1 Microbial community of recently discovered Auka vent field sheds light on vent 2 biogeography and evolutionary history of thermophily

3

4 Daan R. Speth^{1,2,*}, Feiqiao B. Yu^{3,4}, Stephanie A. Connon², Sujung Lim², John S. Magyar²,

5 Manet E. Peña⁵, Stephen R. Quake^{3,4}, and Victoria J. Orphan^{1,2}

6

7 ¹ Division of Biology and Biological Engineering, California Institute of Technology, Pasadena,

8 CA, USA

9 ² Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena,

10 CA, USA

11 ³ Department of Bioengineering, Stanford University, Stanford, CA, USA

12 ⁴ Chan Zuckerberg Biohub, San Francisco, CA, USA

13 ⁵ Facultad de Ciencias Marinas, Universidad Autónoma de Baja California, Ensenada, Mexico

14 * Present address: Max Planck Institute for Marine Microbiology, Bremen, Germany

15

16

17 Correspondence:

18 dspeth@mpi-bremen.de or vorphan@gps.caltech.edu

19

20

21 **Abstract**

22 Hydrothermal vents have been key to our understanding of the limits of life, and the metabolic
 23 and phylogenetic diversity of thermophilic organisms. Here we used environmental
 24 metagenomics combined with analysis of physico-chemical data and 16S rRNA amplicons to
 25 characterize the diversity, temperature optima, and biogeographic distribution of
 26 sediment-hosted microorganisms at the recently discovered Auka vents in the Gulf of California,
 27 the deepest known hydrothermal vent field in the Pacific Ocean. We recovered 325
 28 metagenome assembled genomes (MAGs) representing 54 phyla, over 1/3 of the currently
 29 known phylum diversity, showing the microbial community in Auka hydrothermal sediments is
 30 highly diverse. Large scale 16S rRNA amplicon screening of 227 sediment samples across the
 31 vent field indicates that the MAGs are largely representative of the microbial community.
 32 Metabolic reconstruction of a vent-specific, deeply branching clade within the Desulfobacterota
 33 (Tharopobacteria) suggests these organisms metabolize sulfur using novel octaheme
 34 cytochrome-c proteins related to hydroxylamine oxidoreductase. Community-wide comparison
 35 of the average nucleotide identity of the Auka MAGs with MAGs from the Guaymas Basin vent
 36 field, found 400 km to the Northwest, revealed a remarkable 20% species-level overlap between
 37 vent sites, suggestive of long-distance species transfer and sediment colonization. An adapted
 38 version of a recently developed model for predicting optimal growth temperature to the Auka
 39 and Guaymas MAGs indicates several of these uncultured microorganisms could grow at
 40 temperatures exceeding the currently known upper limit of life. Extending this analysis to
 41 reference data shows that thermophily is a trait that has evolved frequently among Bacteria and
 42 Archaea. Combined, our results show that Auka vent field offers new perspectives on our
 43 understanding of hydrothermal vent microbiology.

44

45 Introduction

46 Microbial communities at hydrothermal vents have long been of interest for their impact on
 47 localized productivity and nutrient cycling in the deep ocean, as surface expressions of the
 48 subsurface biosphere, and as potential analogs for ocean life on icy moons. These
 49 chemosynthetic communities differ strongly from the communities inhabiting the surrounding
 50 seafloor, but the variation between different hydrothermal areas is not well understood. Based
 51 on host lithology, hydrothermal areas can be classified into three groups: basalt-hosted,
 52 ultramafic, and sediment-hosted. The majority of well-studied high temperature vents are
 53 basalt-hosted, with hydrothermal fluid directly discharged from fissures into the overlying
 54 seawater (Dick 2019). In contrast, sediment-hosted hydrothermal vents such as those found in
 55 Guaymas Basin are distinctive for the interaction of the superheated fluid with overlying
 56 sediment. This interaction further alters the fluid composition through incorporating
 57 thermally-degraded organic compounds during advection to the seafloor, resulting in steep
 58 temperature gradients in the sediments and near surface oil production (Procesi et al. 2019).
 59 This additional complexity makes sediment-hosted vent fields attractive study sites for microbial
 60 ecology (Teske 2020). Indeed, Guaymas Basin has proven to be a particularly rich source for
 61 discovery of novel metabolic capabilities of thermophilic microorganisms. Examples include
 62 thermophilic anaerobic oxidation of methane coupled to sulfate reduction by consortia of
 63 *Desulfofervidus* sp. bacteria and ANME-1 archaea (Holler et al. 2011; Schouten et al. 2003),
 64 anaerobic butane degradation by the sulfate-reducing bacterium *Desulfosarcina* BuS5
 65 (Kniemeyer et al. 2007), and anaerobic butane oxidation by consortia of *Synthrophoarchaeum*
 66 sp. and *Desulfofervidus* sp. Bacteria (Laso-Pérez et al. 2016). In addition, some of the most
 67 extreme hyperthermophiles, *Methanopyrus kandleri* and *Pyrodictium abyssi*, with maximum
 68 measured growth temperatures of 122°C and 110°C, respectively, have been isolated from

69 Guaymas Basin sediments and chimneys (Pley et al. 1991; Kurr et al. 1991; Takai et al. 2008).
70 The outsized role of Guaymas Basin in discovery of microbial processes makes the recent
71 discovery of the Auka vent field, a second sediment hosted hydrothermal vent system along the
72 same fault in the Gulf of California (Goffredi et al. 2017; Paduan et al. 2018; Espinosa-Asuar et
73 al. 2020), especially exciting, as it provides a unique opportunity for comparative analyses of
74 sediment hosted hydrothermal vent systems.

75 Auka is located at >3650m water depth in the Southern Pescadero Basin, a pull-apart basin at
76 the southern tip of the Gulf of California and 400 kilometers southeast of Guaymas Basin. The
77 composition of the hydrothermal fluids at both sites is similar. The fluids are slightly acidic (pH
78 6), with high concentrations of methane ($81 \pm 16 \text{ mmol kg}^{-1}$), hydrogen sulfide ($10.8 \pm 6 \text{ mmol}$
79 kg^{-1}), and carbon dioxide ($49.2 \pm 43 \text{ mmol kg}^{-1}$) at Auka and Guaymas respectively, and
80 comparatively low hydrogen gas concentration (2 mmol kg^{-1} at Auka) (Von Damm et al. 1985;
81 Welhan 1988; Paduan et al. 2018). The temperature of the fluids measured at chimney orifices
82 is close to 300°C at both locations. Due to these high temperatures, fluids advecting through the
83 sediments at both sites contain thermogenic hydrocarbons, originating from the catagenesis of
84 sediment organic matter.

85 While the similarities between both sites are striking, there are stark differences as well. At
86 3650m Auka is the deepest known hydrothermal vent system in the Pacific Ocean, and more
87 than twice as deep as Guaymas Basin. The thicker sediment cover (700 - 1000m, (Lizarralde et
88 al. 2007; Procesi et al. 2019), results in higher load of thermogenic hydrocarbons than observed
89 at Pescadero Basin, where the sediment thickness is estimated to be less than 50 m in the
90 areas directly adjacent to the hydrothermal mounds (Paduan et al. 2018). This combination of
91 overlapping and contrasting conditions between the two sites makes them prime targets for

92 comparative analysis of their microbial communities to elucidate the factors shaping microbial
93 populations at either site.

94 To characterize the sediment microbial community at the Auka vent field, we analyzed 325
95 MAGs recovered from metagenomic sequencing combined with a 16S rRNA gene-based
96 diversity survey. The diversity and genomic similarity was then compared between Auka MAGs
97 and those recently reported from Guaymas Basin sediments (Dombrowski et al. 2017;
98 Dombrowski, Teske, and Baker 2018; Seitz et al. 2019), providing important insights into shared
99 microbial species and patterns in vent biogeography. Focusing on one of the novel lineages
100 shared between Auka and Guaymas, a deep-branching lineage of Desulfobacterota, named
101 Tharpobacteria, we conducted a detailed analysis of the metabolic potential within this novel
102 vent-associated clade. Finally, by adapting a genome-based optimal growth temperature
103 prediction model (Sauer and Wang 2019) we determined the distribution of
104 mesophily-hyperthermophily across these diverse lineages, exploring the role temperature plays
105 in shaping microbial communities in Gulf of California hydrothermal vent sediments and, more
106 broadly, evolutionary patterns of thermophily across the tree of life.

