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 **Introduction**

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

53

54 Uncovering the metabolic capabilities of constituents of microbial communities *in situ* requires a
55 coordinated effort that simultaneously identifies taxa and links them to ecosystem functions.
56 Historically, cultured isolates were required in order to draw conclusions about the metabolic
57 capabilities of specific microbial taxa. However, less than 1% of microbes can be reliably
58 maintained in a laboratory setting [9, 10], and enrichment and isolation conditions can differ
59 substantially from *in situ* conditions, making it difficult to draw conclusions about how microbes
60 live in their natural habitat. Further, cultivation efforts are impeded by low biomass inocula and
61 energy limitations in the deep biosphere that select for microorganisms that grow at extremely
62 slow rates [11, 12]. More recently, culture-independent techniques have been used to elucidate
63 connections between individual microorganisms and their metabolic functions under (near) *in situ*
64 conditions. With the rise of high-throughput sequencing technologies and bioinformatics tools,
65 culture-independent techniques have facilitated studies of the deep biosphere with metagenomics

66 [13–15] to identify microbial taxa, or metatranscriptomics [16–20] to reveal which genes are
67 expressed *in situ*. However, while these approaches offer powerful insights into metabolic
68 potential and community level activity, they are unable to resolve activity at the single cell level
69 and thus cannot link specific taxa with metabolic activity *in situ*.
70
71 Bioorthogonal non-canonical amino acid tagging (BONCAT) paired with fluorescence-activated
72 cell sorting (BONCAT-FACS) is a taxonomy-blind Next-Generation Physiology approach [21]
73 that can be used to identify which taxa were translationally active in complex environmental
74 samples during a relatively short incubation without culturing individual organisms [22–26].
75 BONCAT leverages the incorporation of a synthetic amino acid, such as *L*-homopropargylglycine
76 (HPG), an analog of *L*-methionine, into newly synthesized proteins and its later detection using
77 azide-alkyne click chemistry [22, 27]. Using fluorescence as the mode of detection, FACS can be
78 used to sort translationally active cells and study their identity and genetic make-up via 16 rRNA
79 gene amplicon sequencing [21, 28, 29] or shotgun metagenomics [30] and sorted cells can be
80 subjected to cultivation [22, 31]. BONCAT-FACS has been used to identify active taxa in marine
81 sediments [28], soils [26, 32], the human microbiome [33–35], and other complex environments
82 [29, 36–38]. BONCAT has been compared directly to two gold standards of measuring single cell
83 activity - radioactive isotope probing coupled to microautoradiography [39] and stable isotope
84 probing coupled to Nano-scale secondary ion mass spectrometry [22] - and was found to correlate
85 well with both methods. BONACT coupled to fluorescence microscopy can successfully detect
86 anabolic activity of an *Escherichia coli* culture after incubation with *L*-azidohomoalanine for 2%
87 of its generation time [22] and detect the activity of microbial consortia catalyzing the anaerobic
88 oxidation of methane in deep sea sediments after incubation with HPG for approximately 3.7% of

89 the estimated generation time [9, 10]. Here, we used BONCAT-FACS to identify the
90 translationally active community and, coupled with previously generated community shotgun
91 metagenomics datasets, assess the metabolic capacity of translationally active taxa in the Guaymas
92 Basin deep biosphere.
93
94 Guaymas Basin, located in the Gulf of California, is a hydrothermally active rift basin
95 characterized by rapid seafloor spreading and high sedimentation rates. Terrestrial inputs and
96 highly productive waters in the Gulf of California deposit organic matter on the seafloor at rates
97 >1 mm yr^{-1}, generating sediment packages that can be hundreds of meters thick [40].
98 Thermocatalytic alteration of the organic matter, influenced by basaltic sill intrusions and
99 hydrothermal circulation [41, 42], produces a complex mixture of potentially bioavailable organic
100 compounds in the sediment [43]. This includes methane, low-molecular-weight organics,
101 hydrocarbons, and crude oil [44–49]. International Ocean Discovery Program Expedition 385
102 (IODP 385) drilled the subsurface of the Guaymas Basin to investigate how microbial communities
103 drive organic carbon cycling in the hydrothermally altered deep biosphere. In total, holes were
104 drilled at eight sites (U1545-U1552) to capture unaltered sediment and a spectrum of hydrothermal
105 sediments including extremely hot sediments above a more recently intruded sill-sediment system
106 [50]. To explore the microbial ecology of the Guaymas Basin subsurface, we tracked
107 translationally active taxa in six samples, from six depths at five sites drilled during IODP 385.
108 We were limited in the number of samples available for this work from difficult to access sites,
109 but were able to identify translationally active taxa from these samples. We optimized a previously
110 established BONCAT-FACS workflow with 16S rRNA gene amplicon sequencing to identify
111 translationally active taxa [28, 29] and to assess diversity across biogeochemical gradients. We

112 coupled this with metagenomic analyses from IODP 385 [14] to provide direct insight into the
113 metabolic capacity of active taxa within the Guaymas Basin deep biosphere.
114
115 **Materials and Methods**
116 *Sample collection and BONCAT incubation*
117 Sediment cores were collected from Guaymas Basin during IODP expedition 385 onboard drilling
118 vessel *JOIDES Resolution.* Samples were taken at six depths, ranging from 0.76 to 154 mbsf, from
119 five drill holes (Table 1). Intact, round samples were aseptically transferred to the microbiology
120 lab after retrieval on deck and stored at 4°C in heat-sealed gas-tight bags flushed with N_2 gas to
121 maintain anoxic conditions for up to 30 minutes until processing. Material for metagenome
122 analysis was immediately frozen at -80°C. While we cannot rule out the chance of drilling fluid or
123 laboratory contamination in these samples, extensive efforts to minimize the likelihood of this are
124 described in recent metagenomic analyses performed on the same samples [14, 20]. Briefly, during
125 IDOP 385 a tracer, perfluoromethyldecalin, was introduced during drilling of holes at each site
126 intended for microbiology sampling. Any cores where detectable tracer was measured in the core
127 interior (the portion sampled for BONCAT) were not processed [20]. Working in an anoxic glove
128 chamber, sediments were taken from the interior of the core excluding 2 cm of the core perimeter,
129 homogenized with N_2-gassed sterile bottom water, and diluted three times to form a slurry. For
130 each sample, four replicate aliquots of 40 or 120 mL, depending on depth (shallow vs. deep), were
131 distributed into Pyrex bottles and sealed with a butyl stopper and aluminum crimp. The volume of
132 aliquots were larger for deeper samples in order to increase the likelihood of detecting enough
133 active cells for sorting. The amino acid analog *L*-homopropargylglycine (HPG; Vector
134 Laboratories, previously Click Chemistry Tools), a synthetic analog of *L*-methionine, was added

135 to three replicates to a final concentration of 50 µM. The fourth replicate, which served as a control,
136 was not amended with HPG. After HPG addition, samples were incubated at *in situ* temperature
137 (Table 1) for seven days.
138
139 . After seven days, samples were washed with 1x PBS by adding an equal volume of 1x PBS to
140 each sediment slurry. The mixture was briefly vortexed to mix, dispensed to 50ml centrifuge tubes,
141 centrifuged at 3,200 x g for 20 min to pellet the sediment, and the supernatant poured off. A 5mL
142 subsample was frozen in a cryovial at -80°C for whole sediment DNA extraction, and the
143 remaining sediment was amended with 10% glycerol (9:1 sediment to glycerol-TE), and stored at
144 -80°C until further processing. Back at Montana State University, subsequent steps of the
145 BONCAT-FACS workflow including cell extractions, cell sorting, and 16S rRNA gene amplicon
146 sequencing were done following modified protocols [28, 29]. The full details of the workflow are
147 described in the supplementary information.
148
149 *Analysis of Metagenome Assembled Genomes*
150 Metagenome assembled genomes (MAGs) were previously generated from IODP 385 [14]. In
151 short, trimmed reads from all drill holes were pooled and co-assembled using MEGAHIT 1.2.9,
152 and MAGs that were at least 50% complete and contained less than 10% contamination were then
153 used for downstream analyses. Taxa appearing in assembled MAGs that also appeared in
154 contamination controls (laboratory extraction blank and a drilling fluid control) were removed
155 from downstream analyses and are listed in Supplementary data of [14]. Taxonomic identities of
156 these MAGs were assigned using GTDB-Tk v2.1.0 [51]. MAG contamination and completeness
157 were checked using CheckM v1.2.2 [52]. Barrnap (v0.9) was used to identify 16S rRNA gene

