biosphere is one of the largest ecosystems on Earth, supporting an estimated 10^{29}
56 cells (1, 2). Defined as sediments >1 m below the seafloor, the deep biosphere encompasses a wide
57 spectrum of redox conditions that extend to the energetic and thermal limits of life. Biological and
58 geological transformations of organic matter drive a dynamic, albeit slow, carbon cycle in these
59 sub-seafloor environments (3–6). While it is clear that microorganisms are metabolically active in
60 the deep biosphere (7, 8), low cell abundances mean the significance of their activity and
61 contributions to biogeochemical cycles remains largely unconstrained. Direct connections between
62 the taxonomic identity of active microbial populations and metabolic functions are necessary to
63 understand microbial ecology and biogeochemical cycling in the marine deep biosphere.

64

65 Uncovering the metabolic capabilities of constituents of microbial communities *in situ* requires a
66 coordinated effort that simultaneously identifies taxa and links them to ecosystem functions.
67 Historically, cultured isolates were required in order to draw conclusions about the metabolic
68 capabilities of specific microbial taxa. However, less than 1% of microbes can be reliably
69 maintained in a laboratory setting, and enrichment and isolation conditions can differ substantially
70 from *in situ* conditions, making it difficult to draw conclusions about how microbes live in their
71 natural habitat. Further, cultivation efforts are impeded by low biomass inocula and energy
72 limitations in the deep biosphere that select for microorganisms that grow at extremely slow rates
73 (9, 10). More recently, culture-independent techniques have been used to elucidate connections
74 between individual microorganisms and their metabolic functions under (near) *in situ* conditions.
75 With the rise of high-throughput sequencing technologies and bioinformatics tools, culture-
76 independent techniques have facilitated studies of the deep biosphere with metagenomics (11–13)

77 to identify microbial taxa, or metatranscriptomics (14–18) to reveal which genes are expressed *in*
78 *situ*. However, while these approaches offer powerful insights into metabolic potential and
79 community level activity, they do not link specific taxa with metabolic activity *in situ*.
80
81 Bioorthogonal non-canonical amino acid tagging (BONCAT) paired with fluorescence-activated
82 cell sorting (BONCAT-FACS) is a Next-Generation Physiology approach (19) that can be used to
83 identify which taxa were translationally active in complex environmental samples during a
84 relatively short incubation without culturing individual organisms (20–24). BONCAT leverages
85 the incorporation of a synthetic amino acid, such as *L*-homopropargylglycine (HPG), an analog of
86 *L*-methionine, into newly synthesized proteins and its later detection using azide-alkyne click
87 chemistry (20, 25). Using fluorescence as the mode of detection, FACS can be used to sort
88 translationally active cells for downstream genomic or amplicon sequencing and identification of
89 BONCAT-active taxa (21) and even for cultivation efforts (20, 26). BONCAT-FACS, coupled
90 with 16S rRNA gene amplicon sequencing, offers a powerful approach to identify anabolically
91 active taxa on a community level. This approach has been used to identify active taxa in marine
92 sediments (27), soils (24, 28), the human microbiome (29–31), and other complex environments
93 (32–35). Here, we used BONCAT-FACS to identify the translationally active community and,
94 coupled with community shotgun metagenomics, assess the metabolic capacity of translationally
95 active taxa in the Guaymas Basin deep biosphere.
96
97 Guaymas Basin, located in the Gulf of California, is a hydrothermally active rift basin
98 characterized by rapid seafloor spreading and high sedimentation rates. Terrestrial inputs and
99 highly productive waters in the Gulf of California deposit organic matter on the seafloor at rates

100 >1 mm yr^{-1}, generating sediment packages that can be hundreds of meters thick (36).
101 Thermocatalytic alteration of the organic matter, influenced by basaltic sill intrusions and
102 hydrothermal circulation (37, 38), produces a complex mixture of potentially bioavailable organic
103 compounds in the sediment (39). This includes methane, low-molecular-weight organics,
104 hydrocarbons, and crude oil (40–45). International Ocean Discovery Program Expedition 385
105 (IODP 385) drilled the subsurface of the Guaymas Basin to investigate how microbial communities
106 drive organic carbon cycling in the hydrothermally altered deep biosphere. In total, holes were
107 drilled at eight sites (U1545-U1552) to capture unaltered background sediment and a spectrum of
108 hydrothermal sediments including extremely hot sediments above a more recently intruded sill-
109 sediment system (46). To explore the microbial ecology of the Guaymas Basin subsurface, we
110 tracked translationally active taxa from six depths from five sites drilled during IODP 385. We
111 optimized a previously established BONCAT-FACS workflow with 16S rRNA gene amplicon
112 sequencing to identify translationally active taxa (27, 32) and to assess diversity across
113 biogeochemical gradients. We coupled this with metagenomic analyses from IODP 385 (12) to
114 provide direct insight into the metabolic capacity of active taxa within the Guaymas Basin deep
115 biosphere.
116
117 **Materials and Methods**
118 *Sample collection and incubation*
119 Sediment cores were collected from Guaymas Basin during IODP expedition 385 onboard drilling
120 vessel *JOIDES Resolution.* Samples were taken at six depths, ranging from 0.76 to 154 mbsf, from
121 five drill holes (Table 1). Intact, round samples were aseptically transferred to the microbiology
122 lab after retrieval on deck and stored at 4°C in heat-sealed gas-tight bags to maintain anoxic

123 conditions until processing. Working in an anoxic glove chamber, sediments were homogenized
124 with N_2-gassed sterile bottom water diluted three times to form a slurry. For each sample, four
125 replicate aliquots of 40 or 120 mL, depending on depth (shallow vs. deep), were distributed into
126 Pyrex bottles and sealed with a butyl stopper and aluminum crimp. We increased the volume of
127 deeper samples during incubation to increase the likelihood of detecting enough active cells for
128 sorting. The amino acid analog *L*-homopropargylglycine (HPG; Vector Laboratories, previously
129 Click Chemistry Tools), a synthetic analog of *L*-methionine, was added to three replicates to a
130 final concentration of 50 µM. The fourth replicate, which served as a control, was not amended
131 with HPG. After HPG addition, samples were incubated at *in situ* temperature (Table 1) for seven
132 days.
133
134 After seven days, samples were washed with 1X PBS to terminate incubations and remove unused
135 HPG. A 5 mL subsample was frozen in a cryovial at -80°C for whole sediment DNA extraction.
136 The remaining sediment was transferred to 50 mL centrifuge tubes, amended with 10% glycerol-
137 TE (9:1 sediment to glycerol-TE), and stored at -80°C for subsequent steps of the BONCAT-FACS
138 workflow back at Montana State University.
139
140 *Cell Extractions*
141 Cells were extracted following a modified protocol (27, 47) that we optimized to extract cells from
142 low-biomass subsurface marine sediments. Samples preserved in glycerol-TE were thawed
143 overnight at 4°C and centrifuged at 3,200 × *g* for 30 min to pellet sediment. For each sample,
144 sediment was diluted 1:3 in a filter-sterilized 3.5% NaCl solution (0.2 µm pore size) and briefly
145 homogenized by vortex. From this slurry, 200 µL aliquots were further diluted in 100 µL of pure

146 methanol and 700 µL of filter-sterilized 3.5% NaCl solution. For each experimental replicate, eight
147 extraction replicates were prepared in 2 mL microcentrifuge tubes to ensure enough biomass was
148 extracted for downstream processing. Extraction blanks with phosphate-buffered saline (1X PBS)
149 were processed in parallel to assess contaminant introduction during extraction. Extraction
150 replicates were homogenized by vortex for 30 min at room temperature on setting 7 (approximately
151 2,000 rpm, VortexGenie2, Scientific Industries, Inc., Bohemia, NY, United States). Cells were
152 separated by density centrifugation by layering 1 mL of 50% Nycodenz solution below the
153 sediment slurry with a syringe and needle and centrifuging at 3,000 × *g* for 10 min. After
154 centrifugation, the upper and inner phases of the supernatant were collected by pipetting and
155 pooling extraction replicates in a 15 mL centrifuge tube. The remaining supernatant was discarded,
156 leaving only the sediment pellet. To detach cells from particles, the pellet was resuspended in 800
157 µL of 3.5% NaCl solution, 100 µL of pure methanol, and 100 µL of detergent mix (100 mM
158 ethylenediamine tetraacetic acid (EDTA), 100 mM sodium pyrophosphate, 1% Tween80 in 3.5%
159 NaCl solution (47), briefly vortexed to homogenize, and sonicated. The sonication step was done
160 on ice at 125 W and 200 µm amplitude for 5 sec, followed by a 20 sec pause and repeated for 5
161 cycles to minimize cell rupture. Subsequently, samples were subjected to the same Nycodenz
162 density gradient centrifugation described above. The upper and inner phases of the supernatant
163 from each extraction replicate were combined with the pooled supernatant from the first density
164 gradient centrifugation step. The pooled fractions were centrifuged at 16,000 × *g* for 10 min to
165 pellet extracted cells, and the supernatant was gently decanted, leaving ~1 mL of supernatant
166 remaining to resuspend pelleted cells. This cell suspension was transferred to a 1.5 mL
167 microcentrifuge tube, and cells were centrifuged again at 18,000 × *g* for 10 min and concentrated

168 in 100 µL of PBS:ethanol (1:1 ratio of 1X PBS and pure ethanol). Concentrated cells were stored
169 overnight at -20°C to permeabilize cells for click chemistry.
170
171 *Click Chemistry and Fluorescence-Activated Cell Sorting*
172 The following day, cell extracts were click-stained to fluorescently tag cells that incorporated HPG
173 into newly synthesized proteins (48). A bulk click solution was prepared, and 200 µL was added
174 to each cell extract. The final click reaction comprised of 500 µM tris-
175 hydroxypropyltriazolylmethylamine (THPTA; Vector Laboratories), 100 µM copper sulfate
176 pentahydrate, 5 mM sodium *L*-ascorbate, 5 mM aminoguanidine HCl, and 5 µM of Cy5 picolyl-
177 azide dye (Vector Laboratories) in 1X PBS. The click solution and cell extract were homogenized
178 and then incubated at room temperature, in the dark, for 30 min. The click solution was washed
179 away with three washes in 1 mL of 1X PBS with centrifugation at 18,000 × *g* for 5 min between
180 each wash. A 50 µL aliquot of the final suspension was kept aside to be used as a "presort" total
181 cell-extractable fraction for PCR. This subsample was stored in a fresh 1.5 mL microcentrifuge
182 tube at -80°C until freeze-thaw DNA extraction.
183
184 The remaining cell extract was immediately subjected to fluorescence-activated cell sorting
185 (FACS). Prior to loading, each sample was passed through a 35 µm mesh filter to remove sediment
186 particles and diluted to a final volume of ~1.5 mL with 1X PBS. FACS was performed with a Sony
187 SH800S (Sony Biotechnology, San Jose, CA) equipped with a 70 µm sorting chip using 1X PBS
188 as sheath fluid. The sheath fluid was 0.2 µm filtered and UV sterilized overnight while stirring to
189 remove salt crystals in the PBS and minimize contamination of the FACS fluidics system (32). A
190 series of gates was manually drawn to exclude large sediment particles. The final sorting gates

191 were drawn based on the Cy5 relative fluorescence and forward scatter (Fig. S1) of the No-HPG
192 controls. A "Cy5-positive" gate was drawn so that only BONCAT-positive cells that incorporated
193 HPG into newly synthesized proteins were sorted. Each sample was sorted until the sample had
194 been exhausted, regardless of the number of events sorted (Table S1). Sorted fractions were
195 centrifuged at 18,000 × *g* for 5 min to pellet cells. After removing the supernatant, cells were
196 resuspended in 50 µL of nuclease-free water and stored at -80°C for freeze-thaw DNA extraction.
197
198 *DNA extraction*
199 DNA from whole sediments was extracted for each replicate sample after HPG incubation using
200 the FastDNA Spin Kit for Soil (MP Biomedicals, Irvine, CA), following the manufacturer's
201 protocol. After elution, DNA was quantified via the Qubit High Sensitivity assay (Invitrogen,
202 Carlsbad, CA) and stored at -80°C until PCR. DNA was extracted from presort and sorted
203 (BONCAT active) fractions by triplicate repeated freeze-thaw cycles to lyse cells (32). The freeze
204 cycles were done at -80°C for 20 min, and thawing was done at 99°C for 10 min. These DNA
205 extracts were stored at -80°C for further processing.
206
207 *16S rRNA Gene Amplicon Sequencing*
208 Amplification of bacterial and archaeal 16S rRNA genes was performed following the Earth
209 Microbiome protocol (49) using updated primers 515F (forward primer, 5′-
210 GTGYCAGCMGCCGCGGTAA-3' (50) and 806R (reverse primer, 5′-
211 GGACTACNVGGGTWTCTAAT-3' (51). Polymerase chain reaction (PCR) was performed in a
212 final volume of 50 µL consisting of 20 µL of template DNA, 20 µL Invitrogen Platinum Taq II 2x
213 Master Mix, 1 µL of each primer at a concentration of 10 µM, and 8 µL nuclease-free water for