107

108 Results

109 The Auka vent field is located at the western edge of the Southern Pescadero Basin and
 110 occupies an area of approximately 200 by 600 meters (Figure 1). There are five prominent sites
 111 of vigorous fluid venting referred to as P-vent, C-vent, Z-vent, Diane's vent, and Matterhorn,
 112 each characterized by prominent calcite chimneys (Figure 1; (Paduan et al. 2018). Between
 113 these vents lie extensive carbonate platforms dotted with centimeter scale sites of hydrothermal
 114 fluid discharge supporting localized clumps of chemosynthetic *Oasisia* tubeworms (Goffredi et
 115 al. 2017). These hydrothermal features are sediment-hosted, with sediment thickness in the
 116 area directly surrounding the chimneys and carbonate platforms estimated to be less than 50 m
 117 (Paduan et al. 2018). Throughout the vent field, but primarily near Diane's vent and south of
 118 Z-vent, dispersed microbial mats covering the sediment surface indicate widespread advective
 119 hydrothermal discharge through the sediment (Figure 1). The microbial mats exhibited a range
 120 of colors: pink, gray/white, yellow, and contain localized bright white spots surrounding focused
 121 flow of shimmering hydrothermal fluid (Supplemental figure 1). Sediment temperatures at 30cm
 122 depth were elevated relative to the bottom seawater 2.4°C, ranging from 10°C to 177°C, with the
 123 highest temperatures measured below yellow-colored mats and beneath the bright white zones
 124 associated with visible hydrothermal fluid discharge. The release of oil droplets was observed
 125 during sediment coring to the south of Z-vent within a microbial mat, consistent with
 126 near-surface petroleum production within the sediment.

127

128 To characterize the microbial community of Auka vent field, we performed shotgun
 129 metagenomic sequencing on sediments from two paired cores, DR750-PC67 and DR750-PC80,
 130 collected approximately 50 m from the main Z vent chimney (Figure 1). The cores were
 131 collected within sediment covered by patchy yellow microbial mat, 5 cm from visible discharge

132 of hydrothermal fluid and associated bright white spot on the sediment. DR750-PC80 was
133 inserted closest to the fluid discharge, with DR750-PC67 directly adjacent (<2 cm apart), in a
134 straight line away from the white spot (Supplemental Figure 1). Each core was sectioned into
135 two 7 cm horizons (0-7cm and 7-14cm), resulting in four samples. To assess whether there
136 were differences in the microbial community attached to, or forming, larger aggregates in the
137 sediments, <10 µm filtrate of a subsample of each horizon was processed for DNA extraction;
138 doubling the total number of metagenomic samples to eight.

139 After sequencing, assembly, and binning, we retrieved 331 metagenome assembled genomes
140 (MAGs) with estimated completeness over 50%. Recovered MAGs were highly diverse,
141 including 212 Bacteria and 119 Archaea and representing 54 different phyla based on the
142 genome taxonomy database (GTDB) assignment (Chaumeil et al. 2019) (Figure 2,
143 Supplemental Data S5). 105 of these MAGs were estimated to be >90% complete with less than
144 5% contamination, and 111 MAGs contained (fragments of) a 16S rRNA gene (Supplemental
145 Data S5) allowing direct comparison with the more extensive 16S rRNA amplicon survey data
146 from Auka.

147

148 15 MAGs were more than twofold depleted in the <10 µm filtered fraction (Supplemental Data
149 S5). The most depleted MAG in this fraction was a dominant ANME-2c archaeon, which is
150 consistent with members of this methanotrophic clade frequently forming multi-celled consortia
151 >10 µm in association with syntrophic sulfate-reducing bacteria (SRB). Notably, the putative
152 sulfate-reducing partner of ANME-2c (e.g. *Desulfobacterota* SEEP-SRB2), was also depleted
153 (1.88 fold) in the filtered fraction. Another clade depleted in the filtered fraction were the
154 heterotrophic bacteria *Izimaplasma* originally described from methane cold seeps (Skenneron
155 et al. 2016; Zheng et al. 2021). While anaerobic enrichments of this clade consisted of

156 free-living organisms (Skenner et al. 2016), their depletion in the filtered fraction could be
 157 due to association with larger organic particles for heterotrophic growth on DNA (Zheng et al.
 158 2021). Notably, all four recovered MAGs belonging to the *Odinarchaeota* clade of Asgard
 159 Archaea were also strongly depleted in the filtered fraction. The morphology and eco-physiology
 160 of *Odinarchaeota* is poorly characterized and these findings may suggest an association with
 161 sediment particles or other microorganisms, or perhaps linked to an increased effective cell size
 162 due to unusual morphology, as observed for *Prometheoarchaeum* MK-D1, the only cultured
 163 member of the Asgard Archaea (Imachi et al. 2020). None of the four ANME-1 or three
 164 *Desulfosphaeriales* MAGs were strongly depleted in the filtered sample, consistent with previous
 165 observations that sediment-hosted ANME-1 often occur as single cells and tend to form less
 166 well-structured aggregates with SRB relative to ANME-2 consortia (Orphan et al. 2002; Holler et
 167 al. 2011).

168 Conversely, 14 MAGs were >2 fold enriched in the filtered fraction, suggesting limited
 169 association with the sediment matrix relative to other community members. These 14 MAGs
 170 represented 12 bacteria (seven Proteobacteria, two Synergistota, one Desulfobacterota, one
 171 Poribacteria, and one Thermotogota), and two Thermoplasmatota Archaea (Supplemental Data
 172 S5). Six of these MAGs (three Alphaproteobacteria and three Gammaproteobacteria) were
 173 absent from the unfiltered data and were therefore considered contamination and removed from
 174 further analyses, leaving 325 MAGs representing 54 phyla (Figure 2).

175

176 The 325 MAGs recovered from 2 sediment cores describe the microbial community at Auka
 177 vent field at a single location. To better understand the sediment microbial community structure
 178 and distribution patterns in the total vent area in relation to the physicochemical parameters, we
 179 conducted an area-wide survey using 16S rRNA gene amplicon sequencing of 29 additional

180 sediment push cores, representing 227 total samples, collected by remotely operated vehicle
 181 during expeditions in 2017 and 2018 (Figure 1). Porewater geochemical profiles for 22 of these
 182 cores showed a variable degree of mixing of hydrothermal fluids with seawater, as evidenced by
 183 a steep drop in magnesium concentration with depth (Von Damm et al. 1985), and concurrent
 184 increases in calcium and potassium concentration (Supplemental figures S2 and S3,
 185 Supplemental Data S1). Consistent with a variable degree of mixing, temperature measured at
 186 30 cm depth in the sediments varied greatly, ranging from 10°C to 177°C, a range similar to
 187 reports from Guaymas basin sediments (Biddle et al. 2012; McKay et al. 2016; Dowell et al.
 188 2016; Teske et al. 2016). Concentrations of oxidized nitrogen species (nitrate and nitrite) were
 189 below the detection limit in porefluids, while ammonium reached concentrations as high as 16
 190 mM, likely as a result of thermal degradation of organic matter (Haberstroh and Karl 1989).
 191 Porewater sulfide profiles frequently showed maxima up to 12 mM between 5 and 15 cm below
 192 the sediment surface. Sulfate concentrations dropped below the detection limit in the top 20 cm
 193 below the seabed in 14 cores, attributed to a combination of microbial sulfate reduction and
 194 seawater mixing with sulfate-depleted hydrothermal fluid. In cores with the highest inferred flux
 195 of hydrothermal fluid (e.g NA091-119 and S200-PC1), porewater sulfate concentrations were
 196 consistently below 10 mM at all depths (Supplemental figures S2 and S3).