158 sequences from these MAGs. The 16S rRNA gene sequences of these MAGs, if present, were then
159 compared to the 16S rRNA gene amplicon sequences from sorted samples generated in this study
160 for sequence similarity. From a BLAST nucleotide similarity comparison, we used a cutoff of 90%
161 identify match and 99% for hit length to identify sequences that matched.
162
163 The organism identities of these matches were then compared between the MAGs and amplicon
164 sequences to determine if they were identified as the same organism. Importantly, the MAG and
165 16S rRNA gene amplicon sequences were independently sequenced and taxonomy was assigned
166 using different databases (Silva 138.2 for 16S rRNA gene amplicon sequences, and GTDB-Tk
167 2.1.0 for MAGs), and taxonomic identification varied between the two. Organism identity was
168 matched based on nomenclature changes in the literature to the best of our ability. For consistency
169 across our datasets, the Silva 138.2 taxonomic identification is used throughout the manuscript.
170 We were able to successfully match six organisms out of the datasets that met our cutoff criteria
171 of 90% for %ID and 99% for %hit, five bacteria and one archaea.
172
173 MAGs were annotated with Prokka v1.14.6 [53] separately for archaea and bacteria. Individual
174 genes of interest for the environment were selected based on Dombrowski *et al.* [54]. Carbon
175 metabolism pathways were identified using KEGG BlastKOALA v3.1 [55]. Genes for the Wood-
176 Ljungdahl pathway were identified using an internal protein database generated from cultured
177 methanogens because of the presence of MAGs in the same class as other methanogens.
178
179 **Results & Discussion**
180 *Sampling Sites*

181 Samples for BONCAT incubations were collected from drill hole sites, U1545 to U1549, along a
182 transect across Guaymas Basin during IODP Expedition 385 (Table 1). Sites U1545 and U1546
183 are located ~52 km off-axis of the central spreading center, and include a relatively unaltered
184 reference (U1545) site and a more thermally altered (U1546) site with deep sediments (~350-430
185 mbsf) influenced by a now thermally-equilibrated ancient sill emplacement [50]. Sites U1547 and
186 U1548 were drilled within and adjacent to the much hotter Ringvent site. Ringvent is shaped by a
187 shallow, more recently emplaced sill, exhibits a sill-driven hydrothermal mound, and allows us to
188 compare other samples to a site with higher sediment alteration and elevated temperatures. . Site
189 U1549 targeted an off-axis cold seep near a subsurface sill, where methane seepage and hydrates
190 indicate persistent subsurface fluid migration. Sample depths ranged from 0.8 m at site U1546 to
191 155 m at U1545, and temperatures were between 2.8°C and 46.6°C (Table 1). Sediment and
192 porefluid geochemical data relevant to microbial activity are outlined in Table 1, and
193 comprehensive datasets are available in the IODP 385 expedition reports [56–59].
194
195 *Workflow development*
196 We expanded the applications of BONCAT by refining a workflow for detecting translational
197 activity of microbial communities in extremely low-biomass marine deep-biosphere sediments that
198 maximized cell recovery and minimized interference from sediment autofluorescence. To optimize
199 this workflow and test the efficacy of cell extraction protocols, we spiked IODP 385 sediment with
200 *Escherichia coli* (*E. coli*) cells that had been grown with HPG. These sediments had been preserved
201 analogously to the IODP 385 BONCAT samples and incubated without HPG. The presence of
202 BONCAT-positive *E. coli* in cell extracts was used to assess the efficacy of different cell extraction
203 protocols to maximize cell recovery. Additionally, we sought to minimize particulate matter in the

204 final cell suspension. Autofluorescent sediment particles can interfere with FACS analyses when
205 the signal overlaps with the target fluorophore and distorts forward and back scatter, complicating
206 the accurate detection of cells.
207
208 The final cell extraction procedure was adapted from an established protocol [28, 60] and consisted
209 of two primary cell separations. The first wash with methanol only was used to extract cells that
210 are more susceptible to lysis from detergents. The second step combined sonication with a
211 detergent mixture (sodium pyrophosphate, Tween80, and methanol) to detach cells from sediment
212 particles. A 50% Nycodenz gradient was used between detergent wash steps to phase-separate
213 detached cells from sediment. Based on the number of cells and amount of autofluorescent
214 sediment in the extracted suspension, we determined that pooling eight technical replicate extracts
215 was sufficient to extract cells from low-biomass samples for downstream processing with FACS.
216 Field samples analyzed in this study had in situ cell abundances that ranged from 10^9 and $10^8/cm^3$
217 (near-surface samples U1546B-1H2 and U1547B-1H2) to $10^4/cm^3$ (U1547B_8H2 at 65.8 mbsf).
218 Methods for these cell counts and data are shown in [14, 61]. The rapid decline in down-core cell
219 counts in the Guaymas Basin subsurface is consistent with decreasing DNA recovery [61]. We
220 visualized labeling of BONCAT-positive cells from click chemistry in the presort and sorted
221 fractions with fluorescence microscopy (Fig. S1). During method testing and sample analysis, the
222 presort fraction contained cells with and without HPG incorporation, while the sorted fraction only
223 contained HPG-positive cells, demonstrating the ability of our method to successfully sort
224 exclusively BONCAT-positive cells from low-biomass marine deep biosphere sediments.
225 BONCAT has been successfully used on dozens of bacterial and archaeal phyla in many
226 ecosystems, including the human microbiome [33, 34], marine sediments [23, 62] and water

227 column [63, 64], soil [26], wastewater sludge [65], hot springs [29], the terrestrial subsurface [66]
228 and, more specifically, Guaymas basin sediments [28]. However, to the best of our knowledge, it
229 had yet not been applied to fungi. Initially, we aimed to assess the translational activity of fungi
230 with BONCAT-FACS in this study, so we developed this cell extraction protocol to include fungal
231 cells. We used two fungal isolates from Guaymas Basin that are representative of the fungal
232 community in IODP 385 sediment [67], and tested extraction methods similar to our use of *E.coli*
233 above. Based on our testing, this cell extraction method can be used to separate yeast and
234 filamentous fungal cells from sediments. Furthermore, these fungal strains incorporate HPG in
235 liquid media and can be detected with click chemistry and fluorescence microscopy (Fig. S2),
236 demonstrating that BONCAT workflows can be applied to marine fungi. However, due to issues
237 with downstream processing, we could not reliably detect fungi in the presort and sorted fractions
238 with internal transcribed spacer (ITS) amplicon sequencing. While the exact reason for this issue
239 is unknown, poor extraction efficiency, low fungal biomass (or absence of fungi), the inability to
240 sort certain fungal cells with FACS, or poorly matched PCR primers for the ITS region in natural
241 fungal populations [68, 69] may have all impeded the successful detection of fungi in our samples.
242 Therefore, we chose to exclude any further discussion of fungi from this manuscript and focus on
243 bacteria and archaea inhabiting the Guaymas Basin deep biosphere.
244
245 *Diversity of the total, cell extractable, and active prokaryotic community*
246 We combined BONCAT-FACS with 16S rRNA gene amplicon sequencing of sorted cells to
247 identify translationally active bacteria and archaea in the Guaymas Basin deep biosphere. To assess
248 community compositions, the observed diversity (the quantity of unique ASVs) and Shannon's
249 diversity index were determined for each sample (Fig. 1A-B). For all samples, observed and

250 Shannon's diversity indices were higher in the whole sediment DNA extract fraction than in the
251 cell-extracted presort and BONCAT-active sorted fractions (Fig. S3). Diversity was generally
252 higher in the shallower sediments, reflecting higher cell counts in those sediments. . The two
253 deepest samples, 1547B-8H2 and 1545B-19F3 (66 and 155 mbsf, respectively), exhibited the
254 smallest difference in diversity metrics across all three fractions, especially with respect to the
255 whole sediment DNA extract, taken after incubations but before cell extraction. Compared to
256 shallower samples, a larger portion of the total microbial community was detected in the BONCAT
257 active fraction in these deeper sediments, where there are relatively few cells, carbon is more
258 recalcitrant, energetic conditions are less favorable, and geochemical conditions remain stable over
259 extended timescales [70, 71]. This suggests that the microbial population is highly specialized to
260 fill limited available niches in the deep biosphere.
261
262 To evaluate the changes in diversity metrics between the presort and sorted fractions, we calculated
263 fold enrichment (Fig. 1C-F) between these fractions for the most abundant ASVs, which were
264 affiliated with *Gammaproteobacteria*, *Alphaproteobacteria*, *Bacilli*, and *Deinococci* (Fig. 2 and
265 Table S2). Generally, these abundant taxa were not overly represented in either fraction
266 (enrichment values near zero) or were enriched in the presort fraction (*Deinococci* and *Bacilli*),
267 suggesting that only a portion of those taxa detected in the presort fraction were actually
268 translationally active. In several cases, corresponding ASVs were not detected in either the presort
269 or sorted fraction, which skewed enrichment toward the respective other fraction. For instance, in
270 sample 1548B-2H3, no sequences affiliated with *Alphaproteobacteria* (Fig. 1D) were detected in
271 the presort fraction, artificially inflating the fold enrichment toward the sorted fraction. Similarly,
272 no *Deinococci* affiliated ASVs were observed in the sorted fraction for one replicate from 1547B-