214 freeze/thaw extracted cells, and 25 µL consisting of 5 µL of template DNA, 10 µL Invitrogen
215 Platinum Taq II 2x Master Mix, 0.5 µL of each primer at a concentration of 10 µM, and 9 µL
216 nuclease-free water for DNA extracted with the FastDNA Spin Kit for Soil. The thermocycler
217 conditions were: 94ºC for 3 min, then 28 cycles of 94ºC for 45 sec, 50ºC for 60 sec, and 72ºC for
218 90ºC, followed by an elongation step of 72ºC for 10 min. A selection of PCR products was checked
219 for the expected length on a 1% agarose gel. All products were then purified using AMPure XP
220 beads following the manufacturer's protocol with a final elution volume of 35 µL. For Illumina
221 sequencing, Illumina dual index barcodes and sequencing adapters were attached in a second PCR
222 reaction. The final volume for this reaction was 25 µL, consisting of 5 µL of purified amplicon,
223 10 µL Invitrogen Platinum Taq II 2x Master Mix, 0.5 µL of each primer at a concentration of 10
224 µM (i5 and i7 primers), and 9 µL nuclease-free water. The thermocycler conditions were: 95ºC
225 for 3 min, then 8 cycles of 95ºC for 30 sec, 55ºC for 30 sec, 72ºC for 30 sec, followed by an
226 elongation step of 72ºC for 5 min. The PCR products were again purified using AMPure XP beads,
227 and then were quantified in triplicate reactions using Quant-iT Picogreen dsDNA assay
228 (Invitrogen) on a Biotek Synergy H1 Hybrid microplate reader. Purified amplicons were then
229 pooled with 7.5 ng of DNA each, or up to 30 µL if a sample had less than 7.5 ng of DNA. The
230 pooled library was quantified with the Qubit high-sensitivity assay. Purification, quality
231 assessment, and sequencing of pooled libraries were performed at the Molecular Research Core
232 Facility at Idaho State University using Illumina's MiSeq chemistry for 2 x 250 bp paired-end
233 reads.
234
235 *Bioinformatics Analysis*

Amplicon sequence data were processed with QIIME2 (version 2025.7) (52). Barcoding and primer sequences were removed using *cutadapt* before truncating the forward (137 bp) and reverse (139 bp) reads based on sequencing quality. DADA2 was subsequently used to filter, denoise, and merge the reads. Taxonomic assignments for each ASV (amplicon sequence variant) were determined with the Silva 138.2 database using the *sklearn* method (53). Contaminant ASVs were identified using the prevalence model (threshold = 0.5) in the R package *decontam* (54) and manual inspection of the dataset to compare samples to the no HPG controls and process blanks. Observed diversity and Shannon's diversity metrics were calculated using the *phyloseq* package in R (55), with replicate samples pooled for analysis. Linear fold enrichment was calculated in Equation 1 as the ratio of the fractional relative abundance of ASVs affiliated with a given class between the presort and sorted fractions:

$$Fold\ Enrichment = \left(\frac{max(RAsort, RApresort) + 0.1}{min(RAsort, RApresort) + 0.1} - 1\right) * sgn(RA_{\cancel{sort}} - RApresort) \quad (1)$$

Where RA_{sort} and $RA_{presort}$ are the fractional relative abundances of a given taxonomic class in the sorted and presort fractions, respectively. Fold enrichment is calculated in the first part of the function, and a small value (0.1) is added to each relative abundance to avoid dividing by zero. Subtracting that value by one shifts the ratio so that equal relative abundances give zero. The *sgn* function returns the sign of the differences to force the fold enrichment value positive.

Analysis of Metagenome Assembled Genomes

Metagenome assembled genomes (MAGs) were previously generated from IODP 385 (12). In short, trimmed reads from all drill holes were pooled and co-assembled using MEGAHIT 1.2.9,

259 and MAGs that were at least 50% complete and contained less than 10% contamination were then
260 used for downstream analyses. Taxa in control samples were identified as potential contaminants
261 and were removed from downstream analysis. Taxonomic identities of these MAGs were assigned
262 using GTDB-Tk 2 (56). MAG contamination and completeness were checked using CheckM (57).
263 Barrnap (v0.9) was used to identify 16S rRNA gene sequences from these MAGs. The 16S rRNA
264 gene sequences of these MAGs, if present, were then compared to the 16S rRNA gene amplicon
265 sequences from sorted samples generated in this study for sequence similarity. From a BLAST
266 nucleotide similarity comparison, we used a cutoff of 90% for %ID and 99% for %hit to identify
267 sequences that matched.
268
269 The organism identities of these matches were then compared between the MAGs and amplicon
270 sequences to determine if they were identified as the same organism. Importantly, the MAG and
271 16S rRNA gene amplicon sequences were independently sequenced and taxonomy was assigned
272 using different databases (Silva 138.2 for 16S rRNA gene amplicon sequences, and GTDB-Tk
273 2.1.0 for MAGs), and taxonomic identification varied between the two. Organism identity was
274 matched based on nomenclature changes in the literature to the best of our ability. For consistency
275 across our datasets, the Silva 138.2 taxonomic identification is used throughout the manuscript.
276 Six matches were identified that met our cutoff criteria of 90% for %ID and 99% for %hit, five
277 bacteria and one archaea.
278
279 MAGs were annotated with Prokka separately for archaea and bacteria. Individual genes of interest
280 for the environment were selected based on Dombrowski *et al.* (58). Carbon metabolism pathways
281 were identified using KEGG BlastKOALA. Genes for the Wood-Ljungdahl pathway were

identified using an internal protein database generated from cultured methanogens because of the presence of MAGs in the same class as other methanogens.

Results & Discussion

Sampling Sites

Samples for BONCAT incubations were collected from drill hole sites, U1545 to U1549, along a transect across Guaymas Basin during IODP Expedition 385 (Table 1). Sites U1545 and U1546 are located ~52 km off-axis of the central spreading center, and include a relatively unaltered reference (U1545) site and a more thermally altered (U1546) site with deep sediments (~350-430 mbsf) influenced by a now thermally-equilibrated ancient sill emplacement. Sites U1547 and U1548 were drilled within and adjacent to the much hotter Ringvent site. Ringvent is shaped by a shallow, more recently emplaced sill, exhibits a sill-driven hydrothermal mound, and allows us to investigate links between sediment alteration, geochemical gradients, and microbial responses to elevated temperatures. Site U1549 targeted an off-axis cold seep near a subsurface sill, where methane seepage and hydrates indicate persistent subsurface fluid migration. Sample depths ranged from 0.8 m at site U1546 to 155 m at U1545, and temperatures were between 2.8°C and 46.6°C (Table 1). Sediment and porefluid geochemical data relevant to microbial activity are outlined in Table 1, and comprehensive datasets are available in the IODP 385 expedition reports (59–62).

Workflow development

We expanded the applications of BONCAT by refining a workflow for detecting translational activity of microbial communities in extremely low-biomass marine deep-biosphere sediments that

305 maximized cell recovery and minimized interference from sediment autofluorescence. To optimize
306 this workflow and test the efficacy of cell extraction protocols, we spiked IODP 385 sediment with
307 *Escherichia coli* (*E. coli*) cells that had been grown with HPG. These sediments had been preserved
308 analogously to the IODP 385 BONCAT samples and incubated without HPG. The presence of
309 BONCAT-positive *E. coli* in cell extracts was used to assess the efficacy of different cell extraction
310 protocols to maximize cell recovery. Additionally, we sought to minimize particulate matter in the
311 final cell suspension. Autofluorescent sediment particles can interfere with FACS analyses when
312 the signal overlaps with the target fluorophore and distorts forward and back scatter, complicating
313 the accurate detection of cells.
314
315 The final cell extraction procedure was adapted from an established protocol (27, 47) and consisted
316 of two primary cell separations. The first wash with methanol only was used to extract cells that
317 are more susceptible to lysis from detergents. The second step combined sonication with a
318 detergent mixture (sodium pyrophosphate, Tween80, and methanol) to detach cells from sediment
319 particles. A 50% Nycodenz gradient was used between detergent wash steps to phase-separate
320 detached cells from sediment. Based on the number of cells and amount of autofluorescent
321 sediment in the extracted suspension, we determined that pooling eight technical replicate extracts
322 was sufficient to extract cells from low-biomass samples for downstream processing with FACS.
323 We visualized labeling of BONCAT-positive cells from click chemistry in the presort and sorted
324 fractions with fluorescence microscopy (Fig. S2). During method testing and sample analysis, the
325 presort fraction contained cells with and without HPG incorporation, while the sorted fraction only
326 contained HPG-positive cells, attesting to the sorting of BONCAT-positive cells only from low-
327 biomass marine deep biosphere sediments.

328

329 Initially, we aimed to assess the translational activity of fungi with BONCAT-FACS in this study,
330 so we developed this cell extraction protocol to include fungal cells. We used two fungal isolates
331 from Guaymas Basin that are representative of the fungal community in IODP 385 sediment (63),
332 and tested extraction methods similar to our use of *E.coli* above. Based on our testing, this cell
333 extraction method can be used to separate yeast and filamentous fungal cells from sediments.
334 Furthermore, these fungal strains incorporate HPG in liquid media and can be detected with click
335 chemistry and fluorescence microscopy (Fig. S3), demonstrating that BONCAT workflows can be
336 applied to marine fungi. However, due to issues with downstream processing, we could not reliably
337 detect fungi in the presort and sorted fractions with internal transcribed spacer (ITS) amplicon
338 sequencing. While the exact reason for this issue is unknown, poor extraction efficiency, low
339 fungal biomass (or absence of fungi), the inability to sort certain fungal cells with FACS, or poorly
340 matched PCR primers for the ITS region in natural fungal populations (64, 65) may have all
341 impeded the successful detection of fungi in our samples. Therefore, we chose to exclude any
342 further discussion of fungi from this manuscript and focus on bacteria and archaea inhabiting the
343 Guaymas Basin deep biosphere.

344

345 *Diversity of the total, cell extractable, and active prokaryotic community*

346 We combined BONCAT-FACS with 16S rRNA gene amplicon sequencing of sorted cells to
347 identify translationally active bacteria and archaea in the Guaymas Basin deep biosphere. To assess
348 community compositions, the observed diversity (the quantity of unique ASVs) and the Shannon's
349 diversity index were determined for each sample (Fig. 1A-B). For all samples, observed and
350 Shannon's diversity indices were higher in the whole sediment DNA extract fraction than in the

351 cell-extracted presort and BONCAT-active sorted fractions (Fig. S4). Generally, the diversity of
352 the total community decreased with depth, reflecting the decrease in cell abundance with depth.
353 The two deepest samples, 1547B-8H2 and 1545B-19F3 (66 and 155 mbsf, respectively), exhibited
354 the smallest difference in diversity metrics across all three fractions, especially with respect to the
355 DNA extract. Compared to shallower samples, a larger portion of the total microbial community
356 was detected in the BONCAT active fraction in these deeper sediments, where there are relatively
357 few cells, carbon is more recalcitrant, energetic conditions are less favorable, and geochemical
358 conditions remain stable over extended timescales (66, 67). This suggests that the microbial
359 population is highly specialized to fill limited available niches in the deep biosphere.
360
361 To evaluate the changes in diversity metrics between the presort and sorted fractions, we calculated
362 fold enrichment (Fig. 1C-F) between these fractions for the most abundant ASVs, which were
363 affiliated with *Gammaproteobacteria*, *Alphaproteobacteria*, *Bacilli*, and *Deinococci* (Fig. 2 and
364 Table S2). Generally, these abundant taxa were not overly represented in either fraction
365 (enrichment values near zero) or were enriched in the presort fraction (*Deinococci* and *Bacilli*),
366 suggesting that only a portion of those taxa detected in the presort fraction were actually
367 translationally active. In several cases, corresponding ASVs were not detected in either the presort
368 or sorted fraction, which skewed enrichment toward the respective other fraction. For instance, in
369 sample 1548B-2H3, no sequences affiliated with *Alphaproteobacteria* (Fig. 1D) were detected in
370 the presort fraction, artificially inflating the fold enrichment toward the sorted fraction. Similarly,
371 no *Deinococci* affiliated ASVs were observed in the sorted fraction for one replicate from 1547B-
372 1H2 (Fig. 1F). The inability to detect a given ASV in a specific fraction could skew diversity
373 indices, making it difficult to interpret diversity data in those samples.

374

375 Observed diversity of unique ASVs increased from the presort to the sorted fractions in all samples
376 (Fig. 1A), and Shannon's diversity increased in 1549B-3H2 and 1547B-8H2 between those
377 fractions (Fig. 1B). It should be expected that diversity would decrease between the presort and
378 the BONCAT active fraction to reflect that only a portion of the community is translationally active
379 and detectable with BONCAT-FACS. One explanation for this outcome is that there were
380 substantially fewer reads in many of the presort samples than in the sorted samples (Fig. S5), likely
381 due to the relatively small subsample (50 µL) taken from the presort samples for DNA extraction
382 and low amount of biomass in the sample. More ASVs were detected when the number of reads
383 was higher (Fig. S5), which would influence alpha diversity metrics. In the two samples in which
384 the Shannon's diversity increased in the sorted fraction (1547B-8H2 and 1549B-3H2), across all
385 replicates, we detected between three and eleven unique class-level ASVs in the presort samples
386 and between 12 and 30 in the sorted samples. This indicates that sequencing depth directly impacts
387 the detection of additional and/or rare taxa that may have been, in some cases, missed in the presort
388 fraction.