197

198 The taxa detected in our 16S rRNA gene amplicon survey were broadly similar in the 29 cores
 199 (Supplemental Figures S4 - S32). A comparison of the amplicon sequences with 16S rRNA
 200 genes retrieved from the metagenome using PhyloFlash (Gruber-Vodicka, Seah, and Pruesse
 201 2020), and assembled 16S rRNA genes from select MAGs, showed congruence between the
 202 most abundant taxa in all 29 cores and dominant taxa recovered from metagenome sequencing
 203 (Figure 3). This shows that the MAGs assembled from a single location (2 adjacent cores)

provides representative insight of the microbial community within the greater vent field. The overall distribution pattern in Auka sediments indicates widespread distribution of sulfide/sulfur-oxidizing Campylobacterota (formerly Epsilonproteobacteria) and putatively heterotrophic Bacteroidota as part of the surface microbial mat assemblage, with limited recovery of sulfur-oxidizing *Beggiatoa* (Supplemental text). Within the underlying sediment, ANME-2c archaea and Seep-SRB2 sulfate-reducing bacteria (SRB) of the *Dissulfuribacteraceae*, linked to the sulfate-dependent anaerobic oxidation of methane were dominant near the sediment surface, with consortia of ANME-1 Archaea and *Desulfofervidus* sp. SRB were highly abundant in deeper sediment horizons. This ANME niche separation along temperature and geochemical gradients has been reported previously from Guaymas Basin, with ANME-2c observed in lower temperature cores, and ANME-1 in higher temperature cores (Teske et al. 2002; Biddle et al. 2012; Holler et al. 2011). Besides temperature, sulfate concentration may also contribute to niche separation at Auka, with ANME-1 phylotypes dominant in horizons corresponding to low sulfate concentrations, a pattern that has also been described from cold seeps. In several cores (e.g. NA091-118 & 119) amplicon sequence variant (ASV) distribution showed further depth stratification of ANME-1 phylotypes, with ANME-1a ASVs co-occurring with Seep-SRB2 phylotypes in shallower horizons, again consistent with observations from methane cold seeps (Metcalf et al. 2021) (Supplemental figures S10 & S11). Other clades present in high abundance are *Desulfobacteraceae* in the shallow sediments, with Caldatribacterota (JS1), Thermoplasmatota (DHVEG-2), and Thermotogota ASV's common in the deeper horizons (Figure 3). Consistent with the variation in porewater ion concentrations from hydrothermal fluid advection, the depth distribution of taxa varied between sediment cores, with higher inferred hydrothermal fluid input corresponding to limited detection of ANME-2c, higher abundance of ANME-1 in the shallow horizons, and appearance of hyperthermophilic

228 lineages (e.g. Methanopyraceae, Thermoprotei) in the deeper horizons. A full summary of
229 sediment 16S rRNA diversity from the 29 core survey is provided in Supplemental Figures
230 S4-S32 and Supplemental Data S2-S4.

231

232 Both the 16S rRNA diversity analyses and the metagenomic sequencing indicate community
233 stratification by depth in the sediment. We expect temperature to play a major role in shaping
234 community structure, because of the high degree of hydrothermal fluid mixing and observed
235 steep thermal gradients (up to $7^{\circ}\text{C cm}^{-1}$). While broad trends in relative abundance of
236 sediment-hosted microbial community members along a temperature gradient have been shown
237 for Guaymas Basin and other hydrothermal sites (Lutz et al. 2008; Anderson et al. 2013; Ding et
238 al. 2017), the relationship of these trends to optimal growth temperature (OGT) is lacking for the
239 majority of taxa due to the lack of cultured representatives. Genomic data is shedding new light
240 on physiological characteristics like OGT, with strong correlations observed between select
241 genomic features and OGT for cultured bacteria and archaea ((Sauer and Wang 2019) and
242 references therein). This knowledge was recently synthesized and incorporated into an OGT
243 prediction model and used to accurately predict OGTs of phylogenetically diverse cultured
244 microorganisms (Sauer and Wang 2019). Here we apply a modified version of this model (see
245 methods) and used this to estimate OGT for the environmental MAGs recovered from the Auka
246 vents, Guaymas Basin, and a selected set of over 4000 archaeal and bacterial genomes from
247 the GTDB (Figure 2, Supplemental Data S5). For several microorganisms with known OGT or
248 enrichment temperature, these genome-based predictions were largely consistent with reported
249 values. For example, the predicted OGTs of the three Auka *Desulfofervidales* MAGs were 52°C ,
250 53°C , and 60°C , close to the experimentally determined OGT of 60°C for *Desulfofervidus auxilli*
251 in pure culture (Krukenberg et al. 2016) and the predicted OGT of the two ANME-1 MAGs in the

252 GB60 clade were 63°C and 65°C, also close to the 60°C enrichment temperature of the GB60
 253 enrichment culture (Holler et al. 2011) (Figure 2, Supplemental Data S5). Predicted OGT of the
 254 MAGs correlated with their abundance (approximated by read coverage) as a function of
 255 sediment depth, with most MAG's with OGTs over 50°C showing higher abundance in the >7 cm
 256 horizon in both cores. Notably, this OGT depth trend was less pronounced in core DR750-PC80,
 257 collected directly adjacent to localized hydrothermal fluid discharge, compared to core
 258 DR750-PC67 which was collected on the far side of DR750-PC80, with the closest edge
 259 approximately 10 cm further away from the fluid source (Figure 4, Supplemental Figure 1).
 260 While sediment temperatures were not measured for these cores, the visible fluid discharge at
 261 the seabed indicates that core DR750-PC80 likely was exposed to a greater flux of
 262 hydrothermal fluids and steeper temperature gradients relative to PC67, potentially explaining
 263 the higher abundance of MAGs with predicted OGT values over 50°C in the 0-7 cm horizon.

264

265 In addition to the correlation between predicted OGT and increasing abundance in lower
 266 horizons, there also is a clear correlation between predicted OGT and phylogeny (Figure 2).
 267 Using our MAGs and reference data from the genome taxonomy database (GTDB) (Parks et al.
 268 2020), we assessed the phylogenetic signal of predicted OGT in more detail. After taxonomic
 269 assignment (Chaumeil et al. 2019), Auka MAGs, recent Guaymas MAGs (Dombrowski et al.
 270 2017; Dombrowski, Teske, and Baker 2018; Seitz et al. 2019), and reference genomes from the
 271 GTDB (v89) were used to construct clade specific phylogenies with corresponding OGT
 272 predictions, representing a total of 5111 genomes, including all archaea and the majority of
 273 bacterial phyla included in GTDB v89. Within these 14 phylogenies, organisms with predicted
 274 (hyper)thermophilic OGT values grouped together in distinct lineages, often emerging from
 275 lineages with predicted mesophilic OGT values, suggesting thermophilic adaptation is common

276 throughout both the bacterial and archaeal domains of the tree of life, and frequently persists in
 277 a lineage once acquired (Supplemental Figures S33 - S46). In addition, there are several
 278 examples of mesophilic lineages evolving from hyperthermophilic ancestors, such as the
 279 Nitrososphaeria (formerly Thaumarchaea), showing that thermophily is a reversible trait
 280 (Supplemental figure S33). The OGTs predicted for MAGs retrieved from Auka and for
 281 previously published MAGs from Guaymas Basin indicate the occurrence of hyperthermophilic
 282 microorganisms across many phyla at these sites, with one MAG from an uncultured member of
 283 the Thermoprotei from Guaymas Basin having the highest predicted OGT of all organisms
 284 analyzed, at 117°C. Among the Auka MAGs, a member from the same Thermoprotei clade had
 285 a predicted OGT of 111°C (Figure 3, Supplemental Figure S33, Supplemental Data S5).
 286 Surprisingly, the four Asgard Archaea MAGs recovered from Auka are the only predicted
 287 (hyper)thermophiles in this clade, despite several other Asgard MAGs being retrieved from
 288 hydrothermal sites (Supplemental Figure S33). Consistent with surveys of cultured organisms
 289 (Sauer and Wang 2019), the maximum predicted OGT in the 5111 genomes is considerably
 290 higher for Archaea (117°C) than for Bacteria (89°C);(Figure 2), although we note that not all
 291 Bacteria in the GTDB were included in this analysis.