273 1H2 (Fig. 1F). The inability to detect a given ASV in a specific fraction could skew diversity
274 indices, making it difficult to interpret diversity data in those samples.
275
276 Observed diversity of unique ASVs increased from the presort to the sorted fractions in all samples
277 (Fig. 1A), and Shannon's diversity increased in 1549B-3H2 and 1547B-8H2 between those
278 fractions (Fig. 1B). It should be expected that diversity would decrease between the presort and
279 the BONCAT active fraction to reflect that only a portion of the community is translationally active
280 and detectable with BONCAT-FACS. One explanation for this outcome is that there were
281 substantially fewer reads in many of the presort samples than in the sorted samples (Fig. S4), likely
282 due to the relatively small subsample (50 µL) taken from the presort samples for DNA extraction
283 and the low amount of biomass in the sample. More ASVs were detected when the number of
284 reads was higher (Fig. S4), which would influence alpha diversity metrics. In the two samples in
285 which Shannon's diversity increased in the sorted fraction (1547B-8H2 and 1549B-3H2), across
286 all replicates, we detected between three and eleven unique class-level ASVs in the presort samples
287 and between 12 and 30 in the sorted samples. This indicates that sequencing depth directly impacts
288 the detection of additional and/or rare taxa that may have been, in some cases, missed in the presort
289 fraction.
290
291 We observed major differences in Shannon's diversity indices across the BONCAT workflow
292 (DNA extract to presort to sorted fractions). Shannon's diversity was highest in the DNA extract
293 relative to the presort or sorted fractions for nearly every sample (the only exception was the sorted
294 fraction for 1547B-8H2) (Fig. S3). This is expected for the marine subsurface as extracellular DNA
295 in the sediments may have been amplified with PCR, altering community composition and

296 artificially inflating diversity metrics in the initial DNA extract fraction [26]. Diversity indices
297 were generally lower in the presort and sorted fraction as compared to the DNA extract fraction,
298 as cell extraction for BONCAT-FACS is known to introduce biases [28, 29] that could cause large
299 discrepancies in diversity metrics between the DNA extract and presort fractions. To address this,
300 we optimized the cell extraction procedure for these low-biomass deep biosphere samples, but
301 biased cell recovery is a challenge for any complex sample [21, 72], especially from Guaymas
302 Basin sediments that are enriched with hydrothermally altered organic carbon. Ideally, cell
303 extraction procedures should be tuned to specific samples based on experimental goals and
304 biogeochemical conditions.
305
306 *Abundant Taxa and Active Community*
307 To assess the active community within the Guaymas Basin deep biosphere, we analyzed the
308 relative abundances of ASVs in each fraction and compared those frequencies between samples
309 (Fig. 2 and Fig. S5). In the presort cell extract and BONCAT active sorted fractions, we identified
310 a total of 40 bacterial classes and six archaeal classes, which were frequently dominated by a few
311 classes. Across all three fractions (DNA extract, presort, and sorted), ASVs affiliated with
312 *Gammaproteobacteria* were consistently the most abundant taxa detected. *Gammaproteobacteria*
313 had greater than 50% relative abundance within the BONCAT active community in 12 of the sorted
314 fraction samples (Fig. 2). In three of the six samples (1546B-1H2, 1547B-1H2, and 1545B-19F3),
315 all *Gammaproteobacterial* genera were detected in higher relative abundance in the sorted fraction
316 than in the presort fraction (Fig. 2), which indicates members of this group are active in the
317 Guaymas Basin deep biosphere. High abundances of *Gammaproteobacteria* in the BONCAT
318 active fraction likely reflect the heterotrophic nature of this class, which is capable of degrading

319 diverse organic substrates typically found in Guaymas Basin sediments [73–75]. While the relative
320 frequency of *Gammaproteobacteria* was elevated in most samples, differences in relative
321 abundance between fractions could be explained by enrichment of *Gammaproteobacteria* during
322 BONCAT incubation, or methodological biases against other taxonomic groups.
323
324 Across all three fractions, we identified 61 genera affiliated with the *Gammaproteobacteria* (Fig.
325 S6 and Table S3) that are associated with diverse heterotrophic metabolisms. As observed for the
326 whole community analysis, the highest number of gammaproteobacterial genera was observed in
327 the DNA extract, and ASVs affiliated with the genus *Thiomicrohabdus* were amongst the most
328 abundant. In contrast, the presort and sorted fractions were primarily dominated by members of
329 *Halomonas,* and to a lesser extent, *Burkholderia-Caaballerionia-Paraburkholderia* and
330 *Acinetobacter*, which had the highest relative frequency in the deepest sample (1545B-19F3). In
331 several samples, ASVs belonging to *Halomonas* (1546B-1H2, 1547B-1H2, and 1549B-3H2) and
332 *Acinetobacter* (1545B-19F3) exhibited a relative frequency >75% in the BONCAT active fraction
333 in at least two replicates. This was a sharp increase compared to the other fractions which, across
334 all of those samples, were typically less than 20% for the DNA extract and all less than 3% for the
335 presort fractions. The relative abundance of *Halomonas* demonstrated the largest observed
336 increase between the presort and sorted fractions (Fig. S6). In particular, BONCAT active
337 *Halomonas* ASVs were detected at elevated frequencies in samples with temperatures <15°C
338 (Table 1), compared to warmer samples (~38°C). Sediment temperatures can fluctuate by >50°C
339 on relatively small temporal and spatial scales due to hydrothermal activity [76], which may create
340 a selective pressure on the microbial community. This indicates that some members of the
341 *Halomonas* genus remain metabolically active at lower temperatures and withstand large

342 temperature changes characteristic of hydrothermal sediments. Genomic analyses of these
343 communities revealed genes associated with central carbon metabolism, CAZymes, hydrocarbon
344 degradation, and nitrogen, sulfur, and arsenic cycling, suggesting that *Gammaproteobacteria* are
345 key drivers of organic matter remineralization and redox transformations in the subsurface [75,
346 77].
347
348 In addition to *Gammaproteobacteria*, abundant ASVs in the sorted fraction were affiliated with
349 *Alphaproteobacteria* throughout our samples. The relative frequency of translationally active
350 *Alphaproteobacteria* was between 28-52% for three replicates in the relatively warm (~14°C)
351 sediment sample from 1548B-2H3 and up to 19% in the deepest sample from 155 mbsf at the more
352 thermally moderate site 1545B (~39°C). After BONCAT incubation, *Alphaproteobacteria*
353 exhibited higher relative frequencies in the sorted fraction in 4 samples compared to the presort
354 fraction (Fig. 1D). Some of the supposed increase in frequency between fractions was due to low
355 detection of *Alphaproteobacteria* ASVs in the presort fraction. Regardless, the presence of
356 *Alphaproteobacteria* in the sorted fractions demonstrates that members of this class are active in
357 the Guaymas Basin deep biosphere. Similar to our observations for the *Gammaproteobacteria*,
358 this is consistent with previous studies that have demonstrated high abundances of
359 *Alphaproteobacteria* and associated heterotrophic activity in the marine deep biosphere [73, 78–
360 80].
361
362 Members of the class *Deinococci* were detected in all fractions, and consistently exhibited the
363 highest relative abundance in the presort fraction compared to the DNA extract and active sorted
364 fractions (Fig. 2 and Fig. S5). At Ringvent, in sample 1547B-1H2 (2.2 mbsf), *Deinococci* ASVs

365 constituted up to 54% of the presort fraction, compared to less than 1% in the DNA extract. This
366 higher frequency in the presort fraction suggests our cell extraction method is biased in favor of
367 *Deinococci* cells, which may influence relative abundances in the BONCAT active fraction. In
368 sample 1547B-8H2 (66 mbsf), members of this class accounted for up to 15% relative abundance
369 in the BONCAT active fraction and were less than 6% in the corresponding presort fraction. In
370 hydrothermally impacted habitats, *Deinococci*-affiliated sequences have been detected via 16S
371 rRNA gene amplicon survey [81]. Additionally, analysis of 52 MAGs assigned to *Deinococci* from
372 IODP 360 to Atlantis Bank revealed genes associated with broad anaerobic and heterotrophic
373 metabolisms, including fatty acid and aromatic compound degradation (*e.g.*, acyl-CoA
374 dehydrogenase and benzoyl-CoA), acetogenesis, nitrogen and sulfur cycling, and metal reduction
375 [82]. This versatility can promote *Deinococci* metabolism across the dynamic geochemical
376 gradients observed in the Guaymas Basin deep biosphere (Table 1).
377
378 A total of 7 classes affiliated with phylum *Firmicutes* were detected throughout our samples, and
379 the most abundant ASVs from this phylum were assigned to the *Bacilli* class in the presort and
380 active sorted fractions. Although ASVs affiliated with *Bacilli* were detected throughout the
381 sediment, the highest relative abundances in the BONCAT sorted fraction, up to ~20% (one
382 replicate was above 95%), were observed in the shallower samples (1546B-1H2 and 1547B-1H2)
383 (Fig. 2). In those shallower depths, above the sulfate methane transition zone, where sulfate is the
384 dominant terminal electron acceptor (Table 1) for anaerobic metabolisms, BONCAT-active ASVs
385 affiliated with *Firmicutes* may couple sulfate reduction to organic matter degradation [83].
386 *Firmicutes* isolates cultured from several marine deep biosphere sites degrade a broad suite of
387 organic substrates, including acetate, polysaccharides, amino acids, alcohols, nucleotides, and