389

390 We observed major differences in Shannon's diversity indices across the BONCAT workflow
391 (DNA extract to presort to sorted fractions). Shannon's diversity was highest in the DNA extract
392 relative to the presort or sorted fractions for nearly every sample (the only exception was the sorted
393 fraction for 1547B-8H2) (Fig. S4). This is expected for the marine subsurface as extracellular DNA
394 in the sediments may have been amplified with PCR, altering community composition and
395 artificially inflating diversity metrics in the initial DNA extract fraction (24). Diversity indices
396 were generally lower in the presort and sorted fraction, as cell extraction for BONCAT-FACS is

397 known to introduce biases (27, 32) that could cause large discrepancies in diversity metrics
398 between the DNA extract and presort fractions. To address this, we optimized the cell extraction
399 procedure for these low-biomass deep biosphere samples, but biased cell recovery is a challenge
400 for any complex sample (19, 68), especially from Guaymas Basin sediments that are enriched with
401 hydrothermally altered organic carbon. Ideally, cell extraction procedures should be tuned to
402 specific samples based on experimental goals and biogeochemical conditions.
403
404 *Abundant Taxa and Active Community*
405 To assess the active community within the Guaymas Basin deep biosphere, we analyzed the
406 relative abundances of ASVs in each fraction and compared those frequencies between samples
407 (Fig. 2 and Fig. S6). In the presort cell extract and BONCAT active sorted fractions, we identified
408 a total of 40 bacterial classes and six archaeal classes, which were frequently dominated by a few
409 classes. Across all three fractions (DNA extract, presort, and sorted), ASVs affiliated with
410 *Gammaproteobacteria* were consistently the most abundant taxa detected. *Gammaproteobacteria*
411 had greater than 50% relative abundance within the BONCAT active community in 12 of the sorted
412 fraction samples (Fig. 2). In three of the six samples (1546B-1H2, 1547B-1H2, and 1545B-19F3),
413 all *Gammaproteobacterial* genera were detected in higher relative abundance in the sorted fraction
414 than in the presort fraction (Fig. 2), which indicates members of this group are active in the
415 Guaymas Basin deep biosphere. High abundances of *Gammaproteobacteria* in the BONCAT
416 active fraction likely reflect the heterotrophic nature of this class, which is capable of degrading
417 diverse organic substrates typically found in Guaymas Basin sediments (69–71). While the relative
418 frequency of *Gammaproteobacteria* was elevated in most samples, differences in relative

419 abundance between fractions could be explained by enrichment of *Gammaproteobacteria* during
420 BONCAT incubation, or methodological biases against other taxonomic groups.
421
422 Across all three fractions, we identified 61 genera affiliated with the *Gammaproteobacteria* (Fig.
423 S7 and Table S3) that are associated with diverse heterotrophic metabolisms. As observed for the
424 whole community analysis, the highest number of gammaproteobacterial genera was observed in
425 the DNA extract, and ASVs affiliated with the genus *Thiomicrohabdus* were amongst the most
426 abundant. In contrast, the presort and sorted fractions were primarily dominated by members of
427 *Halomonas,* and to a lesser extent, *Burkholderia-Caaballerionia-Paraburkholderia* and
428 *Acinetobacter*, which had the highest relative frequency in the deepest sample (1545B-19F3). In
429 several samples, ASVs belonging to *Halomonas* (1546B-1H2, 1547B-1H2, and 1549B-3H2) and
430 *Acinetobacter* (1545B-19F3) exhibited a relative frequency >75% in the BONCAT active fraction
431 in at least two replicates. This was a sharp increase compared to the other fractions which, across
432 all of those samples, were typically less than 20% for the DNA extract and all less than 3% for the
433 presort fractions. The relative abundance of *Halomonas* demonstrated the largest observed
434 increase between the presort and sorted fractions (Fig. S7). In particular, BONCAT active
435 *Halomonas* ASVs were detected at elevated frequencies in samples with temperatures <15°C
436 (Table 1), compared to warmer samples (~38°C). Sediment temperatures can fluctuate by >50°C
437 on relatively small temporal and spatial scales due to hydrothermal activity (72), which may create
438 a selective pressure on the microbial community. This indicates that some members of the
439 *Halomonas* genus remain metabolically active at lower temperatures and withstand large
440 temperature changes characteristic of hydrothermal sediments. Genomic analyses of these
441 communities revealed genes associated with central carbon metabolism, CAZymes, hydrocarbon

442 degradation, and nitrogen, sulfur, and arsenic cycling, suggesting that *Gammaproteobacteria* are
443 key drivers of organic matter remineralization and redox transformations in the subsurface (71,
444 73).
445
446 In addition to *Gammaproteobacteria*, abundant ASVs in the sorted fraction were affiliated with
447 *Alphaproteobacteria* throughout our samples. The relative frequency of translationally active
448 *Alphaproteobacteria* was between 28-52% for three replicates in the relatively warm (~14°C)
449 sediment sample from 1548B-2H3 and up to 19% in the deepest sample from 155 mbsf at the more
450 thermally moderate site 1545B (~39°C). After BONCAT incubation, *Alphaproteobacteria*
451 exhibited higher relative frequencies in the sorted fraction in 4 samples compared to the presort
452 fraction (Fig. 1D). Some of the supposed increase in frequency between fractions was due to low
453 detection of *Alphaproteobacteria* ASVs in the presort fraction. Regardless, the presence of
454 *Alphaproteobacteria* in the sorted fractions demonstrates that members of this class are active in
455 the Guaymas Basin deep biosphere. Similar to our observations for the *Gammaproteobacteria*,
456 this is consistent with previous studies that have demonstrated high abundances of
457 *Alphaproteobacteria* and associated heterotrophic activity in the marine deep biosphere (69, 74–
458 76).
459
460 Members of the class *Deinococci* were detected in all fractions, and consistently exhibited the
461 highest relative abundance in the presort fraction compared to the DNA extract and active sorted
462 fractions (Fig. 2 and Fig. S6). At Ringvent, in sample 1547B-1H2 (2.2 mbsf), *Deinococci* ASVs
463 constituted up to 54% of the presort fraction, compared to less than 1% in the DNA extract. This
464 higher frequency in the presort fraction suggests our cell extraction method is biased in favor of

465 *Deinococci* cells, which may influence relative abundances in the BONCAT active fraction. In
466 sample 1547B-8H2 (66 mbsf), members of this class accounted for up to 15% relative abundance
467 in the BONCAT active fraction and were less than 6% in the corresponding presort fraction. In
468 hydrothermally impacted habitats, *Deinococci*-affiliated sequences have been detected via 16S
469 rRNA gene amplicon survey (77). Additionally, analysis of 52 MAGs assigned to *Deinococci* from
470 IODP 360 to Atlantis Bank revealed genes associated with broad anaerobic and heterotrophic
471 metabolisms, including fatty acid and aromatic compound degradation (*e.g.*, acyl-CoA
472 dehydrogenase and benzoyl-CoA), acetogenesis, nitrogen and sulfur cycling, and metal reduction
473 (78). This versatility can promote *Deinococci* metabolism across the dynamic geochemical
474 gradients observed in the Guaymas Basin deep biosphere (Table 1).
475
476 A total of 7 classes affiliated with phylum *Firmicutes* were detected throughout our samples, and
477 the most abundant ASVs from this phylum were assigned to the *Bacilli* class in the presort and
478 active sorted fractions. Although ASVs affiliated with *Bacilli* were detected throughout the
479 sediment, the highest relative abundances in the BONCAT sorted fraction, up to ~20% (one
480 replicate was above 95%), were observed in the shallower samples (1546B-1H2 and 1547B-1H2)
481 (Fig. 2). In those shallower depths, above the sulfate methane transition zone, where sulfate is the
482 dominant terminal electron acceptor (Table 1) for anaerobic metabolisms, BONCAT-active ASVs
483 affiliated with *Firmicutes* may couple sulfate reduction to organic matter degradation (79).
484 *Firmicutes* isolates cultured from several marine deep biosphere sites degrade a broad suite of
485 organic substrates, including acetate, polysaccharides, amino acids, alcohols, nucleotides, and
486 microbial necromass, indicating a key role in heterotrophic carbon cycling (27, 69, 80, 81). This
487 metabolic capacity was substantiated by metagenomic analysis that also demonstrated high

488 thermal tolerance, exceeding 80°C (82). Additionally, members of the *Firmicutes* phylum within
489 the *Clostridia* class were associated with degradation of crude oil and polycyclic aromatic
490 hydrocarbon degradation in methanogenic enrichments (83). Hydrothermally induced pyrolysis of
491 organic matter in Guaymas Basin sediments generates petroleum hydrocarbons that are dispersed
492 throughout the sediment column and can fuel diverse microbial metabolisms (41, 42, 84). In the
493 Guaymas Basin specifically, *Firmicutes* have been linked to both sulfate and nitrate reduction (79)
494 and have dominated enrichment cultures amended with methanol, monomethylamine, and
495 ammonium (85). Given the wide metabolic capabilities of the *Firmicutes* and the high frequency
496 in the BONCAT active fraction, members of this phylum likely contribute significantly to organic
497 matter degradation and elemental cycling in Guaymas Basin sediments.
498
499 *Implications for deep biosphere carbon cycling*
500 The BONCAT active fractions were dominated by a heterotrophic microbial population capable
501 of metabolizing a wide range of organic carbon in the Guaymas Basin deep biosphere. To further
502 assess this metabolic capacity and explore potential implications for biogeochemical cycling in
503 hydrothermal sediments, we mapped 16S rRNA gene ASVs from the BONCAT active fraction to
504 annotated metagenome assembled genomes (MAGs) from IODP 385 (12). Six unique MAGs
505 associated with the BONCAT active fraction were identified as members of 5 bacterial phyla
506 (*Planctomycetota*, *Chloroflexota*, *Acidobacteriota*, *Aerophobota*, and *TA06*) and 1 archaeal
507 phylum (*Thermoproteota*). These MAGs were 52.7-95.9% complete and all had contamination
508 below 10% (Table S4).
509

510 The translationally active MAGs harbored genes associated with diverse carbon cycling pathways,
511 including carbohydrate degradation, C1 metabolism, and fermentation, plus nitrogen and sulfur
512 cycling (Fig. 3). MAG_005 (95.91% complete, 9.60% contamination) was assigned to the phylum
513 *Planctomycetota* and encodes complete pathways for glycolysis and galactose degradation, and a
514 nitrite reductase gene. Recently, the first anaerobic culture of *Planctomycetota* isolated from
515 marine sediments experimentally demonstrated functional capacity for these metabolisms (86),
516 validating the genomic potential of that isolate and MAG_005. MAG_054 (52.66% complete,
517 3.54% contamination), assigned to the class *Anaerolineae*, contained genes associated with
518 carbohydrate degradation and fermentation, both of which have been demonstrated by cultured
519 *Anaerolineae* (87). Similarly, NanoSIMS experiments implicated diverse *Chloroflexota* (phylum
520 of *Anaerolineae*) in fermentation and broad heterotrophic metabolisms in hypersaline
521 photoheterotrophic mats (88). In their experiments, Lee *et al*. (2014) ascertained ^{13}C-substrate
522 assimilation by *Chloroflexota* members using phylum-specific FISH probes, but could not resolve
523 taxonomic affiliations on class level or below, or implicate other phyla in carbon cycling (88). In
524 contrast, BONCAT-FACS workflows provide a high-throughput community-level analysis of
525 active taxa, and when coupled with metagenomics, support the investigation of the metabolic
526 potential of sorted cells (89). Taken together, BONCAT-FACS, paired with metagenomics, can
527 corroborate the results from culture-dependent and -independent approaches, while also providing
528 metabolic insights at the community level.
529
530 Most importantly, metagenomic analysis, corroborated by 16S rRNA gene identities, revealed a
531 diverse, translationally active microbial community in the Guaymas Basin deep biosphere with the
532 metabolic capacity for a range of carbon transformations and broad anaerobic metabolisms. Our

533 results confirm that these taxa are active, and that under the right biogeochemical conditions (*e.g.,*
534 substrate availability, temperature, etc.), the organisms represented by these MAGs likely utilize
535 these metabolisms for energy conservation in the deep biosphere. However, experimental evidence
536 is necessary to confirm which of these metabolisms are used *in situ* under different conditions.
537 Additionally, the specific ASVs affiliated with these MAGs were not necessarily amongst the most
538 abundant or widespread taxa in the BONCAT sorted fraction. Thus, the extent to which these
539 microbes drive carbon transformations is yet unknown. Future experiments that combine
540 BONCAT-FACS with substrate amendment (27, 32) and paired metagenomics (89) and/or
541 metatranscriptomics could provide deeper, community-level insights into the metabolic
542 capabilities of translationally active taxa in the marine deep biosphere.
543