292

293 In addition to clade-specific adaptation to high temperature, the phylogenetic analysis showed
 294 that Guaymas Basin and Auka vent fields have substantial community overlap. To further
 295 quantify this overlap in microbial communities between sites, we calculated average nucleotide
 296 identity (ANI) between all MAGs for both the Auka and Guaymas datasets. This analysis
 297 revealed that 68 Auka MAGs, representing 23 bacterial and archaeal phyla, have ANI values
 298 that are >95% to MAGs from Guaymas Basin sediments (Dombrowski et al. 2017; Dombrowski,
 299 Teske, and Baker 2018; Seitz et al. 2019), corresponding to nearly 20% species overlap

300 between these geographically distant vent fields in the Gulf of California (Jain et al. 2018; Olm
301 et al. 2020) (Figure 2 & Figure 5A, Supplemental Data S5). Another 93 Auka MAGs had ANI
302 values between 75% and 95% with MAGs from Guaymas, further showing broad community
303 similarity. Previous work has shown that the community at Guaymas Basin is distinct from those
304 at basalt-hosted and ultramafic systems (Reveillaud et al. 2016). Indeed, a comparison with the
305 MAGs retrieved from sampling deep hydrothermal fluids at Juan de Fuca ridge in the north
306 Pacific (Jungbluth, Amend, and Rappé 2017), and hydrothermal fluids from mafic and ultramafic
307 vent sites Cayman rise in the Caribbean (Anderson et al. 2017) showed only 9 and 12 MAGs
308 with ANI above the 75% threshold value to the MAGs from Auka respectively, and no MAGs
309 with >85% ANI in either dataset (Figure 5A).

310 We hypothesize that the observed 20% species overlap is due to continuous transfer of
311 microbial populations between Auka and Guaymas Basin. The exchange over the ~400 km
312 between both sites may be facilitated by hydrothermal plumes, which have been shown to rise
313 at least 900 m upwards from the vents at Guaymas Basin (Merewether, Olsson, and Lonsdale
314 1985) and harbor distinct microbial communities from the surrounding seawater (Dick and Tebo
315 2010; Dick 2019; Anantharaman, Breier, and Dick 2016). In addition, long distance transfer of
316 thermophilic microorganisms through ocean currents has been documented (Hubert et al. 2009;
317 Müller et al. 2014; Resing et al. 2015), and the open ocean has been shown to contain a “seed
318 bank” of hydrothermal vent taxa (Gonnella et al. 2016). Several organisms in this “seed bank”
319 were also shown to be culturable from seawater particulates (Stetter et al. 1993), indicating that
320 plume mediated microbial population transfer throughout the Gulf of California is highly likely.

321 Considering that hydrothermal plumes likely facilitate a constant flux of organisms between both
322 basins, the species overlap of 20% indicates selection at dispersal, transfer, colonization, or a
323 combination of these. As there is a strong correlation between temperature and community

composition at Auka, we investigated whether the predicted OGT of MAGs correlated with ANI values between Auka and Guaymas Basin, but found no correlation, with both mesophilic and thermophilic groups represented among the 20% (Figure 5B). In addition, there was no correlation with ANI values and MAG abundance in surface sediments (Figure 5C), or MAG abundance in deeper horizons, with both abundant and rare taxa represented (Figure 5D). The latter observation contrasts previous work showing that abundant taxa were more likely to be cosmopolitan (Anderson, Sogin, and Baross 2015). The sediment MAGs with species-level overlap between the Gulf of California vent sites are phylogenetically and physiologically diverse, suggesting there is not a single determinant of transfer and colonization success. Based on these results, we hypothesize that the species overlap between Auka and Guaymas is primarily driven by the niches created from similar environmental conditions in the hydrothermal vent sediments and deeply sourced fluids (Paduan et al. 2018), selecting for colonization by a subset of the transferred microbial population. In addition, the distribution of ANI values may indicate ongoing speciation between populations within the communities at the two sites. Previous studies have shown a gap in pairwise ANI values in the range 85% - 95% which was used to support 95% ANI as the species cutoff (Jain et al. 2018; Olm et al. 2020). While there is a clear peak of ANI values above 95% between Auka and Guaymas MAGs, an additional 25 MAGs (8% of the community) share ANI values between 90% - 95% showing a less pronounced gap in ANI values than previously observed (Figure 5A). These MAGs may represent lineages undergoing speciation after immigration and colonization of the new vent site.

Thus far, the genomic determinants for colonization success have not been established, potentially because such factors are likely to be lineage specific. For example, it is striking that strains of the same archaeal ANME-1 species are the most abundant organism in sediments

348 from both the Auka vent field and the Guaymas basin, but that the other six ANME-1 MAGs
349 (three at each site) belong to distinct lineages, suggesting niche differentiation (Supplemental
350 Figure S36). 16S rRNA gene amplicon sequencing, which was done at higher sediment depth
351 resolution, showed differential distribution of ANME-1 lineages in multiple sediment cores at
352 Auka, supporting niche differentiation between them (Supplemental Figure S4 - S32).

353

354 Beyond implications for biogeography, the lineages largely or fully consisting of organisms found
355 at Auka and Guaymas are of interest for future comparative genomics work. For example,
356 several Aerophobota with OGT's between 50°C and 68°C were detected in both Auka and
357 Guaymas Basin (Supplemental fig. S38). MAGs from this phylum were recently also retrieved
358 from other hydrocarbon-rich environments, including methane cold seeps in the South China
359 Sea (Huang et al. 2019) and areas of petroleum seepage in the Gulf of Mexico (Dong et al.
360 2019), making this clade a promising target for genome guided metabolic predictions and
361 enrichment cultivation focused on anaerobic hydrocarbon degradation. Another example is a
362 deep branching clade within the Desulfobacterota phylum (Supplemental figure S43), likely
363 representing a novel class, that was originally discovered in Guaymas Basin (Dombrowski,
364 Teske, and Baker 2018) and is well represented in the surface layers of Auka vent field
365 sediments. The Guaymas basin MAGs representing this group (indicated as DQWO01) were
366 also included in a recent large scale analysis of Desulfobacterota metabolic potential (Langwig
367 et al. 2021). We recovered five MAGs from this clade, ranging from 55 - 92% estimated
368 completeness, with estimated contamination ranging from 0 - 4.3%. Combined with three MAGs
369 from Guaymas Basin (completeness 55 - 70%, contamination 1.6 - 7.1%) these eight MAGs
370 formed a deep branching monophyletic clade within the Desulfobacterota which we selected for
371 further analysis. The size of the Desulfobacterota MAGs ranged from 1.3 - 3.1 Mbp and, using

372 read mapping as a proxy for abundance, these organisms are enriched in the surface horizons
 373 of cores taken at both Auka and Guaymas. Their predicted OGT values (35 - 41°C) is consistent
 374 with a niche in a mesophilic environment. In accordance with the recent proposal to use
 375 genomic information as type material for naming microbial taxa (Chuvochina et al. 2019), we
 376 propose “*Candidatus Tharpobacterium aukensis*” for the organism represented by the
 377 PB_MBMC_085 MAG (82% estimated completeness, 0% estimated contamination), and will
 378 refer to the clade containing these eight genomes as Tharpobacteria (Supplemental Data S5).
 379 The genus name *Tharpobacterium* was chosen in honor of Marie Tharp, for her work on ocean
 380 floor mapping and plate tectonics that is key to our understanding of hydrothermal vents, where
 381 these organisms are found. The species name *aukensis* represents Auka, the location the MAG
 382 for which the name is proposed was recovered.