388 microbial necromass, indicating a key role in heterotrophic carbon cycling [28, 73, 84, 85]. This
389 metabolic capacity was substantiated by metagenomic analysis that also demonstrated high
390 thermal tolerance, exceeding 80°C [86]. Additionally, members of the *Firmicutes* phylum within
391 the *Clostridia* class were associated with degradation of crude oil and polycyclic aromatic
392 hydrocarbon degradation in methanogenic enrichments [87]. Hydrothermally induced pyrolysis of
393 organic matter in Guaymas Basin sediments generates petroleum hydrocarbons that are dispersed
394 throughout the sediment column and can fuel diverse microbial metabolisms [45, 46, 88]. In the
395 Guaymas Basin specifically, *Firmicutes* have been linked to both sulfate and nitrate reduction [83]
396 and have dominated enrichment cultures amended with methanol, monomethylamine, and
397 ammonium [89]. Given the wide metabolic capabilities of the *Firmicutes* and the high frequency
398 in the BONCAT active fraction, members of this phylum likely contribute significantly to organic
399 matter degradation and elemental cycling in Guaymas Basin sediments.
400
401 *Implications for deep biosphere carbon cycling*
402 The BONCAT active fractions were dominated by a heterotrophic microbial population capable
403 of metabolizing a wide range of organic carbon in the Guaymas Basin deep biosphere. To further
404 assess this metabolic capacity and explore potential implications for biogeochemical cycling in
405 hydrothermal sediments, we mapped 16S rRNA gene ASVs from the BONCAT active fraction to
406 annotated metagenome assembled genomes (MAGs) from IODP 385 [14]. Six unique MAGs
407 associated with the BONCAT active fraction were identified as members of 5 bacterial phyla
408 (*Planctomycetota*, *Chloroflexota*, *Acidobacteriota*, *Aerophobota*, and *TA06*) and 1 archaeal
409 phylum (*Thermoproteota*). These MAGs were 52.7-95.9% complete and all had contamination
410 below 10% (Table S4).

411

412 The translationally active MAGs harbored genes associated with diverse carbon cycling pathways,
413 including carbohydrate degradation, C1 metabolism, and fermentation, plus nitrogen and sulfur
414 cycling (Fig. 3). MAG_005 (95.91% complete, 9.60% contamination) was assigned to the phylum
415 *Planctomycetota* and encodes complete pathways for glycolysis and galactose degradation, and a
416 nitrite reductase gene. Recently, the first anaerobic culture of *Planctomycetota* isolated from
417 marine sediments experimentally demonstrated functional capacity for these metabolisms [90],
418 validating the genomic potential of that isolate and MAG_005. MAG_054 (52.66% complete,
419 3.54% contamination), assigned to the class *Anaerolineae*, contained genes associated with
420 carbohydrate degradation and fermentation, both of which have been demonstrated by cultured
421 *Anaerolineae* [91]. Similarly, NanoSIMS experiments implicated diverse *Chloroflexota* (phylum
422 of *Anaerolineae*) in fermentation and broad heterotrophic metabolisms in hypersaline
423 photoheterotrophic mats [92]. In their experiments, Lee *et al*. (2014) ascertained ^{13}C-substrate
424 assimilation by *Chloroflexota* members using phylum-specific FISH probes, but could not resolve
425 taxonomic affiliations on class level or below, or implicate other phyla in carbon cycling [92]. In
426 contrast, BONCAT-FACS workflows provide a high-throughput community-level analysis of
427 active taxa, and when coupled with metagenomics, support the investigation of the metabolic
428 potential of sorted cells [66]. Taken together, BONCAT-FACS, paired with metagenomics, can
429 corroborate the results from culture-dependent and -independent approaches, while also providing
430 metabolic insights at the community level.

431

432 Most importantly, metagenomic analysis, corroborated by 16S rRNA gene identities, revealed a
433 diverse, translationally active microbial community in the Guaymas Basin deep biosphere with the

434 metabolic capacity for a range of carbon transformations and broad anaerobic metabolisms. Our
435 results confirm that these taxa are active, and that under the right biogeochemical conditions (*e.g.,*
436 substrate availability, temperature, etc.), the organisms represented by these MAGs likely utilize
437 these metabolisms for energy conservation in the deep biosphere. However, experimental evidence
438 is necessary to confirm which of these metabolisms are used *in situ* under different conditions.
439 Additionally, the specific ASVs affiliated with these MAGs were not necessarily amongst the most
440 abundant or widespread taxa in the BONCAT sorted fraction. Thus, the extent to which these
441 microbes drive carbon transformations is yet unknown. Future experiments that combine
442 BONCAT-FACS with substrate amendment [28, 29] and paired metagenomics [66] and/or
443 metatranscriptomics could provide deeper, community-level insights into the metabolic
444 capabilities of translationally active taxa in the marine deep biosphere.
445

446 **Conclusion**

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

457 of heterotrophic metabolisms and drive microbial-mediated recycling of hydrothermally altered
458 organic matter in the Guaymas Basin deep biosphere.
459
460 We demonstrated that BONCAT-FACS can be used to assess microbial activity in low-biomass
461 marine sediments on a community level. Additionally, we paired this with metagenomic analysis
462 from IODP 385 to identify six translationally active MAGs and assess their metabolic capacity.
463 BONCAT-FACS coupled metagenomics links single-cell activity and genetic potential to
464 understand community-level dynamics in a high-throughput manner [21, 66]. This approach could
465 be further optimized in the future to generate metagenomes directly from the sorted fraction and/or
466 by combining experimental amendments with whole sediment metagenomics to more tightly link
467 taxa with metabolic function. Importantly, our workflow, while optimized for low-biomass
468 samples, was likely biased against rare taxa, given the decrease in diversity between fractions.
469 These biases could be exacerbated if combined with metagenomics, which would require greater
470 quantities of biomass for FACS and sequencing compared to gene amplicon sequencing. We were
471 limited in this study by the difficulty of accessing the study site, resulting in a small number of
472 samples. However, we were still able to apply this methodology to deep biosphere samples
473 successfully. With careful benchmarking and optimization, BONCAT-FACS and metagenomics,
474 when combined, offer a powerful approach for connecting microbial taxa and metabolic function
475 in the marine deep biosphere and other complex environmental samples.
476
477 **Acknowledgments**
478 We would like to thank the scientific party and D/V JOIDES Resolution crew from IODP 385 for
479 their assistance in sample collection. Amplicon sequencing was performed by the Molecular

Research Core Facility at Idaho State University (RRID:SCR_012598). This work was supported 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).

503 **References**

504 1. Whitman WB, Coleman DC, Wiebe WJ. Prokaryotes: The unseen majority. *Proceedings of the*
505 *National Academy of Sciences* 1998;**95**:6578–6583. https://doi.org/10.1073/pnas.95.12.6578
506 2. Kallmeyer J et al. Global distribution of microbial abundance and biomass in subseafloor
507 sediment. *Proceedings of the National Academy of Sciences of the United States of America*
508 2012;**109**:16213–16216. https://doi.org/10.1073/pnas.1203849109
509 3. Oremland RS, Culbertson C, Simoneit BRT. Methanogenic activity in sediment from Leg 64,
510 Gulf of California. *Initial reports DSDP, Leg 64, Mazatlan to Long Beach, 1978-79 Part 2*
511 1982;759–762. https://doi.org/10.2973/dsdp.proc.64.122.1982
512 4. D'Hondt S et al. Metabolic Activity of Subsurface Life in Deep-Sea Sediments. *Science*
513 2002;**295**:2067–2070.
514 5. Lever MA. Acetogenesis in the energy-starved deep biosphere-a paradox? *Frontiers in*
515 *Microbiology* 2012;**2**:1–18. https://doi.org/10.3389/fmicb.2011.00284
516 6. Teske A, Callaghan AV, LaRowe DE. Biosphere frontiers of subsurface life in the sedimented
517 hydrothermal system of Guaymas Basin. *Frontiers in Microbiology* 2014;**5**:1–11.
518 https://doi.org/10.3389/fmicb.2014.00362
519 7. Santelli CM et al. Abundance and diversity of microbial life in ocean crust. *Nature*
520 2008;**453**:653–656. https://doi.org/10.1038/nature06899
521 8. Jørgensen BB. Deep subseafloor microbial cells on physiological standby. *Proceedings of the*
522 *National Academy of Sciences* 2011;**108**:18193–18194.
523 https://doi.org/10.1073/pnas.1115421108
524 9. Lloyd KG et al. Phylogenetically Novel Uncultured Microbial Cells Dominate Earth
525 Microbiomes. *mSystems* 2018;**3**. https://doi.org/10.1128/mSystems.00055-18