544 **Conclusion**

545 In this study, we assessed the total, cell-extracted (presort), and translationally active (sorted)
546 microbial community in contrasting anoxic sediments of the Guaymas Basin deep biosphere. We
547 identified a diverse extant microbial community down to 154 meters below the sediment surface,
548 and the DNA extract contained 139 unique classes. We optimized our cell extraction and FACS
549 procedures for low-biomass deep biosphere samples and successfully identified 46 classes in the
550 presort and sorted fractions. Alpha diversity metrics decreased sharply from the DNA extract
551 fraction to the presort fraction, likely due to challenges associated with separating certain taxa
552 from the sediments, but were comparable between the presort and sorted fractions..
553 *Gammaproteobacteria* overwhelmingly dominated the BONCAT-active fraction, followed by
554 *Alphaproteobacteria*, *Firmicutes*, and *Deinococci*. These phyla have the capacity for a wide range

555 of heterotrophic metabolisms and drive microbial-mediated recycling of hydrothermally altered
556 organic matter in the Guaymas Basin deep biosphere.
557
558 We demonstrated that BONCAT-FACS can be used to assess microbial activity in low-biomass
559 marine sediments on a community level. Additionally, we paired this with metagenomic analysis
560 from IODP 385 to identify six translationally active MAGs and assess their metabolic capacity.
561 BONCAT-FACS coupled metagenomics links single-cell activity and genetic potential to
562 understand community-level dynamics in a high-throughput manner (19, 89). This approach could
563 be further optimized in the future to generate metagenomes directly from the sorted fraction and/or
564 by combining experimental amendments with whole sediment metagenomics to more tightly link
565 taxa with metabolic function. Importantly, our workflow, while optimized for low-biomass
566 samples, was likely biased against rare taxa, given the decrease in diversity between fractions.
567 These biases could be exacerbated if combining with metagenomics, which would require greater
568 quantities of biomass for FACS and sequencing compared to gene amplicon sequencing. With
569 careful benchmarking and optimization, BONCAT-FACS and metagenomics, when combined,
570 offer a powerful approach for connecting microbial taxa and metabolic function in the marine deep
571 biosphere and other complex environmental samples.
572
573 **Acknowledgments**
574 We would like to thank the scientific party and D/V JOIDES Resolution crew from IODP 385 for
575 their assistance in sample collection. Amplicon sequencing was performed by the Molecular
576 Research Core Facility at Idaho State University (RRID:SCR_012598). This work was supported
577 by NSF funding to A.M. (PRFB 2010880), R.H. (OCE-2049445), R.H. and V.E. (OCE-2046056),

V.E. and R.H. (OCE-2046799), and V.E. (OCE-1829903). Large language models (LLMs) were used to assist in developing data visualization scripts.

Author Contributions

R.H. and V.E. conceptualized and supervised this study. V.E. collected the samples and performed shipboard sample processing for BONCAT incubations. A.M., S.N., and C.G. developed the methodology, performed experiments, and curated data. A.M., S.N., and Z.J. performed bioinformatics analysis. A.M. and S.N. analyzed the data and wrote the manuscript. All authors edited and gave critical feedback on the final manuscript. A.M., V.E., and R.H. raised funding. A.M. and S.N. contributed to this work equally.

Data Availability

Amplicon sequence data generated for this manuscript are publicly available under NCBI Sequence Read Archive (SRA) Bio Project accession number PRJNA1356905 (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA1356905) and reads are available under accession numbers SRR35956508-SRR35956585. The metagenome assembled genomes are publicly available under NCBI SRA Bio Project accession number PRJNA909197 (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA909197), and reads are under accession numbers SRR22580794-SRR22580807 and SRR23614663-SRR23614677. The full biogeochemical datasets from IODP 385, including data presented in this manuscript, are available in the IODP expedition report (http://publications.iodp.org/proceedings/385/385title.html

References

1. Whitman WB, Coleman DC, Wiebe WJ. 1998. Prokaryotes: The unseen majority. Proceedings of the National Academy of Sciences 95:6578–6583.

2. Kallmeyer J, Pockalny R, Adhikari RR, Smith DC, D'Hondt S. 2012. Global distribution of microbial abundance and biomass in subseafloor sediment. Proceedings of the National Academy of Sciences of the United States of America 109:16213–16216.

3. Oremland RS, Culbertson C, Simoneit BRT. 1982. Methanogenic activity in sediment from Leg 64, Gulf of California. Initial reports DSDP, Leg 64, Mazatlan to Long Beach, 1978-79 Part 2 759–762.

4. D'Hondt S, Rutherford S, Spivack AJ, Hondt SD, Rutherford S, Spivack AJ. 2002. Metabolic Activity of Subsurface Life in Deep-Sea Sediments. Science 295:2067–2070.

5. Lever MA. 2012. Acetogenesis in the energy-starved deep biosphere-a paradox? Frontiers in Microbiology 2:1–18.

6. Teske A, Callaghan AV, LaRowe DE. 2014. Biosphere frontiers of subsurface life in the sedimented hydrothermal system of Guaymas Basin. Frontiers in Microbiology 5:1–11.

7. Santelli CM, Orcutt BN, Banning E, Bach W, Moyer CL, Sogin ML, Staudigel H, Edwards KJ. 2008. Abundance and diversity of microbial life in ocean crust. Nature 453:653–656.

8. Jørgensen BB. 2011. Deep subseafloor microbial cells on physiological standby. Proceedings of the National Academy of Sciences 108:18193–18194.

618 9. Lever MA, Alperin M, Engelen B, Inagaki F, Nakagawa S, Steinsbu BO, Teske A. 2006.
619 Trends in basalt and sediment core contamination during IODP expedition 301. Geomicrobiology
620 Journal 23:517–530.

621 10. Orcutt BN, LaRowe DE, Biddle JF, Colwell FS, Glazer BT, Reese BK, Kirkpatrick JB,
622 Lapham L, Mills HJ, Sylvan JB, Wankel SD, Wheat CG. 2013. Microbial activity in the marine
623 deep biosphere: Progress and prospects. Frontiers in Microbiology 4:1–15.

624 11. Gittins DA, Desiage P-A, Morrison N, Rattray JE, Bhatnagar S, Chakraborty A, Zorz J, Li
625 C, Horanszky O, Cramm MA, Bisiach F, Bennett R, Webb J, MacDonald A, Fowler M, Campbell
626 DC, Hubert CRJ. 2022. Geological processes mediate a microbial dispersal loop in the deep
627 biosphere. Science Advances 8:eabn3485.

628 12. Mara P, Geller-McGrath D, Edgcomb V, Beaudoin D, Morono Y, Teske A. 2023.
629 Metagenomic profiles of archaea and bacteria within thermal and geochemical gradients of the
630 Guaymas Basin deep subsurface. Nat Commun 14:7768.

631 13. Sun J, Hirai M, Takaki Y, Evans PN, Nunoura T, Rinke C. 2025. Metagenomic insights
632 into taxonomic and functional patterns in shallow coastal and deep subseafloor sediments in the
633 Western Pacific. Microbial Genomics 11:001351.

634 14. Orsi WD, Edgcomb VP, Christman GD, Biddle JF. 2013. Gene expression in the deep
635 biosphere. Nature 499:205–208.

636 15. Coskun ÖK, Pichler M, Vargas S, Gilder S, Orsi WD. 2018. Linking Uncultivated
637 Microbial Populations and Benthic Carbon Turnover by Using Quantitative Stable Isotope
638 Probing. Applied and Environmental Microbiology 84.

639 16. Coskun ÖK, Özen V, Wankel SD, Orsi WD. 2019. Quantifying population-specific growth
640 in benthic bacterial communities under low oxygen using H218O. The ISME Journal 13.

641 17. Mara P, Zhou Y-L, Teske A, Morono Y, Beaudoin D, Edgcomb V. 2023. Microbial gene
642 expression in Guaymas Basin subsurface sediments responds to hydrothermal stress and energy
643 limitation. 11. ISME J 17:1907–1919.

644 18. Mara P, Beaudoin D, Aiello I, Morono Y, Geller-McGrath D, Edgcomb VP, Teske A.
645 2024. Deep subseafloor sediments in Guaymas Basin harbor cosmopolitan microbiota and traces
646 of hydrothermal populations. Commun Earth Environ 5:505.

647 19. Hatzenpichler R, Krukenberg V, Spietz RL, Jay ZJ. 2020. Next-generation physiology
648 approaches to study microbiome function at single cell level. Nature Reviews Microbiology
649 18:241–256.

650 20. Hatzenpichler R, Scheller S, Tavormina PL, Babin BM, Tirrell DA, Orphan VJ. 2014. In
651 situ visualization of newly synthesized proteins in environmental microbes using amino acid
652 tagging and click chemistry. Environmental Microbiology 16:2568–2590.

653 21. Hatzenpichler R, Connon SA, Goudeau D, Malmstrom RR, Woyke T, Orphan VJ. 2016.
654 Visualizing in situ translational activity for identifying and sorting slow-growing archaeal -
655 bacterial consortia. Proceedings of the National Academy of Sciences 113:E4069–E4078.

656 22. Samo TJ, Smriga S, Malfatti F, Sherwood BP, Azam F. 2014. Broad distribution and high
657 proportion of protein synthesis active marine bacteria revealed by click chemistry at the single cell
658 level. Frontiers in Marine Science 1:1–18.

659 23. Leizeaga A, Estrany M, Forn I, Sebastián M. 2017. Using click-chemistry for visualizing
660 in situ changes of translational activity in planktonic marine bacteria. Frontiers in Microbiology
661 8:1–11.

662 24. Couradeau E, Sasse J, Goudeau D, Nath N, Hazen TC, Bowen BP, Chakraborty R,
663 Malmstrom RR, Northen TR. 2019. Probing the active fraction of soil microbiomes using
664 BONCAT-FACS. Nature Communications 10.

665 25. Kiick KL, Saxon E, Tirrell DA, Bertozzi CR. 2002. Incorporation of azides into
666 recombinant proteins for chemoselective modification by the Staudinger ligation. Proceedings of
667 the National Academy of Sciences 99:19–24.

668 26. Mulay SA, Vishnivetskaya TA, Hochanadel LH, Klingeman DM, Lloyd KG, Pelletier DA,
669 Podar M. 2025. BONCAT-Live for isolation and cultivation of active environmental bacteria.
670 mBio.

671 27. Krukenberg V, Reichart NJ, Spietz RL, Hatzenpichler R. 2021. Microbial Community
672 Response to Polysaccharide Amendment in Anoxic Hydrothermal Sediments of the Guaymas
673 Basin. Frontiers in Microbiology 12.

674 28. Harris JE, Bledsoe RB, Guha S, Omari H, Crandall SG, Burghardt LT, Couradeau E. 2025.
675 The activity of soil microbial taxa in the rhizosphere predicts the success of root colonization.
676 mSystems 0:e00458-25.

677 29. Riva A, Kolimár D, Spittler A, Wisgrill L, Herbold CW, Abrankó L, Berry D. 2020.
678 Conversion of Rutin, a Prevalent Dietary Flavonol, by the Human Gut Microbiota. Front Microbiol
679 11.

680 30. Valentini TD, Lucas SK, Binder KA, Cameron LC, Motl JA, Dunitz JM, Hunter RC. 2020.
681 Bioorthogonal non-canonical amino acid tagging reveals translationally active subpopulations of
682 the cystic fibrosis lung microbiota. Nat Commun 11:2287.

683 31. Taguer M, Shapiro BJ, Maurice CF. 2021. Translational activity is uncoupled from nucleic
684 acid content in bacterial cells of the human gut microbiota. Gut Microbes 13:1903289.

685 32. Reichart NJ, Jay ZJ, Krukenberg V, Parker AE, Spietz RL, Hatzenpichler R. 2020.
686 Activity-based cell sorting reveals responses of uncultured archaea and bacteria to substrate
687 amendment. The ISME Journal 14:2851–2861.

688 33. Chen L, Zhao B, Li X, Cheng Z, Wu R, Xia Y. 2021. Isolating and characterizing
689 translationally active fraction of anammox microbiota using bioorthogonal non-canonical amino
690 acid tagging. Chemical Engineering Journal 418:129411.

691 34. Du Z, Behrens SF. 2021. Tracking *de novo* protein synthesis in the activated sludge
692 microbiome using BONCAT-FACS. Water Research 205:117696.

693 35. Trexler RV, Van Goethem MW, Goudeau D, Nath N, Malmstrom RR, Northen TR,
694 Couradeau E. 2023. BONCAT-FACS-Seq reveals the active fraction of a biocrust community
695 undergoing a wet-up event. Front Microbiol 14.

696 36. Calvert SE. 1966. Origin of Diatom-Rich, Varved Sediments from the Gulf of California.
697 The Journal of Geology 74:546–565.

698 37. Einsele G, Gieskes JM, Curray J, Moore DM, Aguayo E, Aubry M-P, Fornari D, Guerrero
699 J, Kastner M, Kelts K, Lyle M, Matoba Y, Molina-Cruz A, Niemitz J, Rueda J, Saunders A,

Schrader H, Simoneit B, Vacquier V. 1980. Intrusion of basaltic sills into highly porous sediments, and resulting hydrothermal activity. Nature 283:441–445.

38. Lonsdale P, Becker K. 1985. Hydrothermal plumes, hot springs, and conductive heat flow in the Southern Trough of Guaymas Basin. Earth and Planetary Science Letters 73:211–225.