383 To investigate the metabolic potential of the Tharpobacteria, we performed functional
 384 enrichment analysis (Shaiber et al. 2020) on the 427 genomes within the Desulfobacterota
 385 obtained from Auka, Guaymas Basin and GTDB v89 (Supplemental Figure S43). This analysis
 386 indicated the Tharpobacteria MAGs encode the potential for beta oxidation of long chain fatty
 387 acids and degradation of benzoyl-CoA, indicating a possible role in aromatic hydrocarbon
 388 degradation (Boll and Fuchs 1995). Furthermore, Tharpobacteria are likely capable of
 389 degradation of butyrate, as recently reported for *Ca. Phosphitivorax sp.* in the UBA1062 order
 390 (Hao et al. 2020), another deep branching group within the Desulfobacterota. Unlike the
 391 UBA1062 genomes, Tharpobacteria MAGs encode NADH dehydrogenase (complex I),
 392 bc₁-complex (complex III), and a heme biosynthesis pathway. We propose that Tharpobacteria
 393 MAGs use the electrons obtained from the oxidation of fatty acids and other hydrocarbons, and
 394 subsequently directed into the quinone pool via complex I to reduce a periplasmic electron
 395 acceptor (Figure 6A). However, the nature of this electron acceptor was not directly obvious

396 from the genomic potential of the Tharpobacteria MAGs. Unlike many Desulfobacterota, the
 397 Tharpobacteria do not encode the capability to use sulfate as electron acceptor, and neither the
 398 functional enrichment analysis nor a follow up manual investigation of the genomes revealed
 399 terminal reductases for utilization of common electron acceptors.

400 We screened for other potential terminal reductases among the Tharpobacteria genomes by
 401 analyzing genes containing heme binding motifs (CxxCH), as hemes can be involved in both
 402 electron transfer and catalytic centers. A protein affiliated with the multiheme cytochrome-c
 403 (MCC) fold family, harboring well-characterized proteins involved in nitrogen and sulfur cycling
 404 (Simon et al. 2011), was present in 7 of the 8 MAGs. An exploratory analysis of all 5855 MCC
 405 family proteins retrieved from the GTDB genomes (v95) indicated this protein of interest was a
 406 member of the hydroxylamine oxidoreductase (HAO) family (Figure 6B). Proteins of the HAO
 407 family are known to play key roles in the nitrogen cycle, first identified as an essential protein in
 408 aerobic ammonia oxidizing bacteria (AOB) (Igarashi et al. 1997), and later shown to be essential
 409 for anaerobic ammonium oxidation (anammox) as well (Kartal and Keltjens 2016). HAO family
 410 proteins from Campylobacterota have been shown to catalyze nitrite reduction to ammonium *in*
 411 *vitro*, although their physiological role remains unclear (Haase et al. 2017). Notably, none of the
 412 HAO family proteins have suggested roles outside of the nitrogen cycle. However, our analysis
 413 found members of the HAO protein family in the genomes of many organisms not associated
 414 with nitrogen cycling, spanning 65 phyla (GTDB v95), and most commonly in the
 415 Desulfobacterota phylum. Phylogenetic analysis of the 1502 HAO family proteins showed that
 416 those with a known function in aerobic and anaerobic ammonium oxidation form a monophyletic
 417 clade. The Campylobacterota HAOs form a separate clade, while proteins of unknown function
 418 form several other clades within the family, with the Tharpobacteria proteins clustering within a
 419 clade with proteins from other Desulfobacterota (Figure 6C). Based on this distribution, and the

420 absence of detectable oxidized nitrogen species from the Auka sediments, we hypothesize this
 421 protein catalyzes the reduction of sulfur species, rather than nitrogen. There is precedence for
 422 this as sulfur cycling has been shown for other members of the MCC fold family such as sulfite
 423 reductase *mccA* (Hermann et al. 2015), and octaheme tetrathionate reductase (Mowat et al.
 424 2004). We propose that this HAO family protein catalyzes the reduction of the terminal electron
 425 acceptor in the Tharpobacteria clade, based on its prevalence across the MAGs.

426 It is also important to consider the possibility that individual Tharpobacteria MAGs use different
 427 electron acceptors, as there are several other protein complexes representing potential
 428 candidates as the site of the final reduction in the electron flow. Most prominently, several
 429 Tharpobacteria MAGs encode molybdopterin oxidoreductases, key complexes of many
 430 anaerobic metabolisms (Grimaldi et al. 2013). However, none of these molybdopterin
 431 oxidoreductase complexes have a known function, or are conserved in more than three of the
 432 eight genomes. In addition, there are several other cytochrome-containing proteins that could be
 433 candidate sites for terminal electron acceptor reduction in specific MAGs. Further research is
 434 needed to resolve the metabolism of this deep branching Desulfobacterota clade, but this
 435 metagenomic analysis offers an entry point for further characterization of the closely related
 436 HAO family proteins from isolated organisms as well as designing targeted enrichments,
 437 transcriptomic analyses, or stable isotope probing experiments using vent samples harboring
 438 Tharpobacteria.

439

440

441

442 Discussion

443 Genome-resolved metagenomics is rapidly yielding genomic information of organisms across
444 the tree of life, which is valuable for hypotheses about the ecology and physiology of organisms
445 in uncultured lineages, and to provide targets for subsequent experimental study. Our analyses
446 of the Auka sediments revealed a highly diverse microbial community, comprising many
447 understudied lineages affiliated with sediment hosted hydrothermal vents. An in-depth look at
448 one of these lineages, which we designated Tharpobacteria, revealed respiratory fatty acid
449 degradation coupled to an unidentified electron acceptor, for which we propose reduction of a
450 sulfur compound catalyzed by a MCC fold family protein. This provides a clear target for
451 experimental verification, and confirmation of our hypothesis would have implications for
452 interpretation of the biological sulfur cycle well beyond the Tharpobacteria.

453 In addition to gene function, investigating the genomic factors involved in colonization success
454 at Auka and Guaymas is an exciting direction for future work, and a distinct advantage of using
455 MAGs over the 16S rRNA gene amplicon analyses more frequently used for biogeography
456 studies (Nemergut et al. 2011; Anderson, Sogin, and Baross 2015; Ruff et al. 2015). The
457 geography and geological setting of the Gulf of California makes this region exceptionally well
458 suited as a model system for hydrothermal vent biogeography. The narrow gulf constrains the
459 currents and thus the direction of hydrothermal plumes and the increasing sediment thickness
460 with distance from the mouth of the gulf differentiates conditions between the two known vent
461 sites. The spreading centers on the Carmen and Farallon segments, located between the
462 Guaymas and Pescadero segments (Lizarralde et al. 2007), could also harbor as yet
463 undiscovered hydrothermal vents that act as stepping stones between the two sites.