526 10. Steen AD et al. High proportions of bacteria and archaea across most biomes remain
527 uncultured. *The ISME Journal* 2019. https://doi.org/10.1038/s41396-019-0484-y
528 11. Lever MA et al. Trends in basalt and sediment core contamination during IODP expedition
529 301. *Geomicrobiology Journal* 2006;**23**:517–530.
530 https://doi.org/10.1080/01490450600897245
531 12. Orcutt BN et al. Microbial activity in the marine deep biosphere: Progress and prospects.
532 *Frontiers in Microbiology* 2013;**4**:1–15. https://doi.org/10.3389/fmicb.2013.00189
533 13. Gittins DA et al. Geological processes mediate a microbial dispersal loop in the deep biosphere.
534 *Science Advances* 2022;**8**:eabn3485. https://doi.org/10.1126/sciadv.abn3485
535 14. Mara P et al. Metagenomic profiles of archaea and bacteria within thermal and geochemical
536 gradients of the Guaymas Basin deep subsurface. *Nat Commun* 2023;**14**:7768.
537 https://doi.org/10.1038/s41467-023-43296-x
538 15. Sun J et al. Metagenomic insights into taxonomic and functional patterns in shallow coastal
539 and deep subseafloor sediments in the Western Pacific. *Microbial Genomics* 2025;**11**:001351.
540 https://doi.org/10.1099/mgen.0.001351
541 16. Orsi WD et al. Gene expression in the deep biosphere. *Nature* 2013;**499**:205–208.
542 https://doi.org/10.1038/nature12230
543 17. Coskun ÖK et al. Linking Uncultivated Microbial Populations and Benthic Carbon Turnover
544 by Using Quantitative Stable Isotope Probing. *Applied and Environmental Microbiology*
545 2018;**84**. https://doi.org/10.1128/AEM.01083-18
546 18. Coskun ÖK et al. Quantifying population-specific growth in benthic bacterial communities
547 under low oxygen using H218O. *The ISME Journal* 2019;**13**. https://doi.org/10.1038/s41396-
548 019-0373-4

549 19. Mara P et al. Microbial gene expression in Guaymas Basin subsurface sediments responds to
550 hydrothermal stress and energy limitation. *ISME J* 2023;**17**:1907–1919.
551 https://doi.org/10.1038/s41396-023-01492-z
552 20. Mara P et al. Deep subseafloor sediments in Guaymas Basin harbor cosmopolitan microbiota
553 and traces of hydrothermal populations. *Commun Earth Environ* 2024;**5**:505.
554 https://doi.org/10.1038/s43247-024-01662-7
555 21. Hatzenpichler R et al. Next-generation physiology approaches to study microbiome function
556 at single cell level. *Nature Reviews Microbiology* 2020;**18**:241–256.
557 https://doi.org/10.1038/s41579-020-0323-1
558 22. Hatzenpichler R et al. In situ visualization of newly synthesized proteins in environmental
559 microbes using amino acid tagging and click chemistry. *Environmental Microbiology*
560 2014;**16**:2568–2590. https://doi.org/10.1111/1462-2920.12436
561 23. Hatzenpichler R et al. Visualizing in situ translational activity for identifying and sorting slow-
562 growing archaeal - bacterial consortia. *Proceedings of the National Academy of Sciences*
563 2016;**113**:E4069–E4078. https://doi.org/10.1073/pnas.1603757113
564 24. Samo TJ et al. Broad distribution and high proportion of protein synthesis active marine
565 bacteria revealed by click chemistry at the single cell level. *Frontiers in Marine Science*
566 2014;**1**:1–18. https://doi.org/10.3389/fmars.2014.00048
567 25. Leizeaga A et al. Using click-chemistry for visualizing in situ changes of translational activity
568 in planktonic marine bacteria. *Frontiers in Microbiology* 2017;**8**:1–11.
569 https://doi.org/10.3389/fmicb.2017.02360
570 26. Couradeau E et al. Probing the active fraction of soil microbiomes using BONCAT-FACS.
571 *Nature Communications* 2019;**10**. https://doi.org/10.1038/s41467-019-10542-0

572 27. Kiick KL et al. Incorporation of azides into recombinant proteins for chemoselective
573 modification by the Staudinger ligation. *Proceedings of the National Academy of Sciences*
574 2002;**99**:19–24. https://doi.org/10.1073/pnas.012583299
575 28. Krukenberg V et al. Microbial Community Response to Polysaccharide Amendment in Anoxic
576 Hydrothermal Sediments of the Guaymas Basin. *Frontiers in Microbiology* 2021;**12**.
577 https://doi.org/fig
578 29. Reichart NJ et al. Activity-based cell sorting reveals responses of uncultured archaea and
579 bacteria to substrate amendment. *The ISME Journal* 2020;**14**:2851–2861.
580 https://doi.org/10.1038/s41396-020-00749-1
581 30. Rinke C et al. Obtaining genomes from uncultivated environmental microorganisms using
582 FACS-based single-cell genomics. *Nature Protocols* 2014;**9**:1038–1048.
583 https://doi.org/10.1038/nprot.2014.067
584 31. Mulay SA et al. BONCAT-Live for isolation and cultivation of active environmental bacteria.
585 *mBio* 2025.
586 32. Harris JE et al. The activity of soil microbial taxa in the rhizosphere predicts the success of
587 root colonization. *mSystems* 2025;**0**:e00458-25. https://doi.org/10.1128/msystems.00458-25
588 33. Riva A et al. Conversion of Rutin, a Prevalent Dietary Flavonol, by the Human Gut Microbiota.
589 *Front Microbiol* 2020;**11**. https://doi.org/10.3389/fmicb.2020.585428
590 34. Valentini TD et al. Bioorthogonal non-canonical amino acid tagging reveals translationally
591 active subpopulations of the cystic fibrosis lung microbiota. *Nat Commun* 2020;**11**:2287.
592 https://doi.org/10.1038/s41467-020-16163-2

593 35. Taguer M, Shapiro BJ, Maurice CF. Translational activity is uncoupled from nucleic acid
594 content in bacterial cells of the human gut microbiota. *Gut Microbes* 2021;**13**:1903289.
595 https://doi.org/10.1080/19490976.2021.1903289
596 36. Chen L et al. Isolating and characterizing translationally active fraction of anammox
597 microbiota using bioorthogonal non-canonical amino acid tagging. *Chemical Engineering*
598 *Journal* 2021;**418**:129411. https://doi.org/10.1016/j.cej.2021.129411
599 37. Du Z, Behrens SF. Tracking *de novo* protein synthesis in the activated sludge microbiome
600 using BONCAT-FACS. *Water Research* 2021;**205**:117696.
601 https://doi.org/10.1016/j.watres.2021.117696
602 38. Trexler RV et al. BONCAT-FACS-Seq reveals the active fraction of a biocrust community
603 undergoing a wet-up event. *Front Microbiol* 2023;**14**.
604 https://doi.org/10.3389/fmicb.2023.1176751
605 39. Amano C et al. Single-cell heterotrophic activity in deep-ocean prokaryotic communities
606 quantified by BONCAT and microautoradiography. *ISME Commun* 2026;ycag038.
607 https://doi.org/10.1093/ismeco/ycag038
608 40. Calvert SE. Origin of Diatom-Rich, Varved Sediments from the Gulf of California. *The*
609 *Journal of Geology* 1966;**74**:546–565. https://doi.org/10.1086/627188
610 41. Einsele G et al. Intrusion of basaltic sills into highly porous sediments, and resulting
611 hydrothermal activity. *Nature* 1980;**283**:441–445.
612 42. Lonsdale P, Becker K. Hydrothermal plumes, hot springs, and conductive heat flow in the
613 Southern Trough of Guaymas Basin. *Earth and Planetary Science Letters* 1985;**73**:211–225.
614 https://doi.org/10.1016/0012-821X(85)90070-6

43. Lin YS et al. Near-surface heating of young rift sediment causes mass production and discharge of reactive dissolved organic matter. *Scientific Reports* 2017;**7**:1–10. https://doi.org/10.1038/srep44864

44. Simoneit BRT et al. Organic geochemistry of recent sediments from Guaymas Basin, Gulf of California. *Deep Sea Research Part A, Oceanographic Research Papers* 1979;**26**:879–891. https://doi.org/10.1016/0198-0149(79)90102-X

45. Kawka OE, Simoneit BRT. Survey of hydrothermally-generated petroleums from the Guaymas Basin spreading center. *Organic Geochemistry* 1987;**11**:311–328. https://doi.org/10.1016/0146-6380(87)90042-8

46. Bazylinski DA, Farrington JW, Jannasch HW. Hydrocarbons in surface sediments from a Guaymas Basin hydrothermal vent site. *Organic Geochemistry* 1988;**12**:547–558. https://doi.org/10.1016/0146-6380(88)90146-5

47. Whelan JK, Simoneit BRT, Tarafa ME. C1-C8 hydrocarbons in sediments from Guaymas Basin, Gulf of California-Comparison to Peru Margin, Japan Trench and California Borderlands. *Organic Geochemistry* 1988;**12**:171–194. https://doi.org/10.1016/0146-6380(88)90253-7

48. Martens CS. Generation of short chain acid anions in hydrothermally altered sediments of the Guaymas Basin, Gulf of California. *Applied Geochemistry* 1990;**5**:71–76. https://doi.org/10.1016/0883-2927(90)90037-6

49. Zhuang G et al. Generation and Utilization of Volatile Fatty Acids and Alcohols in Hydrothermally Altered Sediments in the Guaymas Basin, Gulf of California. *Geophysical Research Letters* 2019;**46**:2637–2646. https://doi.org/10.1029/2018GL081284