39. Lin YS, Koch BP, Feseker T, Ziervogel K, Goldhammer T, Schmidt F, Witt M, Kellermann MY, Zabel M, Teske A, Hinrichs KU. 2017. Near-surface heating of young rift sediment causes mass production and discharge of reactive dissolved organic matter. Scientific Reports 7:1–10.

40. Simoneit BRT, Mazurek MA, Brenner S, Crisp PT, Kaplan IR. 1979. Organic geochemistry of recent sediments from Guaymas Basin, Gulf of California. Deep Sea Research Part A, Oceanographic Research Papers 26:879–891.

41. Kawka OE, Simoneit BRT. 1987. Survey of hydrothermally-generated petroleums from the Guaymas Basin spreading center. Organic Geochemistry 11:311–328.

42. Bazylinski DA, Farrington JW, Jannasch HW. 1988. Hydrocarbons in surface sediments from a Guaymas Basin hydrothermal vent site. Organic Geochemistry 12:547–558.

43. Whelan JK, Simoneit BRT, Tarafa ME. 1988. C1-C8 hydrocarbons in sediments from Guaymas Basin, Gulf of California-Comparison to Peru Margin, Japan Trench and California Borderlands. Organic Geochemistry 12:171–194.

44. Martens CS. 1990. Generation of short chain acid anions in hydrothermally altered sediments of the Guaymas Basin, Gulf of California. Applied Geochemistry 5:71–76.

45. Zhuang G, Montgomery A, Samarkin VA, Song M, Liu J, Schubotz F, Teske A, Hinrichs K-U, Joye SB. 2019. Generation and Utilization of Volatile Fatty Acids and Alcohols in Hydrothermally Altered Sediments in the Guaymas Basin, Gulf of California. Geophysical Research Letters 46:2637–2646.

46. Teske A, Lizarralde D, Höfig TW, Aiello IW, Ash JL, Bojanova DP, Buatier MD, Edgcomb VP, Galerne CY. 2021. "Expedition 385 Summary" in: Guaymas Basin Tectonics and Biosphere. International Ocean Discovery Program.

47. Kallmeyer J, Smith DC, Spivack AJ, D'Hondt S. 2008. New cell extraction procedure applied to deep subsurface sediments. Limnology and Oceanography: Methods 6:236–245.

48. Hatzenpichler R, Orphan VJ. 2015. Detection of Protein-Synthesizing Microorganisms in the Environment via Bioorthogonal Noncanonical Amino Acid Tagging (BONCAT). Hydrocarbon and Lipid Microbiology Protocols Vol 7: Single-Cell and Single-Molecule Methods 171–195.

49. Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, Locey KJ, Prill RJ, Tripathi A, Gibbons SM, Ackermann G, Navas-Molina JA, Janssen S, Kopylova E, Vázquez-Baeza Y, González A, Morton JT, Mirarab S, Zech Xu Z, Jiang L, Haroon MF, Kanbar J, Zhu Q, Jin Song S, Kosciolek T, Bokulich NA, Lefler J, Brislawn CJ, Humphrey G, Owens SM, Hampton-Marcell J, Berg-Lyons D, McKenzie V, Fierer N, Fuhrman JA, Clauset A, Stevens RL, Shade A, Pollard KS, Goodwin KD, Jansson JK, Gilbert JA, Knight R. 2017. A communal catalogue reveals Earth's multiscale microbial diversity. Nature 551:457–463.

739 50. Parada AE, Needham DM, Fuhrman JA. 2016. Every base matters: Assessing small subunit
740 rRNA primers for marine microbiomes with mock communities, time series and global field
741 samples. Environmental Microbiology 18:1403–1414.

742 51. Apprill A, McNally S, Parsons R, Weber L. 2015. Minor revision to V4 region SSU rRNA
743 806R gene primer greatly increases detection of SAR11 bacterioplankton. Aquatic Microbial
744 Ecology 75:129–137.

745 52. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H,
746 Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown
747 CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein
748 PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz
749 JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H,
750 Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang KB, Keefe CR,
751 Keim P, Kelley ST, Knights D, Koester I, Kosciolek T, Kreps J, Langille MGI, Lee J, Ley R, Liu
752 Y-X, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik
753 AV, Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian
754 SB, Pearson T, Peoples SL, Petras D, Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Robeson
755 MS, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD,
756 Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, van der Hooft JJJ,
757 Vargas F, Vázquez-Baeza Y, Vogtmann E, von Hippel M, Walters W, Wan Y, Wang M, Warren
758 J, Weber KC, Williamson CHD, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R,
759 Caporaso JG. 2019. Reproducible, interactive, scalable and extensible microbiome data science
760 using QIIME 2. Nat Biotechnol 37:852–857.

761 53. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO.
762 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based
763 tools. Nucleic Acids Res 41:D590-596.

764 54. Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ. 2018. Simple statistical
765 identification and removal of contaminant sequences in marker-gene and metagenomics data.
766 Microbiome 6:226.

767 55. McMurdie PJ, Holmes S. 2013. phyloseq: An R Package for Reproducible Interactive
768 Analysis and Graphics of Microbiome Census Data. PLOS ONE 8:e61217.

769 56. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2020. GTDB-Tk: a toolkit to classify
770 genomes with the Genome Taxonomy Database. Bioinformatics 36:1925–1927.

771 57. Chklovski A, Parks DH, Woodcroft BJ, Tyson GW. 2023. CheckM2: a rapid, scalable and
772 accurate tool for assessing microbial genome quality using machine learning. Nat Methods 1–10.

773 58. Dombrowski N, Seitz KW, Teske A, Baker BJ. 2017. Genomic insights into potential
774 interdependencies in microbial hydrocarbon and nutrient cycling in hydrothermal sediments.
775 Microbiome 5:106.

776 59. Teske AP, Lizarralde D, Höfig TW, Aiello IW, Ash JL, Bojanova DP, Buatier MD,
777 Edgcomb VP, Galerne CY, Gontharet S, Heuer VB, Jiang S, Kars MAC, Singh SK, Kim J-H,
778 Koornneef LMT, Marsaglia KM, Meyer N, Morono Y, Negrete-Aranda R, Neumann F, Pastor LC,
779 Peña-Salinas ME, Pérez-Cruz LL, Ran L, Riboulleau A, Sarao JA, Schubert F, Stock JM, Toffin
780 LMAA, Xie W, Yamanaka T, Zhuang G. 2021. Site U1545. In Teske, A., Lizarralde, D., Höfig,

T. W., and the Expedition 385 Scientists, Guaymas Basin Tectonics and Biosphere. Proceedings of the International Ocean Discovery Program 385.

60. Teske AP, Lizarralde D, Höfig TW, Aiello IW, Ash JL, Bojanova DP, Buatier MD, Edgcomb VP, Galerne CY, Gontharet S, Heuer VB, Jiang S, Kars MAC, Singh SK, Kim J-H, Koornneef LMT, Marsaglia KM, Meyer N, Morono Y, Negrete-Aranda R, Neumann F, Pastor LC, Peña-Salinas ME, Pérez-Cruz LL, Ran L, Riboulleau A, Sarao JA, Schubert F, Stock JM, Toffin LMAA, Xie W, Yamanaka T, Zhuang G. 2021. Site U1546. In Teske, A., Lizarralde, D., Höfig, T. W., and the Expedition 385 Scientists, Guaymas Basin Tectonics and Biosphere. Proceedings of the International Ocean Discovery Program 385.

61. Teske AP, Lizarralde D, Höfig TW, Aiello IW, Ash JL, Bojanova DP, Buatier MD, Edgcomb VP, Galerne CY, Gontharet S, Heuer VB, Jiang S, Kars MAC, Singh SK, Kim J-H, Koornneef LMT, Marsaglia KM, Meyer N, Morono Y, Negrete-Aranda R, Neumann F, Pastor LC, Peña-Salinas ME, Pérez-Cruz LL, Ran L, Riboulleau A, Sarao JA, Schubert F, Stock JM, Toffin LMAA, Xie W, Yamanaka T, Zhuang G. 2021. Sites U1547 and U1548. In Teske, A., Lizarralde, D., Höfig, T. W., and the Expedition 385 Scientists, Guaymas Basin Tectonics and Biosphere. Proceedings of the International Ocean Discovery Program 385.

62. Teske AP, Lizarralde D, Höfig TW, Aiello IW, Ash JL, Bojanova DP, Buatier MD, Edgcomb VP, Galerne CY, Gontharet S, Heuer VB, Jiang S, Kars MAC, Singh SK, Kim J-H, Koornneef LMT, Marsaglia KM, Meyer N, Morono Y, Negrete-Aranda R, Neumann F, Pastor LC, Peña-Salinas ME, Pérez-Cruz LL, Ran L, Riboulleau A, Sarao JA, Schubert F, Stock JM, Toffin LMAA, Xie W, Yamanaka T, Zhuang G. 2021. Site U1549. In Teske, A., Lizarralde, D., Höfig,

802 T. W., and the Expedition 385 Scientists, Guaymas Basin Tectonics and Biosphere. Proceedings
803 of the International Ocean Discovery Program 385.

804 63. Keeler E, Burgaud G, Teske A, Beaudoin D, Mehiri M, Dayras M, Cassand J, Edgcomb
805 V. 2021. Deep-sea hydrothermal vent sediments reveal diverse fungi with antibacterial activities.
806 FEMS Microbiology Ecology 97:fiab103.

807 64. Taylor DL, Walters WA, Lennon NJ, Bochicchio J, Krohn A, Caporaso JG, Pennanen T.
808 2016. Accurate Estimation of Fungal Diversity and Abundance through Improved Lineage-
809 Specific Primers Optimized for Illumina Amplicon Sequencing. Applied and Environmental
810 Microbiology 82:7217–7226.

811 65. Tedersoo L, Anslan S, Bahram M, Põlme S, Riit T, Liiv I, Kõljalg U, Kisand V, Nilsson
812 H, Hildebrand F, Bork P, Abarenkov K. 2015. Shotgun metagenomes and multiple primer pair-
813 barcode combinations of amplicons reveal biases in metabarcoding analyses of fungi. MycoKeys
814 10:1–43.

815 66. Inagaki F, Hinrichs KU, Kubo Y, Bowles MW, Heuer VB, Hong WL,
816 Hoshino T, Ijiri A, Imachi H, Ito M, Kaneko M, Lever MA, Lin YS, Methé BA,
817 Morita S, Morono Y, Tanikawa W, Bihan M, Bowden SA, Elvert M, Glombitza
818 C, Gross D, Harrington GJ, Hori T, Li K, Limmer D, Liu CH, Murayama M,
819 Ohkouchi N, Ono S, Park YS, Phillips SC, Prieto-Mollar X, Purkey M,
820 Riedinger N, Sanada Y, Sauvage J, Snyder G, Susilawati R, Takano Y,
821 Tasumi E, Terada T, Tomaru H, Trembath-Reichert E, Wang DT, Yamada Y.

2015. Exploring deep microbial life in coal-bearing sediment down to ~2.5 km below the ocean floor. Science 349:420–424.

67. D'Hondt S, Pockalny R, Spivack AJ, Fulfer VM. 2019. Subseafloor life and its biogeochemical impacts. Nature Communications 10:1–13.

68. Morono Y, Terada T, Kallmeyer J, Inagaki F. 2013. An improved cell separation technique for marine subsurface sediments: applications for high-throughput analysis using flow cytometry and cell sorting. Environmental Microbiology 15:2841–2849.

69. Batzke A, Engelen B, Sass H, Cypionka H. 2007. Phylogenetic and physiological diversity of cultured deep-biosphere bacteria from Equatorial Pacific Ocean and Peru Margin sediments. Geomicrobiology Journal 24:261–273.

70. Biddle JF, Sylvan JB, Brazelton WJ, Tully BJ, Edwards KJ, Moyer CL, Heidelberg JF, Nelson WC. 2012. Prospects for the study of evolution in the deep biosphere. Frontiers in Microbiology 2:1–7.

71. McKay L, Klokman VW, Mendlovitz HP, Larowe DE, Hoer DR, Albert D, Amend JP, Teske A. 2016. Thermal and geochemical influences on microbial biogeography in the hydrothermal sediments of Guaymas Basin, Gulf of California. Environmental Microbiology Reports 8:150–161.

72. Teske A, De Beer D, McKay L, Tivey MK, Biddle JF, Hoer D, Lloyd KG, Lever MA, Røy H, Albert DB, Mendlovitz HP, MacGregor BJ. 2016. The Guaymas Basin hiking guide to hydrothermal mounds, chimneys, and microbial mats: Complex seafloor expressions of subsurface hydrothermal circulation. Frontiers in Microbiology 7.

843 73. D'Angelo T, Goordial J, Poulton NJ, Seyler L, Huber JA, Stepanauskas R, Orcutt BN.
844 2022. Oceanic Crustal Fluid Single Cell Genomics Complements Metagenomic and
845 Metatranscriptomic Surveys With Orders of Magnitude Less Sample Volume. Frontiers in
846 Microbiology 12:1–13.

847 74. Orcutt BN, Sylvan JB, Knab NJ, Edwards KJ. 2011. Microbial Ecology of the Dark Ocean
848 above, at, and below the Seafloor. Microbiology and Molecular Biology Reviews 75:361–422.