464 Furthermore, the sediment-hosted hydrothermal vents of the Gulf of California provide an
465 excellent study site for further discoveries of (hyper)thermophilic organisms. The maximum

466 predicted OGT for organisms at Auka (111°C) and Guaymas (117°C) were substantially higher
 467 than the highest experimentally determined OGT values of 106°C for *Pyrolobus fumarii* (Blöchl
 468 *et al.* 1997), and 105°C for *Methanopyrus kandleri* grown at high pressure (Takai *et al.* 2008).
 469 Interestingly, the predicted OGT for *M. kandleri* is 98°C, identical to its observed optimum at
 470 ambient pressure (Kurr *et al.* 1991), while the OGT prediction for *Pyrolobus fumarii* is 102°C,
 471 slightly lower than the reported OGT. Strain 121, with a predicted optimum of 103°C (Kashefi
 472 and Lovley 2003), does not have a publicly available genome sequence; but the predicted OGT
 473 of 98°C for *Pyrodictium abyssi*, the most closely related organism with a sequenced genome, is
 474 consistent with experimental observation (Pley *et al.* 1991). The MAGs with predicted OGTs
 475 higher than the highest experimentally observed OGTs are not restricted to a single lineage, but
 476 rather distributed over several clades in the Thermoproteota (formerly Crenarchaeota). The
 477 Thermoproteota are severely undersampled, with the MAGs from Auka and Guaymas combined
 478 representing 40% of the Thermoproteota genomes analysed in this study. This indicates further
 479 sampling will likely yield organisms with even higher predicted OGTs than those predicted here,
 480 and suggests the upper temperature limit of life could be considerably higher than currently
 481 known. Both *in situ* and *in vitro* experiments on the sediments of the Gulf of California
 482 hydrothermal vent sites could push our knowledge of the upper temperature boundary of life.
 483 Finally, OGT prediction and genome-resolved metagenomics provide a powerful combination to
 484 examine the evolutionary history of thermophily. Our analysis shows that thermophily has
 485 evolved frequently in both the Bacteria and Archaea (Supplemental figures S33-S46),
 486 suggesting adaptation of mesophiles to the colonization of high temperature environments. The
 487 rapidly increasing genomic representation of lineages in the tree of life makes phylogenomics
 488 combined with optimal growth temperature predictions an exciting angle on the debate about
 489 the conditions in which life on earth arose and diversified.

490 **Methods**

491

492 Sample collection & shipboard processing

493 Samples were collected from Auka vent field in Pescadero Basin (Gulf of California, Mexico,
 494 23.954, -108.863) on R/V Western Flyer in 2015 (MBARI2015), E/V Nautilus in 2017 (NA091)
 495 and R/V Falkor in 2018 (FK181031); information for the required sampling permits are provided
 496 in the funding acknowledgement section. Pushcore samples were collected using remotely
 497 operated vehicle (ROV) Doc Ricketts (MBARI2015), ROV Hercules (NA091), and ROV
 498 SuBastian (FK181031) from sediment covered areas with microbial mat cover and/or visible
 499 hydrothermal fluid flow (Figure 1). Sites were deemed suitable for sampling if a 28 cm core
 500 could be fully inserted into the sediment. During MBARI2015, two sediment cores were
 501 collected ~10cm and ~15cm distance from a site of focused hydrothermal fluid discharge close
 502 to Z-vent (Figure 1, Supplementary Figure 1). Cores were split in surface (0-7 cm) and deep (7+
 503 cm) horizons, and stored in heat sealed mylar at 4 °C, under nitrogen gas. During NA091, eight
 504 sediment cores were collected from locations across Auka vent field (Figure 1, Supplemental
 505 Figure 1). Cores were sectioned in 1-3 cm thick horizons (Supplemental Data S1 & S2). 2 mL
 506 subsamples of each horizon were frozen at -80 °C for DNA extraction, and 6 mL sediment was
 507 centrifuged at 16000g for 2 minutes to collect sediment porewater. 0.25 mL filtered porewater
 508 was preserved in 0.25 mL 0.5M zinc acetate solution for later sulfide analysis. 0.25 mL filtered
 509 porewater was stored at -20 °C for subsequent ion chromatography. During FK181031, 22
 510 sediment cores were collected from locations across Auka vent field (Figure 1, Supplemental
 511 Figure 1). Cores were sectioned in 1 or 3 cm horizons (Supplemental Data S1 & S2). 2 mL
 512 subsamples of each horizon were frozen at -80 °C for DNA extraction, and porewater was
 513 extracted from ~15 mL (1 cm horizons) or ~50 mL (3 cm horizons) sediment under nitrogen gas

514 using a pneumatic sediment squeezer (KC Denmark A/S, Silkeborg, Denmark). 0.25 mL filtered
515 porewater was preserved in 0.25 mL 0.5M zinc acetate solution for sulfide analysis. 0.25 mL
516 filtered porewater was stored at -20 °C for ion chromatography.

517

518 Geochemistry

519 All filtered water samples were stored at -20°C until analysis. Major ions were measured with a
520 Dionex ICS-2000 (Dionex, Sunnyvale, CA, USA) ion chromatography system (Environmental
521 Analysis Center, Caltech) with anion and cation columns running in parallel. An autosampler
522 loads samples diluted 1:50 in 18 MΩ water run through an LC-Pak polisher (MilliporeSigma,
523 Burlington, MA, USA) serially to a 10 µL sample loop on the anions channel, then a 10 µL
524 sample loop on the cations channel. Both columns and detectors are maintained at 30°C. The
525 ion chromatography system was run as described previously (Green-Saxena et al. 2014) with
526 the following modifications. Anions were resolved by a 2mm Dionex IonPac AS19 analytical
527 column protected by a 2mm Dionex IonPac AG19 guard column (ThermoFisher, Waltham, MA,
528 USA). A potassium hydroxide eluent generator cartridge generated a hydroxide gradient that
529 was pumped at 0.25 mL/min. The gradient was constant at 10 mM for 5 minutes, increased
530 linearly to 48.5 mM at 27 minutes, then increased linearly to 50 mM at 40 minutes. A Dionex
531 AERS 500 suppressor provided suppressed conductivity detection running on recycle mode
532 with an applied current of 30 mA. Cations were resolved by a 4mm Dionex IonPac CS16
533 analytical column protected by a 4mm Dionex IonPac CG16 guard column. A methanesulfonic
534 acid eluent generator cartridge generated a methanesulfonic acid gradient that was pumped at
535 0.36 mL/min. The gradient was constant at 10 mM for 5 minutes, nonlinearly increased to 20
536 mM at 20 min (Chromeleon curve 7, concave up), and nonlinearly increased to 40 mM at 40 min
537 (Chromeleon curve 1, concave down). A Dionex CERS 2mm suppressor provided suppressed

conductivity detection with an applied current of 32 mA. Chromatographic peaks were integrated by Chromeleon 7.2 using the Cobra algorithm, and were correlated to concentration by running known standards. The threshold of detection was approximately 10 µM for bromide and thiosulfate, 50 µM for ammonium, 100 µM for calcium, potassium, and sulfate, and 400 µM for magnesium.

Sulfide was determined colorimetrically using a protocol based on (Cline 1969). Briefly, 22 µL of a 1:1 mixture of sample and 0.5 M zinc acetate was added to 198 µL milliQ water in a 96 well plate and mixed thoroughly. Immediately after mixing, 20 µL of the diluted sample was transferred to a second 96 well plate containing 180 µL milliQ water, resulting in one 96 well plate containing 10-fold diluted samples and one 96 well plate containing 100-fold diluted samples with 200 µL final volume in each well. 20 µL of 1:1 mixture of Cline reagents (30 mM FeCl₃ • 6H₂O in 6N HCl, and 11.5 mM N,N-dimethylphenylenediamine dihydrochloride (DPDD) in 6N HCl) was added to each well and allowed to react for 1hr before measuring on a Tecan Sunrise 4.2 plate reader, using the Magellan software (version 7.3). Duplicates of all samples were run on the same 96-well plate. Sample concentration was determined by comparison to a 12 point standard curve added to each plate in duplicate, and treated identically to the samples..