637 50. Teske A et al. 'Expedition 385 Summary' in: Guaymas Basin Tectonics and Biosphere.
638 International Ocean Discovery Program, 2021.
639 51. Chaumeil P-A et al. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy
640 Database. *Bioinformatics* 2020;**36**:1925–1927. https://doi.org/10.1093/bioinformatics/btz848
641 52. Chklovski A et al. CheckM2: a rapid, scalable and accurate tool for assessing microbial genome
642 quality using machine learning. *Nat Methods* 2023;1–10. https://doi.org/10.1038/s41592-023-
643 01940-w
644 53. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014;**30**:2068–
645 2069. https://doi.org/10.1093/bioinformatics/btu153
646 54. Dombrowski N et al. Genomic insights into potential interdependencies in microbial
647 hydrocarbon and nutrient cycling in hydrothermal sediments. *Microbiome* 2017;**5**:106.
648 https://doi.org/10.1186/s40168-017-0322-2
649 55. Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG Tools for
650 Functional Characterization of Genome and Metagenome Sequences. *Journal of Molecular*
651 *Biology* 2016;**428**:726–731. https://doi.org/10.1016/j.jmb.2015.11.006
652 56. Teske AP et al. Site U1545. In Teske, A., Lizarralde, D., Höfig, T. W., and the Expedition 385
653 Scientists, Guaymas Basin Tectonics and Biosphere. *Proceedings of the International Ocean*
654 *Discovery Program* 2021;**385**. https://doi.org/10.14379/iodp.proc.385.103.2021
655 57. Teske AP et al. Site U1546. In Teske, A., Lizarralde, D., Höfig, T. W., and the Expedition 385
656 Scientists, Guaymas Basin Tectonics and Biosphere. *Proceedings of the International Ocean*
657 *Discovery Program* 2021;**385**. https://doi.org/10.14379/iodp.proc.385.104.2021
658 58. Teske AP et al. Sites U1547 and U1548. In Teske, A., Lizarralde, D., Höfig, T. W., and the
659 Expedition 385 Scientists, Guaymas Basin Tectonics and Biosphere. *Proceedings of the*

660 *International Ocean Discovery Program* 2021;**385**.
661 https://doi.org/10.14379/iodp.proc.385.105.2021
662 59. Teske AP et al. Site U1549. In Teske, A., Lizarralde, D., Höfig, T. W., and the Expedition 385
663 Scientists, Guaymas Basin Tectonics and Biosphere. *Proceedings of the International Ocean*
664 *Discovery Program* 2021;**385**. https://doi.org/10.14379/iodp.proc.385.106.2021
665 60. Kallmeyer J et al. New cell extraction procedure applied to deep subsurface sediments.
666 *Limnology and Oceanography: Methods* 2008;**6**:236–245.
667 https://doi.org/10.4319/lom.2008.6.236
668 61. Ramírez GA et al. Data report: the DNA event horizon in the Guaymas Basin subsurface
669 biosphere: technical advances and redefined limits in bulk extractions of nucleic acids from
670 deep marine sediments, IODP Expedition 385. *Proceedings of the International Ocean*
671 *Discovery Program* 2024;**385**. https://doi.org/10.14379/iodp.proc.385.204.2024
672 62. Marlow J et al. Spatially resolved correlative microscopy and microbial identification reveal
673 dynamic depth- and mineral-dependent anabolic activity in salt marsh sediment.
674 *Environmental Microbiology* 2021;**00**:1462-2920.15667. https://doi.org/10.1111/1462-
675 2920.15667
676 63. Sebastián M et al. High growth potential of long-term starved deep ocean opportunistic
677 heterotrophic bacteria. *Frontiers in Microbiology* 2019;**10**:1–12.
678 https://doi.org/10.3389/fmicb.2019.00760
679 64. Martinez-Varela A et al. Bacterial responses to background organic pollutants in the northeast
680 subarctic Pacific Ocean. *Environ Microbiol* 2021;**23**:4532–4546.
681 https://doi.org/10.1111/1462-2920.15646

682 65. Madill MBW et al. Activity-Based Cell Sorting Reveals Resistance of Functionally Degenerate
683 Nitrospira during a Press Disturbance in Nitrifying Activated Sludge. *mSystems*
684 2021;**6**:10.1128/msystems.00712-21. https://doi.org/10.1128/msystems.00712-21
685 66. Schweitzer HD et al. Subsurface hydrocarbon degradation strategies in low- and high-sulfate
686 coal seam communities identified with activity-based metagenomics. *npj Biofilms*
687 *Microbiomes* 2022;**8**:7. https://doi.org/10.1038/s41522-022-00267-2
688 67. Keeler E et al. Deep-sea hydrothermal vent sediments reveal diverse fungi with antibacterial
689 activities. *FEMS Microbiology Ecology* 2021;**97**:fiab103.
690 https://doi.org/10.1093/femsec/fiab103
691 68. Taylor DL et al. Accurate Estimation of Fungal Diversity and Abundance through Improved
692 Lineage-Specific Primers Optimized for Illumina Amplicon Sequencing. *Applied and*
693 *Environmental Microbiology* 2016;**82**:7217–7226. https://doi.org/10.1128/AEM.02576-16
694 69. Tedersoo L et al. Shotgun metagenomes and multiple primer pair-barcode combinations of
695 amplicons reveal biases in metabarcoding analyses of fungi. *MycoKeys* 2015;**10**:1–43.
696 https://doi.org/10.3897/mycokeys.10.4852
697 70. Inagaki F et al. Exploring deep microbial life in coal-bearing sediment down to ∼2.5 km below
698 the ocean floor. *Science* 2015;**349**:420–424. https://doi.org/10.1126/science.aaa6882
699 71. D'Hondt S et al. Subseafloor life and its biogeochemical impacts. *Nature Communications*
700 2019;**10**:1–13. https://doi.org/10.1038/s41467-019-11450-z
701 72. Morono Y et al. An improved cell separation technique for marine subsurface sediments:
702 applications for high-throughput analysis using flow cytometry and cell sorting.
703 *Environmental Microbiology* 2013;**15**:2841–2849. https://doi.org/10.1111/1462-2920.12153

704 73. Batzke A et al. Phylogenetic and physiological diversity of cultured deep-biosphere bacteria
705 from Equatorial Pacific Ocean and Peru Margin sediments. *Geomicrobiology Journal*
706 2007;**24**:261–273. https://doi.org/10.1080/01490450701456453
707 74. Biddle JF et al. Prospects for the study of evolution in the deep biosphere. *Frontiers in*
708 *Microbiology* 2012;**2**:1–7. https://doi.org/10.3389/fmicb.2011.00285
709 75. McKay L et al. Thermal and geochemical influences on microbial biogeography in the
710 hydrothermal sediments of Guaymas Basin, Gulf of California. *Environmental Microbiology*
711 *Reports* 2016;**8**:150–161. https://doi.org/10.1111/1758-2229.12365
712 76. Teske A et al. The Guaymas Basin hiking guide to hydrothermal mounds, chimneys, and
713 microbial mats: Complex seafloor expressions of subsurface hydrothermal circulation.
714 *Frontiers in Microbiology* 2016;**7**. https://doi.org/10.3389/fmicb.2016.00075
715 77. D'Angelo T et al. Oceanic Crustal Fluid Single Cell Genomics Complements Metagenomic
716 and Metatranscriptomic Surveys With Orders of Magnitude Less Sample Volume. *Frontiers*
717 *in Microbiology* 2022;**12**:1–13. https://doi.org/10.3389/fmicb.2021.738231
718 78. Orcutt BN et al. Microbial Ecology of the Dark Ocean above, at, and below the Seafloor.
719 *Microbiology and Molecular Biology Reviews* 2011;**75**:361–422.
720 https://doi.org/10.1128/mmbr.00039-10
721 79. Meier DV et al. Heterotrophic Proteobacteria in the vicinity of diffuse hydrothermal venting.
722 *Environmental Microbiology* 2016;**18**:4348–4368. https://doi.org/10.1111/1462-2920.13304
723 80. Dong X et al. Thermogenic hydrocarbon biodegradation by diverse depth-stratified microbial
724 populations at a Scotian Basin cold seep. *Nat Commun* 2020;**11**:5825.
725 https://doi.org/10.1038/s41467-020-19648-2

81. Wee SY et al. Microbial Abundance and Diversity in Subsurface Lower Oceanic Crust at Atlantis Bank, Southwest Indian Ridge. *Applied and Environmental Microbiology* 2021;**87**. https://doi.org/10.1128/AEM.01519-21

82. Zhou Z et al. Global patterns of diversity and metabolism of microbial communities in deep-sea hydrothermal vent deposits. *Microbiome* 2022;**10**:241. https://doi.org/10.1186/s40168-022-01424-7

83. Callac N et al. Microbial colonization of basaltic glasses in hydrothermal organic-rich sediments at Guaymas Basin. *Frontiers in Microbiology* 2013;**4**.