849 75. Meier DV, Bach W, Girguis PR, Gruber-Vodicka HR, Reeves EP, Richter M, Vidoudez
850 C, Amann R, Meyerdierks A. 2016. Heterotrophic Proteobacteria in the vicinity of diffuse
851 hydrothermal venting. Environmental Microbiology 18:4348–4368.

852 76. Dong X, Rattray JE, Campbell DC, Webb J, Chakraborty A, Adebayo O, Matthews S, Li
853 C, Fowler M, Morrison NM, MacDonald A, Groves RA, Lewis IA, Wang SH, Mayumi D,
854 Greening C, Hubert CRJ. 2020. Thermogenic hydrocarbon biodegradation by diverse depth-
855 stratified microbial populations at a Scotian Basin cold seep. Nat Commun 11:5825.

856 77. Wee SY, Edgcomb VP, Beaudoin D, Yvon-Lewis S, Sylvan JB. 2021. Microbial
857 Abundance and Diversity in Subsurface Lower Oceanic Crust at Atlantis Bank, Southwest Indian
858 Ridge. Applied and Environmental Microbiology 87.

859 78. Zhou Z, St. John E, Anantharaman K, Reysenbach A-L. 2022. Global patterns of diversity
860 and metabolism of microbial communities in deep-sea hydrothermal vent deposits. Microbiome
861 10:241.

79. Callac N, Rommevaux-Jestin C, Rouxel O, Lesongeur F, Liorzou C, Bollinger C, Ferrant A, Godfroy A. 2013. Microbial colonization of basaltic glasses in hydrothermal organic-rich sediments at Guaymas Basin. Frontiers in Microbiology 4.

80. L'Haridon S, Miroshnichenko ML, Kostrikina NA, Tindall BJ, Spring S, Schumann P, Stackebrandt E, Bonch-Osmolovskaya EA, Jeanthon C. 2006. Vulcanibacillus modesticaldus gen. nov., sp. nov., a strictly anaerobic, nitrate-reducing bacterium from deep-sea hydrothermal vents. International Journal of Systematic and Evolutionary Microbiology 56:1047–1053.

81. Guo L-Y, Ling S-K, Li C-M, Chen G-J, Du Z-J. 2018. Bacillus marinisedimentorum sp. nov., isolated from marine sediment. Int J Syst Evol Microbiol 68:198–203.

82. Bell E, Rattray JE, Sloan K, Sherry A, Pilloni G, Hubert CRJ. 2022. Hyperthermophilic endospores germinate and metabolize organic carbon in sediments heated to 80°C. Environmental Microbiology n/a.

83. Berdugo-Clavijo C, Dong X, Soh J, Sensen CW, Gieg LM. 2012. Methanogenic biodegradation of two-ringed polycyclic aromatic hydrocarbons. FEMS Microbiology Ecology 81:124–133.

84. Edgcomb VP, Teske AP, Mara P. 2022. Microbial Hydrocarbon Degradation in Guaymas Basin—Exploring the Roles and Potential Interactions of Fungi and Sulfate-Reducing Bacteria. Front Microbiol 13.

85. Trembath-Reichert E, Morono Y, Ijiri A, Hoshino T, Dawson KS, Inagaki F, Orphan VJ. 2017. Methyl-compound use and slow growth characterize microbial life in 2-km-deep subseafloor coal and shale beds. Proceedings of the National Academy of Sciences 114:E9206–E9215.

883 86. Zheng R, Wang C, Liu R, Cai R, Sun C. 2024. Physiological and metabolic insights into
884 the first cultured anaerobic representative of deep-sea Planctomycetes bacteria. eLife 12:RP89874.

885 87. Sekiguchi Y, Yamada T, Hanada S, Ohashi A, Harada H, Kamagata Y. 2003. Anaerolinea
886 thermophila gen. nov., sp. nov. and Caldilinea aerophila gen. nov., sp. nov., novel filamentous
887 thermophiles that represent a previously uncultured lineage of the domain Bacteria at the
888 subphylum level. International Journal of Systematic and Evolutionary Microbiology 53:1843–
889 1851.

890 88. Lee JZ, Burow LC, Woebken D, Everroad RC, Kubo MD, Spormann AM, Weber PK, Pett-
891 Ridge J, Bebout BM, Hoehler TM. 2014. Fermentation couples Chloroflexi and sulfate-reducing
892 bacteria to Cyanobacteria in hypersaline microbial mats. Front Microbiol 5.

893 89. Schweitzer HD, Smith HJ, Barnhart EP, McKay LJ, Gerlach R, Cunningham AB,
894 Malmstrom RR, Goudeau D, Fields MW. 2022. Subsurface hydrocarbon degradation strategies in
895 low- and high-sulfate coal seam communities identified with activity-based metagenomics. npj
896 Biofilms Microbiomes 8:7.

Table 1. Depth and temperature data for BONCAT incubations.BD = below detection. Geochemistry data are compiled from IODP 385 site chapters as published by Mara *et al.* (2023). Sediment temperatures are based on linear interpolation from published temperature gradients

Hole	Sample ID	Depth (mbsf)	Temp (°C)	Alkalinity (mM)	SO_4^{2-} (mM)	PO_4^{3-} (µM)	H_2S (mM)	NH_4^+ (mM)	CH_4 (mM)	CO (nM)	DOC (mg/L)	DIC (mM)	TOC (wt%)	TN (wt%)	TOC/TN
U1545B	1545B-19F3	155.0	39.8	35.3	0.3	46.9	3.2	15.9	1.9	116	48.2	10.7	2.17	0.34	7.5
U1546B	1546B-1H2	0.8	2.8	5.1	27.7	12.4	428	0.7	BD	665	21.8	2.2	4.36	0.52	9.7
U1547B	1547B-1H2	2.2	14.2	3	27.9	16.5	68	0.1	BD	350	13.9	1.2	3.76	0.45	9.7
U1547B	1547B-8H2	65.8	46.6	13	20.2	25.4	6606	2.9	BD	91	23.3	1.5	1.25	0.15	9.5
U1548B	1548B-2H3	8.9	13.7	4.8	26.8	24.5	732.5	0.3	BD	554	13.1	1.4	2.4	0.22	12.9
U1549B	1549B-3H2	16.5	6.4	18.7	17.5	98.1	2567	3.3	BD	188	23.6	5.7	3.12	0.59	6.1

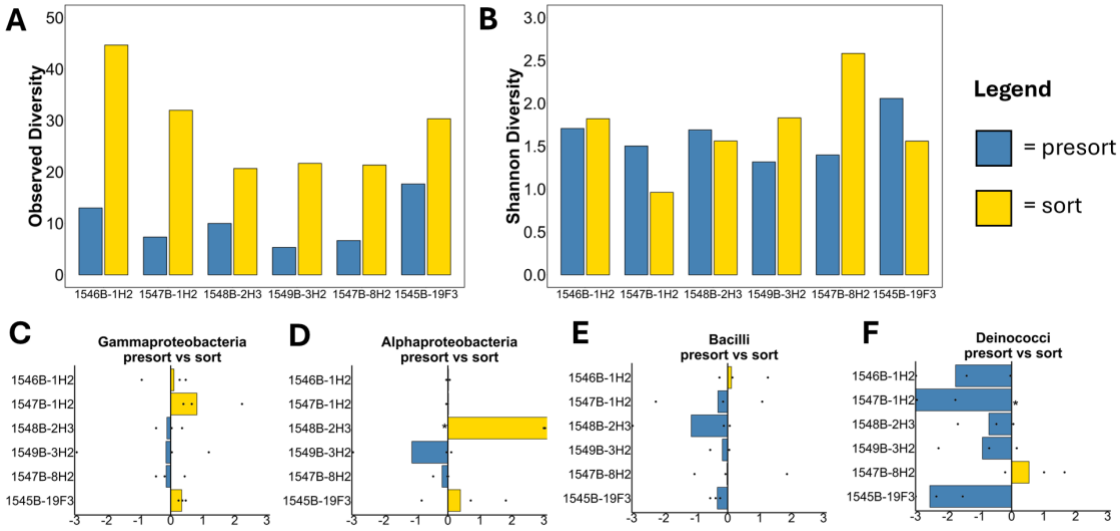

Figure 1. Alpha diversity indices of each sample and fold enrichment changes of the 4 most abundant Classes. Plots of Observed Diversity (A) and Shannon's diversity index (B) for all sites. Fold enrichment of relative abundance between the pre-sort and sorted fractions for Gammaproteobacteria (C), Alphaproteobacteria (D), Bacilli (E), and Deinococci (F). Fold enrichment was calculated as the ratio of relative abundances of ASVs affiliated with a given class between the pre-sort and sorted fraction, normalized to zero so that equal frequencies equate to zero. In fold-enrichment plots (C-F), yellow bars indicate that a class was more represented in the sorted fraction, and blue bars show that the class was more abundant in the pre-sort fraction. Bars represent the average of triplicate samples, and each point shows the value of each replicate. A * indicated that no ASVs affiliated with that class were detected in either the presort or sorted fraction.

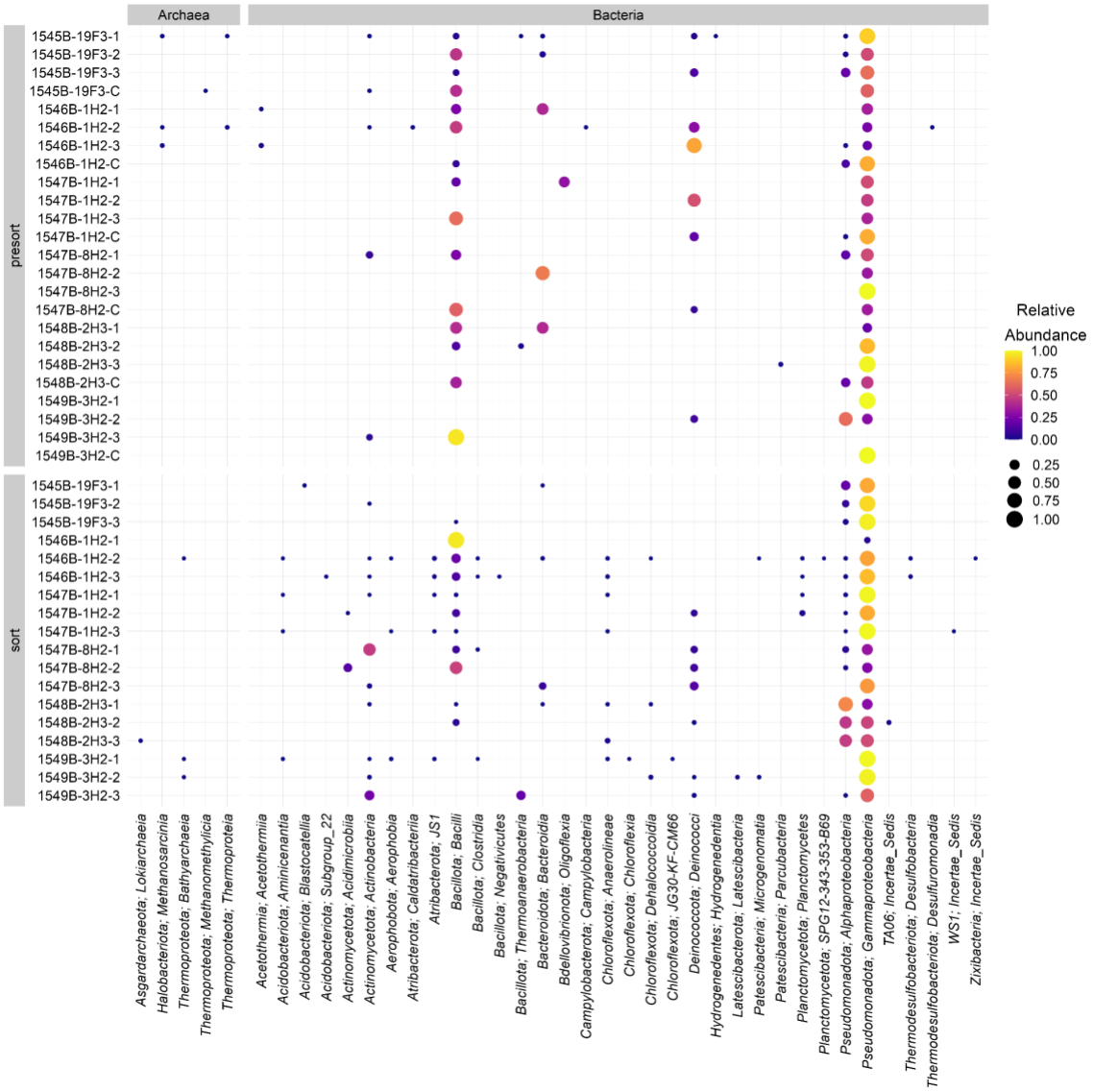

Figure 2. Community composition of cell extract (pre-sort) and BONCAT-positive sorted samples depicted as the relative abundance of ASVs of the 16S rRNA gene collapsed to class level. Fractions are grouped separately on the y-axis. Replicate numbers are indicated by 1-3; C refers to the no-HPG control