DNA extraction and metagenome sequencing

The core horizons (0-7 cm and 7+ cm) from both cores obtained during MBARI2015, DR750-PC67 and DR750-PC80, were split in two subsamples each. Of these 8 samples, half was filtered through a 10 µm mesh to deplete sediment-attached organisms or those forming large aggregates, and the other half was untreated. DNA was extracted from 0.25 g wet sediment using a CTAB/phenol/chloroform organic solvent extraction protocol modified from Zhou et al. (Zhou, Bruns, and Tiedje 1996) as previously described (Speth et al. 2016). Cell lysis

was achieved in a CTAB extraction buffer (10 g L⁻¹ CTAB, 100 mM Tris, 100 mM EDTA, 100 mM sodium phosphate, 1.5 M NaCl pH 8) with serial addition of, and incubation with, lysozyme, proteinase K, and sodium dodecyl sulfate SDS. After cell lysis, DNA was recovered using phenol/chloroform/isoamylalcohol (25:24:1) at 65°C, followed by an additional chloroform/isoamylalcohol (24:1) extraction to remove traces of phenol. After recovery of the aqueous phase, DNA was precipitated using isopropanol at room temperature, and washed with ice-cold (-20 °C) 70 % ethanol and resuspended in nuclease-free water. Barcoded Nextera XT V2 libraries (Illumina) were made with dual sequencing indices, pooled, and purified with 0.75 volumes of AMPure beads (Agencourt). The resulting libraries were sequenced on a HiSeq2500 with 2x125 protocol in high output mode. (Elim Biopharmaceuticals, Hayward, CA, USA), resulting in 415 million paired reads.

573

574 Assembly and Binning

To recover high quality metagenome assembled genomes (MAGs), an iterative assembly and binning workflow was adopted. Each single dataset was subsampled to 10 million reads, assembled using SPAdes v3.14.1 (Bankevich et al. 2012) and manually binned using Anvi'o (Murat Eren et al. 2015). Reads were mapped to the manual bins, using BWA (version 0.7.12-r1039) (H. Li and Durbin 2009), filtered at >95 % identity over >80 % of the read length using BamM, (<http://ecogenomics.github.io/BamM/>) with matching reads removed from the dataset, and the process was repeated for all datasets until no more manual bins could be obtained, resulting in 50 manual bins representing abundant community members (Supplemental Data S5). Subsequently, all unmapped reads of all eight datasets, 194 million paired reads total, were pooled, co-assembled using Megahit (version 1.2.9) (D. Li et al. 2015), and binned using Metabat2 (version 2.12.2) (Kang et al. 2019), and then manually inspected

586 and corrected using Anvi'o (version 6.2) (Murat Eren et al. 2015). The automated correction
587 removed >1 % of bases in 166 metabat bins and >20 % of bases from 54 metabat bins,
588 highlighting the value of manual inspection. Bin quality was assessed using checkM (version
589 1.1.2), and the single copy marker gene sets included in Anvi'o. Bins with >50 % completeness
590 and <10% redundancy as determined by CheckM were retained. 8 additional bins with >10%
591 redundancy were included after a second manual inspection (Supplemental Data S5). This
592 workflow resulted in 331 metagenome assembled genomes (MAGs), six of which were not
593 detected in the unfiltered datasets. Those six MAGs were considered contamination and
594 removed from all subsequent analyses. Taxonomic assignment of the MAGs was done using
595 GTDB-tk (version 1.3.0) (Chaumeil et al. 2019).

596

597 MAG phylogeny and annotation

598 The two-domain phylogenetic tree of all 325 MAGs was generated using the single copy marker
599 gene sets "Archaea_76" and "Bacteria_71" included in Anvi'o, using the 25 genes shared
600 between these two single copy marker sets. Genes matching the 25 selected markers were
601 extracted, aligned using muscle (version 3.8.1551) (Edgar 2004), and the alignment
602 concatenated using "anvi-get-sequences-for-hmm-hits". The resulting concatenated alignment
603 was converted to PHYLIP format using the ElConcatenero.py script
604 (<https://github.com/ODiogoSilva/ElConcatenero>), and a phylogeny was calculated using RAXML
605 (v8.2.12) (Stamatakis 2014), with the PROTGAMMALG4X model (Le, Dang, and Gascuel 2012)
606 and the autoMRE bootstopping criterion resulting in 100 bootstrap replicates (Pattengale et al.
607 2009).

608 Gene calling on the MAGs was done using Prodigal (version 2.6.3) (Hyatt et al. 2010), and
609 predicted genes were annotated using the DIAMOND (Buchfink, Xie, and Huson 2015) against

the NCBI-NR database, PFAM (El-Gebali et al. 2019), KEGG (Kanehisa and Goto 2000), COG (Tatusov et al. 2003), CDD (Marchler-Bauer et al. 2015), EGGNOG (Huerta-Cepas et al. 2019, 2017), and CATH (Sillitoe et al. 2019; Lewis et al. 2018) (Supplemental Data S6). KEGGDecoder (Graham, Heidelberg, and Tully 2018) was used to assess broad metabolic capabilities of the MAGs.

A clade we refer to as Tharpobacteria, comprising eight genomes (five obtained from Auka and three from Guaymas), was chosen for further analysis. The metabolic capabilities of this clade were compared to 419 other *Desulfobacterota* genomes (Supplemental figure S43) by functional enrichment analysis with *anvi-compute-functional-enrichment* using KEGG module assignments from *anvi-estimate-metabolism* (Anvi'o version 7). Multiheme cytochrome c fold family proteins were obtained from the GTDB using two iterations of sequence recruitment and filtering using a bit score ratio (Rasko, Myers, and Ravel 2005) for each of the five constituent protein families. The resulting sequence sets were merged and dereplicated, and the resulting 5855 proteins sequences were classified using ASM-clust with t-distributed stochastic neighborhood embedding (tSNE) perplexity value set to 500 (Speth and Orphan 2019).

MAG reference phylogenies and optimal growth temperature prediction

Species overlap between the Auka vent field MAGs and other hydrothermal vent field MAGs was assessed using FastANI (version 1.3). 666 Guaymas basin MAGs (Dombrowski et al. 2017; Dombrowski, Teske, and Baker 2018; Seitz et al. 2019) and 99 MAGs obtained from Juan de Fuca ridge hydrothermal fluid (Jungbluth, Amend, and Rappé 2017) were downloaded from NCBI. 348 bins from the Cayman Rise hydrothermal vent field (Anderson et al. 2017) were retrieved as Anvi'o databases (version 2.1.0) from Figshare (https://figshare.com/projects/Mid-Cayman_Rise_Metagenome_Assembled_Genomes/20783).

634 Contig fasta files were exported from the Anvi'o databases and 131 bins over 50%
635 completeness were used in ANI analysis.

636 Detailed phylogenetic placement of the Auka MAGs was obtained by downloading the genomes
637 comprising the identified phyla, as well as sister phyla, from the genome taxonomy database
638 (GTDB, version 89) (Parks et al. 2020). In addition 666 MAGs recently obtained from Guaymas
639 basin were included in the reference set (Dombrowski et al. 2017; Dombrowski, Teske, and
640 Baker 2018; Seitz et al. 2019). The genome set was split in 14 subsets, based on taxonomy and
641 the GTDB reference tree, for alignment and tree calculation. Anvi'o databases were generated
642 for all genomes, and the "Archaea_76" and "Bacteria_71" gene sets included in Anvi'o were
643 used to generate concatenated alignments for Archaeal and Bacterial subsets, respectively,
644 using muscle (version 3.8.1551) (Edgar 2004). Phylogenies were calculated using FastTree
645 (version 2.1.7) (Price, Dehal, and Arkin 2010).

646 Optimal growth temperatures (OGT) were predicted for the 325 MAGs and the reference
647 genomes using the method described by Sauer and Wang (Sauer and Wang 2019)
648 (https://github.com/DavidBSauer/OGT_prediction). Regression models modified by David Sauer
649 to exclude 16S rRNA and genome size, to account for absence of 16S rRNA and genome
650 incompleteness, were downloaded from
651 [https://github.com/DavidBSauer/OGT_prediction/tree/master/data/calculations/prediction/regres](https://github.com/DavidBSauer/OGT_prediction/tree/master/data/calculations/prediction/regression_models)
652 [sion_models](https://github.com/DavidBSauer/OGT_prediction/tree/master/data/calculations/prediction/regression_models). The regression models for "Superkingdom Archaea" and "Superkingdom Bacteria"
653 were chosen because of the diversity of the organisms in the dataset. The
654 prediction_pipeline.py script was run as described in the documentation, with a custom
655 "genomes_retrieved.txt" file (two tab-delimited columns, no header, columns: filename of gzip
656 compressed contig fasta <tab> genome ID) and "species_taxonomic.txt" file (two tab-delimited

657 columns, headers: "species <tab> superkingdom", columns: "Genome ID <tab>
658 Archaea|Bacteria").