84. L'Haridon S et al. Vulcanibacillus modesticaldus gen. nov., sp. nov., a strictly anaerobic, nitrate-reducing bacterium from deep-sea hydrothermal vents. *International Journal of Systematic and Evolutionary Microbiology* 2006;**56**:1047–1053. https://doi.org/10.1099/ijs.0.64012-0

85. Guo L-Y et al. Bacillus marinisedimentorum sp. nov., isolated from marine sediment. *Int J Syst Evol Microbiol* 2018;**68**:198–203. https://doi.org/10.1099/ijsem.0.002482

86. Bell E et al. Hyperthermophilic endospores germinate and metabolize organic carbon in sediments heated to 80°C. *Environmental Microbiology* 2022;**n/a**. https://doi.org/10.1111/1462-2920.16167

87. Berdugo-Clavijo C et al. Methanogenic biodegradation of two-ringed polycyclic aromatic hydrocarbons. *FEMS Microbiology Ecology* 2012;**81**:124–133. https://doi.org/10.1111/j.1574-6941.2012.01328.x

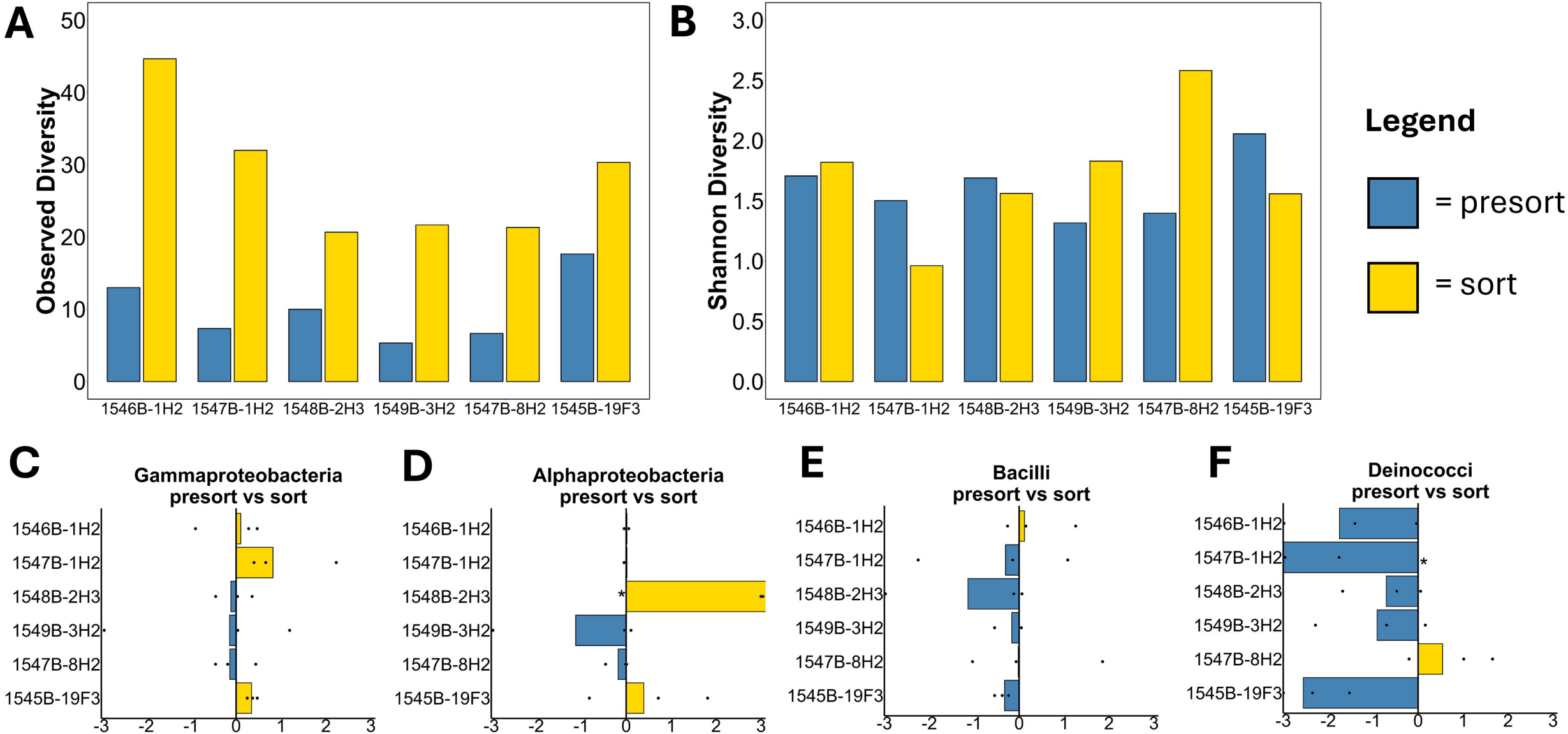
88. Edgcomb VP, Teske AP, Mara P. Microbial Hydrocarbon Degradation in Guaymas Basin—Exploring the Roles and Potential Interactions of Fungi and Sulfate-Reducing Bacteria. *Front Microbiol* 2022;**13**. https://doi.org/10.3389/fmicb.2022.831828

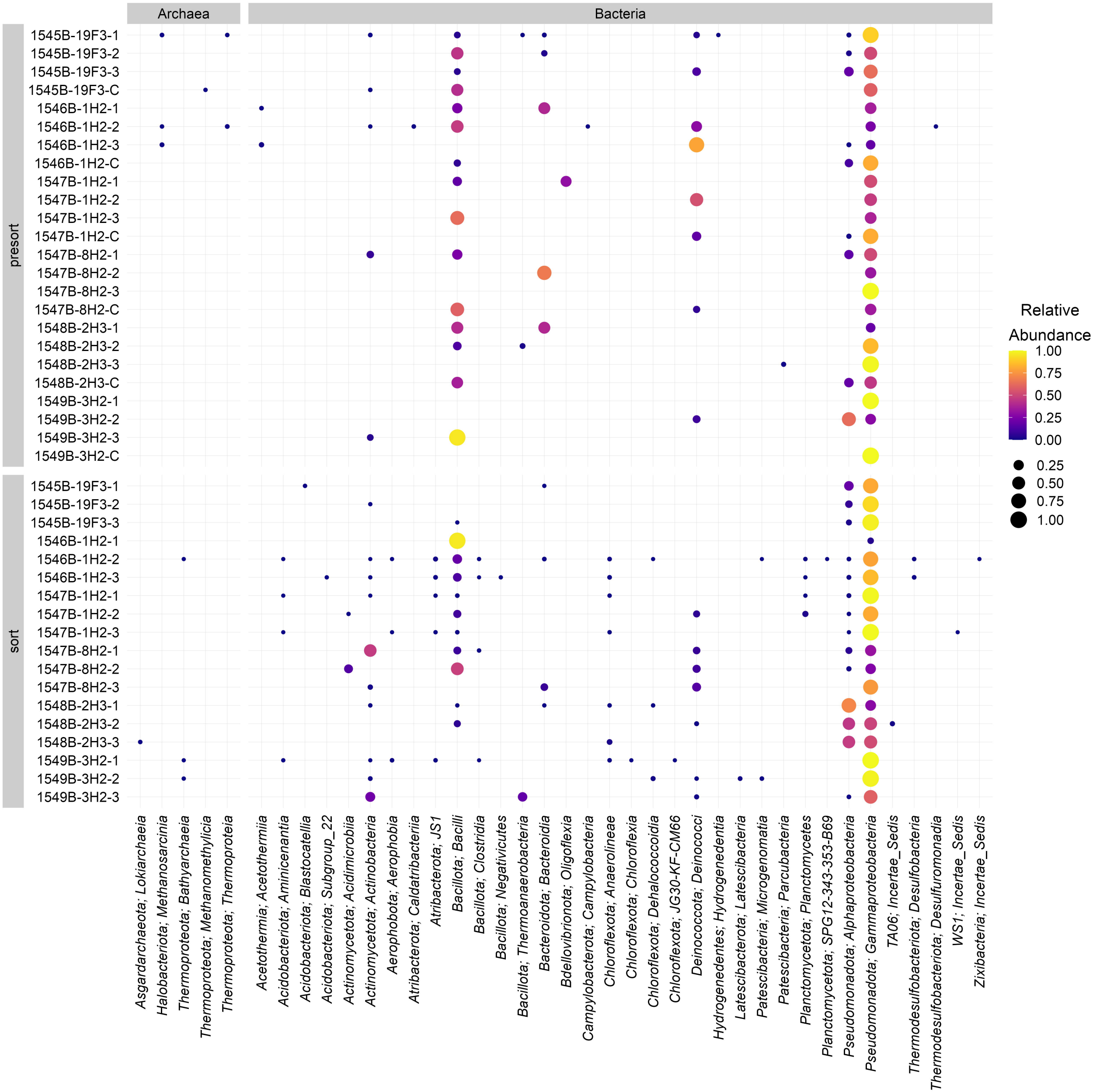
749 89. Trembath-Reichert E et al. Methyl-compound use and slow growth characterize microbial life
750 in 2-km-deep subseafloor coal and shale beds. *Proceedings of the National Academy of*
751 *Sciences* 2017;**114**:E9206–E9215. https://doi.org/Tre
752 90. Zheng R et al. Physiological and metabolic insights into the first cultured anaerobic
753 representative of deep-sea Planctomycetes bacteria. *eLife* 2024;**12**:RP89874.
754 https://doi.org/10.7554/eLife.89874
755 91. Sekiguchi Y et al. Anaerolinea thermophila gen. nov., sp. nov. and Caldilinea aerophila gen.
756 nov., sp. nov., novel filamentous thermophiles that represent a previously uncultured lineage
757 of the domain Bacteria at the subphylum level. *International Journal of Systematic and*
758 *Evolutionary Microbiology* 2003;**53**:1843–1851. https://doi.org/10.1099/ijs.0.02699-0
759 92. Lee JZ et al. Fermentation couples Chloroflexi and sulfate-reducing bacteria to Cyanobacteria
760 in hypersaline microbial mats. *Front Microbiol* 2014;**5**.
761 https://doi.org/10.3389/fmicb.2014.00061
762