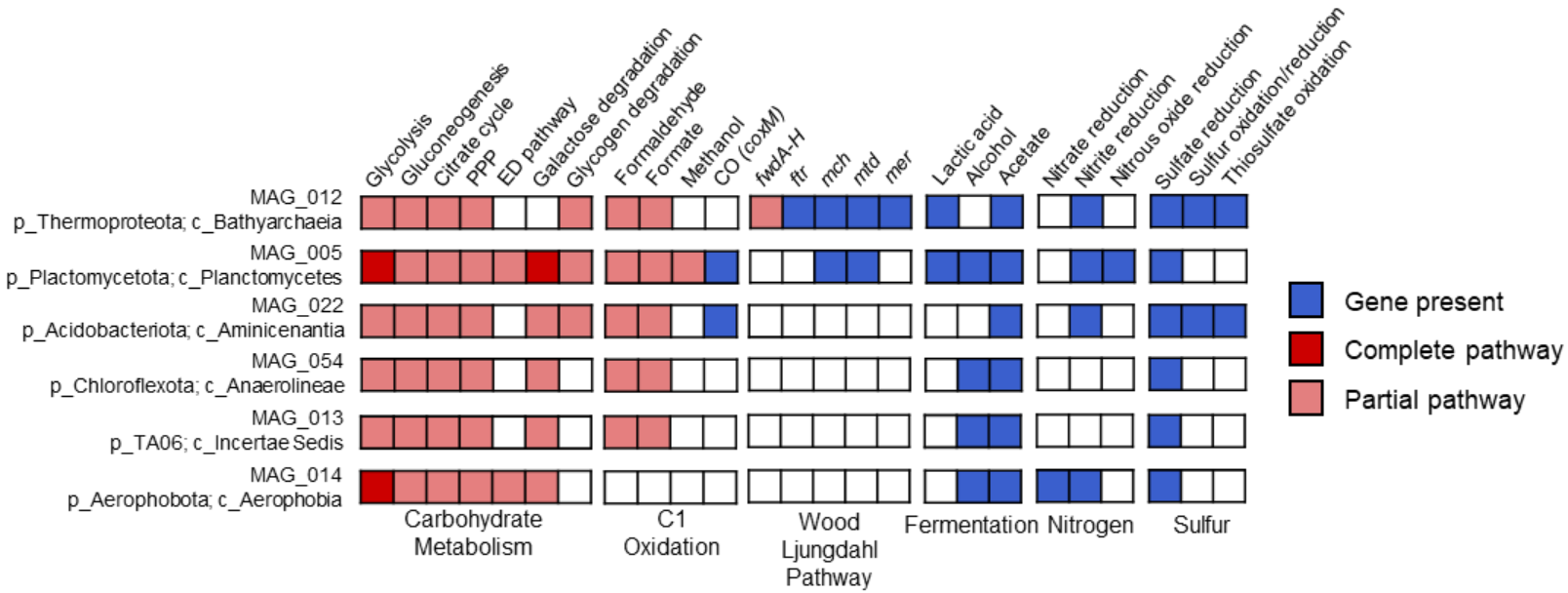

Figure 3. Heat map of metabolic potential of the six translationally active MAGs. A subset of genes was assessed. Squares indicate the presence of a functional gene (blue), complete metabolic pathway (red), or partial metabolic pathway (pink). White squares indicate that no gene(s) affiliated with that pathway were detected. MAGs are numbered following identities in (Mara et al., 2023). Abbreviations: PPP (pentose phosphate pathway), ED pathway (Entner-Doudoroff pathway). MAG_012 (84.65% complete, 9.60% contamination), MAG_005 (95.91% complete, 7.52% contamination), MAG_022 (87.01% complete, 4.23% contamination), MAG_054 (52.66% complete, 3.54% contamination), MAG_013 (86.40% complete, 0.78% contamination), MAG_014 (88.80% complete, 8.25% contamination).

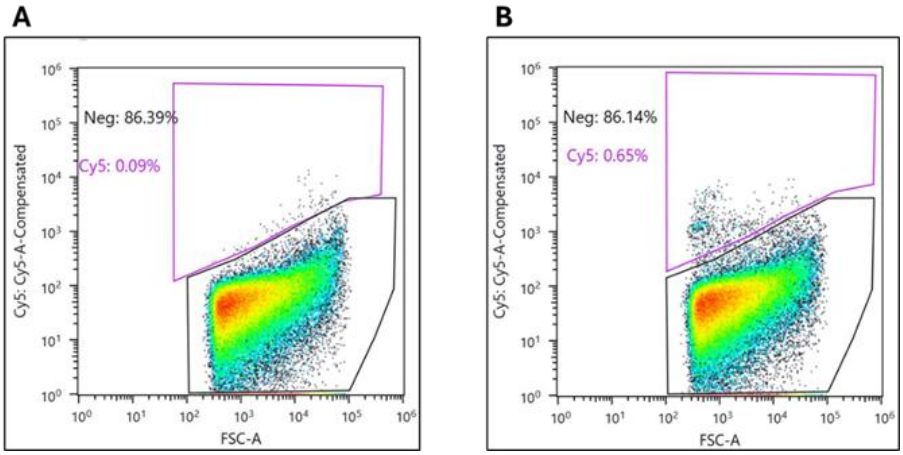

Figure S1. Representative FACS plots for 1545B-19F3 No HPG control (A) and 1545B-19F3 Replicate A (B). No-HPG control samples were used to draw the positive (Cy5; purple polygon) and negative (Neg; black polygon) gates for each sample set. The Cy5 positive gate represents BONCAT positive events and is sorted into collection tubes for downstream analysis. The percent values represent the fraction of the total events in each gate.

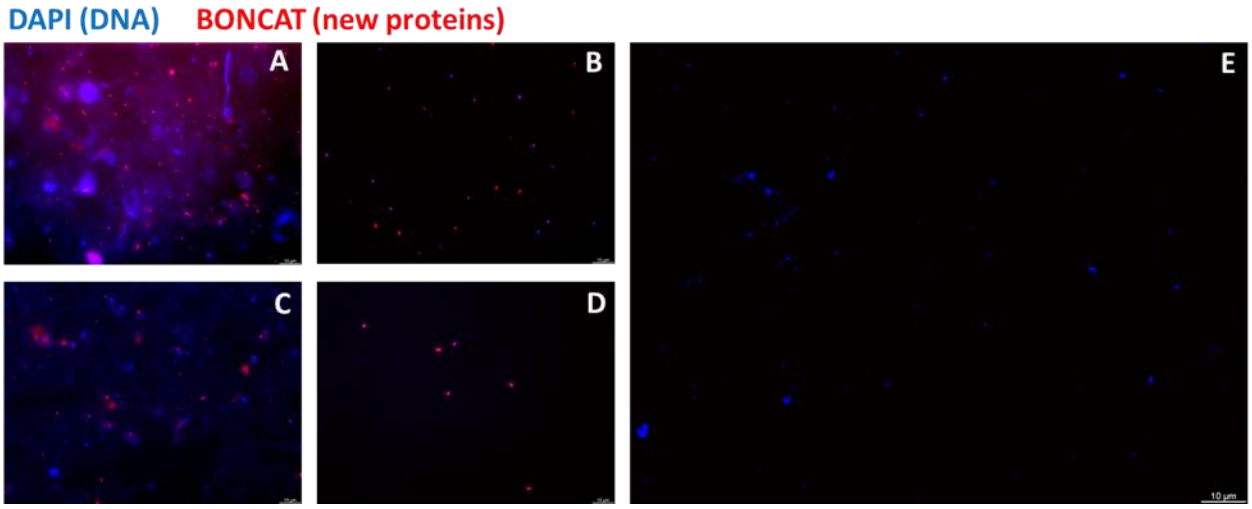

Figure S2. Fluorescence microscopy images of IODP 385 BONCAT samples. Images from 1545B-19F3 Replicate A pre-sort (A) and sorted (B), 1548B-3H2 Replicate B pre-sort (C) and sorted (D), and 1548B-3H2 no HPG control pre-sort (E). All cells are stained with the general nucleic acid stain 4′,6-diamidino-2-phenylindole (DAPI) in blue. Cells that incorporated HPG were stained with Cy5 using click chemistry and are depicted as red. Note that the pre-sort samples (A and C) contain cells with and without HPG incorporation, while sorted samples only contain cells that had incorporated HPG (red). Some sediment leftover from cell extraction appears blue, either from autofluorescence or non-specific DAPI staining. Scale bars depicted on each image are 10 µm.

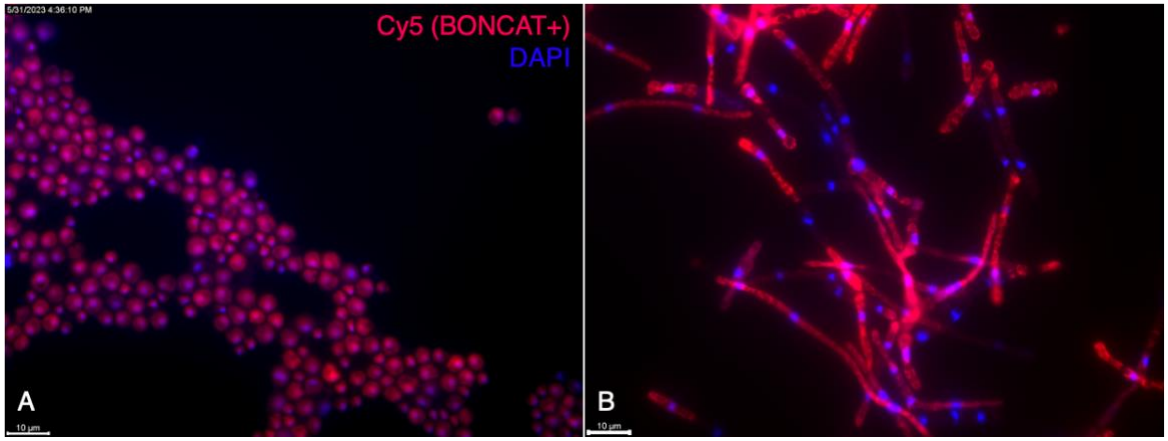

Figure S3. Fluorescence microscopy image of fungal isolates showing successful BONCAT labeling. *Rhodotorula sp.* (A) and *Cadophora sp.* (B) were grown with HPG in liquid media, preserved with GlyTE, and subjected to click chemistry to identify HPG incorporation. The red stain indicates BONCAT-positive cells labelled with Cy5, and the blue stain represents the general DNA stain DAPI. The scale bar is 10 µm.

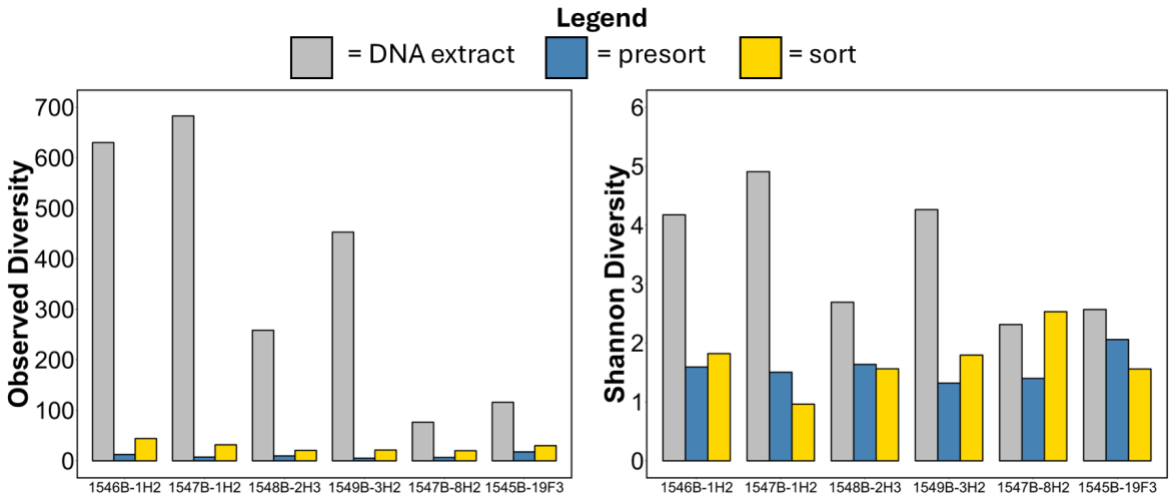

Figure S4. Alpha diversity indices of each sample. Observed diversity (A) and Shannon's diversity (B) for DNA extract, cell extracted pre-sort, and BONCAT-sorted fractions from all samples.

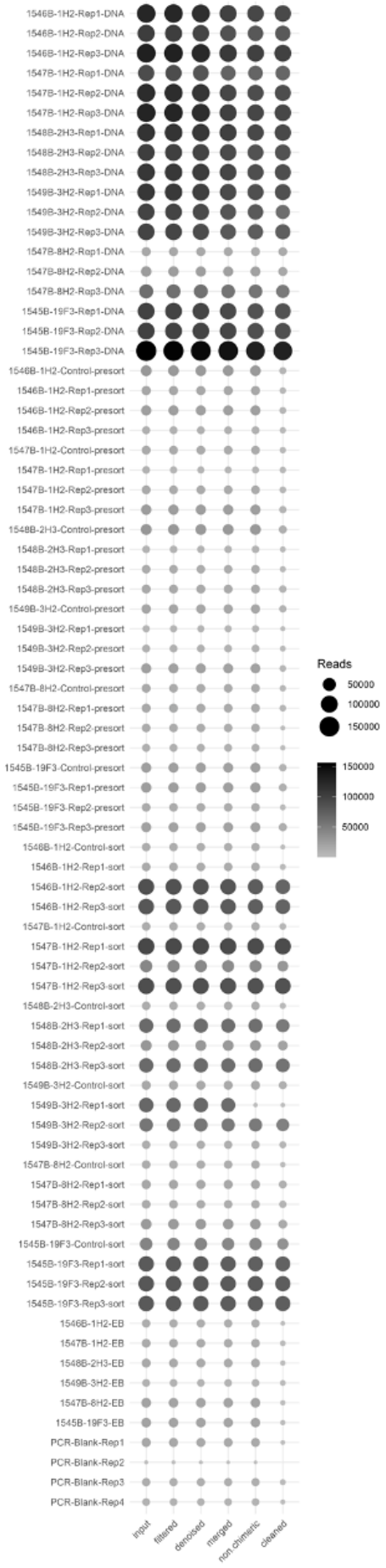
1546B-1H2-Rep1-DNA
1546B-1H2-Rep2-DNA
1546B-1H2-Rep3-DNA
1547B-1H2-Rep1-DNA
1547B-1H2-Rep2-DNA
1547B-1H2-Rep3-DNA
1548B-2H3-Rep1-DNA
1548B-2H3-Rep2-DNA
1548B-2H3-Rep3-DNA
1549B-3H2-Rep1-DNA
1549B-3H2-Rep2-DNA
1549B-3H2-Rep3-DNA
1547B-8H2-Rep1-DNA
1547B-8H2-Rep2-DNA
1547B-8H2-Rep3-DNA
1545B-19F3-Rep1-DNA
1545B-19F3-Rep2-DNA
1545B-19F3-Rep3-DNA
1546B-1H2-Control-presort
1546B-1H2-Rep1-presort
1546B-1H2-Rep2-presort
1546B-1H2-Rep3-presort
1547B-1H2-Control-presort
1547B-1H2-Rep1-presort
1547B-1H2-Rep2-presort
1547B-1H2-Rep3-presort
1548B-2H3-Control-presort
1548B-2H3-Rep1-presort
1548B-2H3-Rep2-presort
1548B-2H3-Rep3-presort
1549B-3H2-Control-presort
1549B-3H2-Rep1-presort
1549B-3H2-Rep2-presort
1549B-3H2-Rep3-presort
1547B-8H2-Control-presort
1547B-8H2-Rep1-presort
1547B-8H2-Rep2-presort
1547B-8H2-Rep3-presort
1545B-19F3-Control-presort
1545B-19F3-Rep1-presort
1545B-19F3-Rep2-presort
1545B-19F3-Rep3-presort
1546B-1H2-Control-sort
1546B-1H2-Rep1-sort
1546B-1H2-Rep2-sort
1546B-1H2-Rep3-sort
1547B-1H2-Control-sort
1547B-1H2-Rep1-sort
1547B-1H2-Rep2-sort
1547B-1H2-Rep3-sort
1548B-2H3-Control-sort
1548B-2H3-Rep1-sort
1548B-2H3-Rep2-sort
1548B-2H3-Rep3-sort
1549B-3H2-Control-sort
1549B-3H2-Rep1-sort
1549B-3H2-Rep2-sort
1549B-3H2-Rep3-sort
1547B-8H2-Control-sort
1547B-8H2-Rep1-sort
1547B-8H2-Rep2-sort
1547B-8H2-Rep3-sort
1545B-19F3-Control-sort
1545B-19F3-Rep1-sort
1545B-19F3-Rep2-sort
1545B-19F3-Rep3-sort
1546B-1H2-EB
1547B-1H2-EB
1548B-2H3-EB
1549B-3H2-EB
1547B-8H2-EB
1545B-19F3-EB
PCR-Blank-Rep1
PCR-Blank-Rep2
PCR-Blank-Rep3
PCR-Blank-Rep4
input
filtered
denoised
merged
non chimeric
cleaned
Reads
50000
100000
150000
150000
100000
50000

Figure S5. Sequence processing for 16S rRNA gene amplicons. Samples are grouped by fraction (DNA extract, cell extracted fraction/presort, sorted, and controls) and depth. The size and color of points reflect the number of reads at each processing step. Cleaned = number of reads after processing and removal of contaminant sequences. Control replicates for presort and sort samples are the no-HPG controls. Extraction blanks (EB) were extracted alongside samples and then treated as presort samples. Sorted control replicates, extraction blanks, and PCR blanks were used to assess contaminants and were removed from the dataset prior to community analysis.

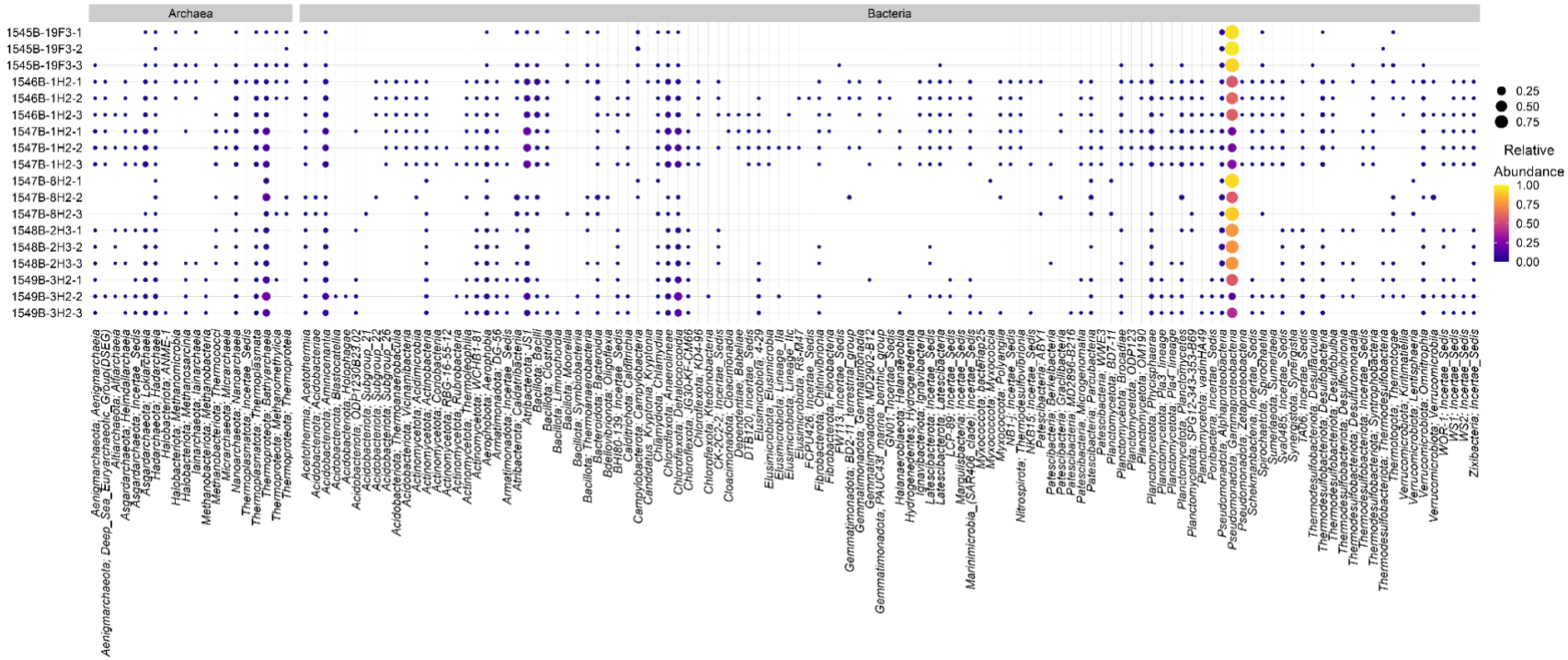

Figure S6. Relative abundance of ASVs from 16S rRNA gene amplicon sequencing from whole sediment DNA extracts. Replicates are denoted Rep1-3.

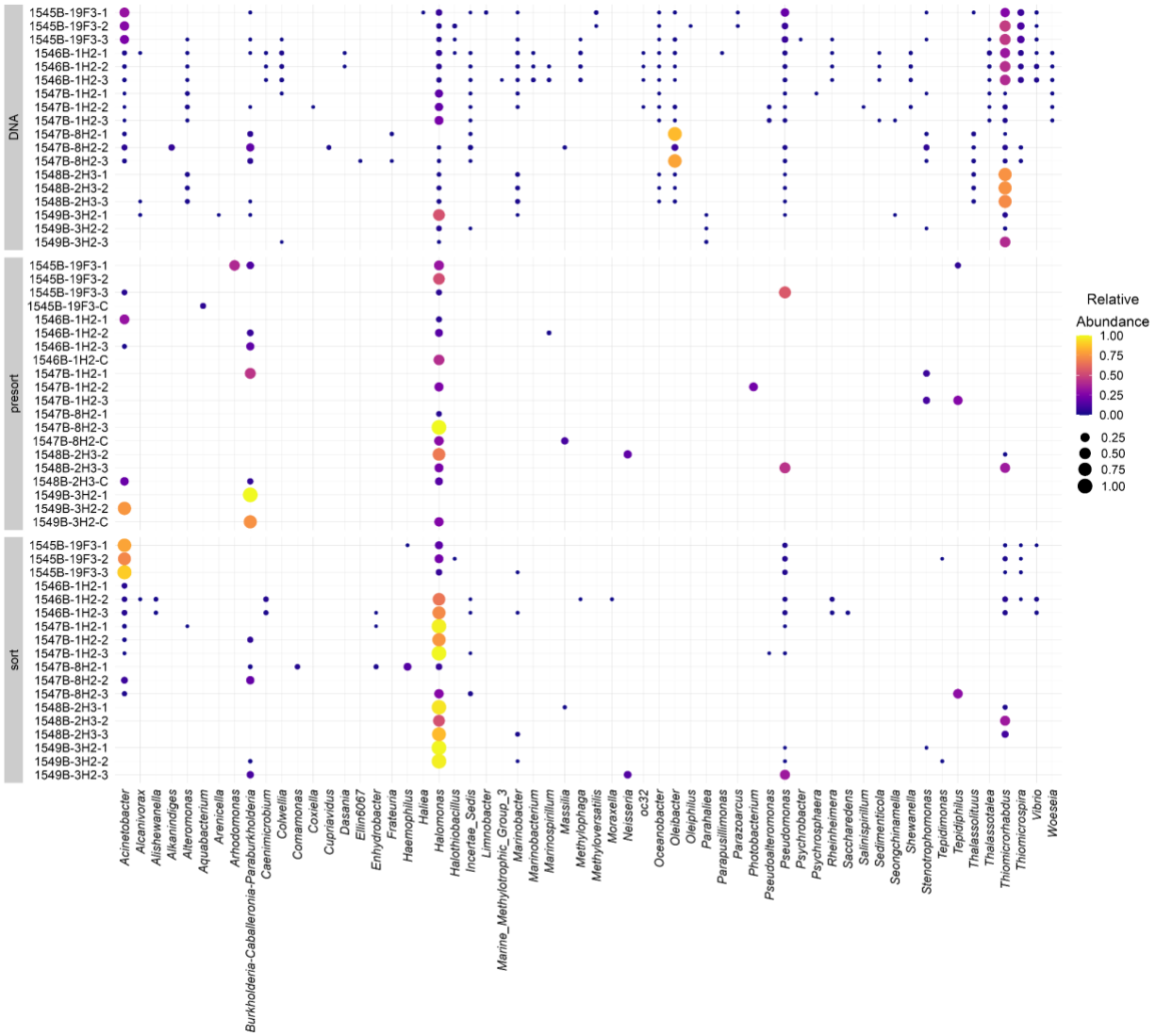

Figure S7. Relative abundance of genus-level ASVs, affiliated with the Gammaproteobacteria, from 16S rRNA gene amplicon sequencing for each fraction. Fractions are faceted on the y-axis for the DNA extract (DNA), cell extract (presort), and BONCAT positive fractions (sort). Replicate numbers are 1-3; C refers to the no-HPG control.

Table S1. Sorting metrics from FACS for each sample. The number of events sorted for each sample and replicate is listed. The percent of total events is the number of positive, sorted events out of the total number of events detected by FACS.

Sample	Replicate	Sorted BONCAT+ events	Percent of total events
1546B-1H2	1	97991	0.69%
1546B-1H2	2	208025	1.31%
1546B-1H2	3	119443	0.58%
1545B-19F3	1	136468	0.63%
1545B-19F3	2	135643	0.59%
1545B-19F3	3	143970	0.57%
1547B-1H2	1	132251	0.55%
1547B-1H2	2	266329	0.77%
1547B-1H2	3	398515	0.91%
1549B-3H2	1	723234	6.72%
1549B-3H2	2	8567	0.06%
1549B-3H2	3	7786	0.06%
1548B-3H2	1	54497	0.28%
1548B-3H2	2	33976	0.18%
1548B-3H2	3	38075	0.15%
1547B-8H2	1	12239	0.06%
1547B-8H2	2	6959	0.06%
1547B-8H2	3	8718	0.08%