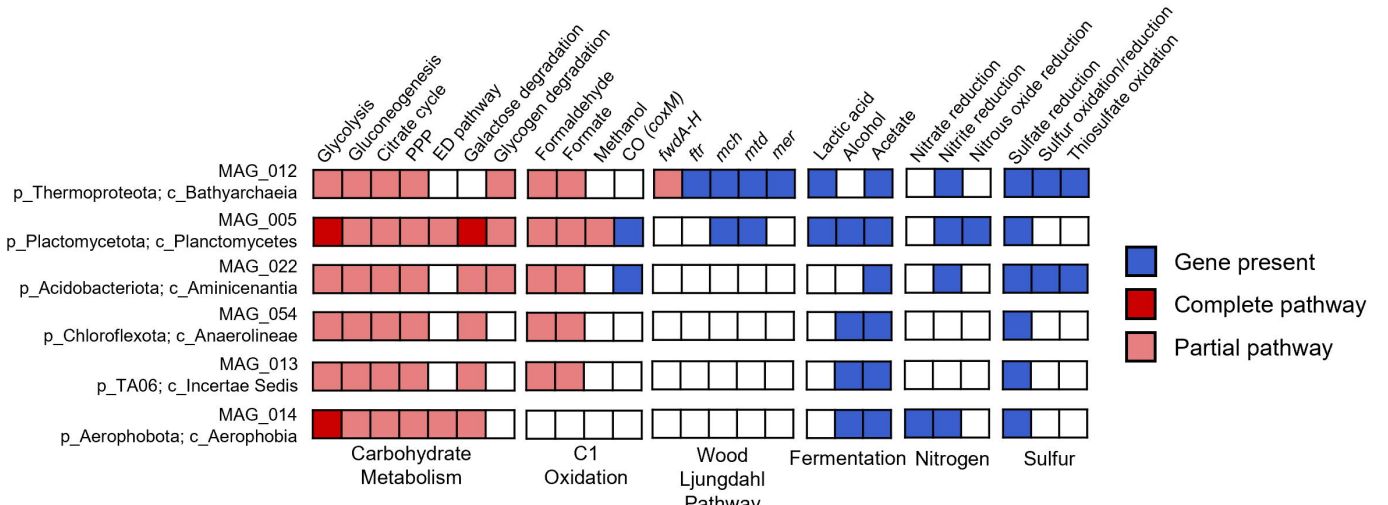
763 **Figures legends**

764 **Figure 1**. Alpha diversity indices of each sample and fold enrichment changes of the 4 most
765 abundant Classes. Plots of Observed Diversity (A) and Shannon's diversity index (B) for all sites.
766 Fold enrichment of relative abundance between the pre-sort and sorted fractions for
767 Gammaproteobacteria (C), Alphaproteobacteria (D), Bacilli (E), and Deinococci (F). Fold
768 enrichment was calculated as the ratio of relative abundances of ASVs affiliated with a given class
769 between the pre-sort and sorted fraction, normalized to zero so that equal frequencies equate to
770 zero. In fold-enrichment plots (C-F), yellow bars indicate that a class was more represented in the
771 sorted fraction, and blue bars show that the class was more abundant in the pre-sort fraction. Bars

772 represent the average of triplicate samples, and each point shows the value of each replicate. A *
773 indicated that no ASVs affiliated with that class were detected in either the presort or sorted
774 fraction. Note that samples are presented in order of increasing depth.
775
776 **Figure 2**. Community composition of cell extract (pre-sort) and BONCAT-positive sorted samples
777 depicted as the relative abundance of ASVs of the 16S rRNA gene collapsed to class level.
778 Fractions are grouped separately on the y-axis. Replicate numbers are indicated by 1-3; C refers
779 to the no-HPG control.
780
781 **Figure 3**. Heat map of metabolic potential of the six translationally active MAGs. A subset of
782 genes was assessed. Squares indicate the presence of a functional gene (blue), complete metabolic
783 pathway (red), or partial metabolic pathway (pink). White squares indicate that no gene(s)
784 affiliated with that pathway were detected. MAGs are numbered following identities in (Mara et
785 al., 2023). Abbreviations: PPP (pentose phosphate pathway), ED pathway (Entner-Doudoroff
786 pathway). MAG_012 (84.65% complete, 9.60% contamination), MAG_005 (95.91% complete,
787 7.52% contamination), MAG_022 (87.01% complete, 4.23% contamination), MAG_054 (52.66%
788 complete, 3.54% contamination), MAG_013 (86.40% complete, 0.78% contamination),
789 MAG_014 (88.80% complete, 8.25% contamination).


A
Observed Diversity
50
40
30
20
10
0
1546B-1H2
1547B-1H2
1548B-2H3
1549B-3H2
1547B-8H2
1545B-19F3
B
Shannon Diversity
3.0
2.5
2.0
1.5
1.0
0.5
0.0
1546B-1H2
1547B-1H2
1548B-2H3
1549B-3H2
1547B-8H2
1545B-19F3
Legend
= presort
= sort
C
Gammaproteobacteria
presort vs sort
1546B-1H2
1547B-1H2
1548B-2H3
1549B-3H2
1547B-8H2
1545B-19F3
-3
-2
-1
0
1
2
3
D
Alphaproteobacteria
presort vs sort
*
E
Bacilli
presort vs sort
F
Deinococci
presort vs sort
*


Archaea
Bacteria
presort
sort
1545B-19F3-1
1545B-19F3-2
1545B-19F3-3
1545B-19F3-C
1546B-1H2-1
1546B-1H2-2
1546B-1H2-3
1546B-1H2-C
1547B-1H2-1
1547B-1H2-2
1547B-1H2-3
1547B-1H2-C
1547B-8H2-1
1547B-8H2-2
1547B-8H2-3
1547B-8H2-C
1548B-2H3-1
1548B-2H3-2
1548B-2H3-3
1548B-2H3-C
1549B-3H2-1
1549B-3H2-2
1549B-3H2-3
1549B-3H2-C
1545B-19F3-1
1545B-19F3-2
1545B-19F3-3
1546B-1H2-1
1546B-1H2-2
1546B-1H2-3
1547B-1H2-1
1547B-1H2-2
1547B-1H2-3
1547B-8H2-1
1547B-8H2-2
1547B-8H2-3
1548B-2H3-1
1548B-2H3-2
1548B-2H3-3
1549B-3H2-1
1549B-3H2-2
1549B-3H2-3
Asgardarchaeota; Lokiarchaeia
Halobacteriota; Methanosarcinia
Thermoproteota; Bathyarchaeia
Thermoproteota; Methanomethylicia
Thermoproteota; Thermoproteia
Acetothermia; Acetothermiia
Acidobacteriota; Aminicenantia
Acidobacteriota; Blastocatellia
Acidobacteriota; Subgroup_22
Actinomycetota; Acidimicrobiia
Actinomycetota; Actinobacteria
Aerophobota; Aerophobia
Atribacterota; Caldatribacteriia
Atribacterota; JS1
Bacillota; Bacilli
Bacillota; Clostridia
Bacillota; Negativicutes
Bacillota; Thermoanaerobacteria
Bacteroidota; Bacteroidia
Bdellovibrionota; Oligoflexia
Campylobacterota; Campylobacteria
Chloroflexota; Anaerolineae
Chloroflexota; Chloroflexia
Chloroflexota; Dehalococcoidia
Chloroflexota; JG30-KF-CM66
Deinococcota; Deinococci
Hydrogenedentes; Hydrogenedentia
Latescibacterota; Latescibacteria
Patescibacteria; Microgenomatia
Patescibacteria; Parcubacteria
Planctomycetota; Planctomycetes
Planctomycetota; SPG12-343-353-B69
Pseudomonadota; Alphaproteobacteria
Pseudomonadota; Gammaproteobacteria
TA06; Incertae_Sedis
Thermodesulfobacteriota; Desulfobacteria
Thermodesulfobacteriota; Desulfuromonadia
WS1; Incertae_Sedis
Zixibacteria; Incertae_Sedis
Relative Abundance
1.00
0.75
0.50
0.25
0.00
0.25
0.50
0.75
1.00


Glycolysis
Gluconeogenesis
Citrate cycle
PPP
ED pathway
Galactose degradation
Glycogen degradation
Formaldehyde
Formate
Methanol
CO (coxM)
fwdA-H
ftr
mch
mtd
mer
Lactic acid
Alcohol
Acetate
Nitrate reduction
Nitrite reduction
Nitrous oxide reduction
Sulfate reduction
Sulfur oxidation/reduction
Thiosulfate oxidation
MAG_012
p_Thermoproteota; c_Bathyarchaeia
MAG_005
p_Plactomycetota; c_Planctomycetes
MAG_022
p_Acidobacteriota; c_Aminicenantia
MAG_054
p_Chloroflexota; c_Anaerolineae
MAG_013
p_TA06; c_Incertae Sedis
MAG_014
p_Aerophobota; c_Aerophobia
Carbohydrate Metabolism
C1 Oxidation
Wood Ljungdahl Pathway
Fermentation
Nitrogen
Sulfur
Gene present
Complete pathway
Partial pathway