659

660 16S rRNA gene analyses

661 To gain insight in the microbial diversity of the Auka vent field, 216 samples derived from 30
662 sediment cores, and 8 samples from microbial mats or biofilms, were used for DNA extraction
663 using the Qiagen Dneasy PowerSoil kit (Valencia, CA, USA) following the manufacturer's
664 protocol, with the exception that cells were lysed using MP Biomedicals FastPrep-24 (Irvine,
665 CA, USA) for 45s, at 5.5 m s⁻¹. The V4-V5 region of the 16S rRNA gene was PCR amplified
666 from the resulting 224 DNA extracts using the 515f/926r primer set (Walters et al. 2016)
667 modified with Illumina (San Diego, CA, USA) adapters on 5' end (515F
668 5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-GTGYCAGCMGCCGCGGTAA-3'and
669 926R

670 5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-CCGYCAATTYMTTTRAGTTT-3').

671 Duplicate PCR reactions were set up for each sample with Q5 Hot Start High-Fidelity 2x Master
672 Mix (New England Biolabs, Ipswich, MA, USA) in a 15 µL reaction volume, with annealing at
673 54°C, for 28 cycles. The number of cycles was increased if no product was obtained, as detailed
674 in the sample metadata (Supplemental Data S2). Duplicate PCR samples were then pooled and
675 barcoded with Illumina Nextera XT index 2 primers that include unique 8-bp barcodes (P5
676 5'-AATGATACGGCGACCACCGAGATCTACAC-XXXXXXXXX-TCGTCGGCAGCGTC-3' and P7
677 5'-CAAGCAGAAGACGGCATACGAGAT-XXXXXXXXX-GTCTCGTGGGCTCGG-3'). Amplification
678 with barcoded primers used Q5 Hot Start PCR mixture but used 2.5 µL of product in 25 µL of
679 total reaction volume, annealed at 66°C, and cycled 10 times. Products were purified using
680 Millipore-Sigma (St. Louis, MO, USA) MultiScreen Plate MSNU03010 with vacuum manifold and

681 quantified using ThermoFisher Scientific (Waltham, MA, USA) QuantIT PicoGreen dsDNA
 682 Assay Kit P11496 on the BioRad CFX96 Touch Real-Time PCR Detection System. Barcoded
 683 samples were combined in equimolar amounts into single tubes and purified with Qiagen PCR
 684 Purification Kit 28104 before sequencing on Illumina MiSeq with the addition of 15-20% PhiX
 685 and with either a 2x250 or a 2x300 protocol (Laragen Inc., Culver City, CA, USA). Demultiplexed
 686 sequencing data was processed using QIIME2 (v2020.2) (Bolyen et al. 2019), with DADA2 for
 687 amplicon sequence variant (ASV) calling (Callahan et al. 2016) and Cutadapt for sequence
 688 quality trimming and primer removal (Martin 2011). resulted in 18777 ASVs representing
 689 6,244,164 trimmed and filtered reads.

690 16S rRNA genes were recovered from the MAGs using the HMMs included in Anvi'o (Murat
 691 Eren et al. 2015), and by *de novo* assembly of the combined metagenomic sequencing reads
 692 using phyloFlash (Gruber-Vodicka, Seah, and Pruesse 2020). After combining and dereplication
 693 of the two sets of sequences at 97% identity using usearch (v11.0.667) (Edgar 2010), the
 694 resulting 284 sequences were aligned using muscle (v3.8.1551) (Edgar 2004), and a
 695 phylogenetic tree was calculated using RAxML (v8.2.12) (Stamatakis 2014), with the
 696 GTRGAMMA model and the autoMRE bootstrapping criterion resulting in 400 bootstrap
 697 replicates (Pattengale et al. 2009). Metagenome abundance of the 284 dereplicated 16S rRNA
 698 gene sequences was determined by mapping the combined reads of the shallow (0-7cm) and
 699 deep (7+ cm) horizons of both cores onto the 16S sequences using bbmap, with a 95% identity
 700 filter (Bushnell 2014). Approximately 0.06% of the reads mapped on the genes, which is
 701 consistent with expectation based on the 16S rRNA gene, and average genome length.
 702 Abundance of the 284 dereplicated 16S rRNA gene sequences in the amplicon sequencing
 703 (ITag) data was determined by aligning all 18,777 ASVs to the 16S rRNA genes using BLAST

704 version 2.10.1 (Altschul et al. 1990), and summing the number of reads represented by all ASVs
705 with >97% identity to the assembled 16S rRNA genes.

706 **Acknowledgments**

707 We thank the pilots, crew, and participants on the cruises to the southern Gulf of California: R/V
 708 Western Flyer operated by the Monterey Bay Aquarium Research Institute (MBARI), E/V
 709 Nautilus operated by the Ocean Exploration Trust, with cruise NA091 supported by the Dalio
 710 Foundation and Woods Hole Oceanographic Institute, and R/V Falkor operated by the Schmidt
 711 Ocean Institute. We appreciate the support and opportunity to sail with chief scientists Scott
 712 Wankel and Anna Michel on NA091. We also thank David W. Caress and Jennifer B. Paduan
 713 (MBARI) for providing the high resolution version of the bathymetric map in Figure 1, David
 714 Sauer for recalculating the optimal growth temperature prediction model to be appropriate for
 715 metagenomics, and Haley Sapers for providing a framework for the 16S rRNA gene amplicon
 716 processing. This research used samples provided by the Ocean Exploration Trust's Nautilus
 717 Exploration Program, cruise NA091.

718

719 **Data availability:**

720 Raw metagenome reads, assembled metagenome bins, and 16S rRNA gene amplicon
 721 sequencing data are available in GenBank under BioProject accession number PRJNA713414

722

723 **Funding:**

724 This work was supported by the Center for Dark Energy Biosphere investigations (C-DEBI),
 725 Canadian Institute for Advanced Science (CIFAR), the US Department of Energy, Office of
 726 Science, Office of Biological and Environmental Research under award number DE-SC0016469
 727 to Victoria J. Orphan. Daan R. Speth was supported by the Netherlands Organisation for
 728 Scientific Research, Rubicon award 019.153LW.039. Feiqiao B. Yu and Stephen R. Quake were
 729 supported by the John Templeton Foundation grant 51250 and the Chan Zuckerberg Biohub.

730 Victoria J. Orphan is a CIFAR fellow in the Earth 4D program.

731 Sample collection permits were granted by la Dirección General de Ordenamiento Pesquero y

732 Acuícola, Comisión Nacional de Acuacultura y Pesca (CONAPESCA: Permiso de Pesca de

733 Fomento No. PPFE/DGOPA-200/18) and la Dirección General de Geografía y Medio Ambiente,

734 Instituto Nacional de Estadística y Geografía (INEGI: Autorización EG0122018), with the

735 associated Diplomatic Note number 18-2083 (CTC/07345/18) from la Secretaría de Relaciones

736 Exteriores - Agencia Mexicana de Cooperación Internacional para el Desarrollo / Dirección

737 General de Cooperación Técnica y Científica. Sample collection permit for cruise NA091 was

738 obtained by the Ocean Exploration Trust under permit number EG0072017.

739

740 Figures and figure legends:

742 **Figure 1.** Overview of sampling area and impressions of Auka vent field.

743 **A)** Bathymetric map of Auka vent field with the most prominent sites of focused hydrothermal
744 venting labeled by name. The sampling locations of push cores analyzed in this study are
745 indicated, with the cores used for metagenomic sequencing highlighted in orange. **B)** Diane's
746 vent, a chimney with distinctive clear hydrothermal fluid discharge clearly visible. **C)** Top of the
747 Matterhorn, a ~10m high free standing chimney, with the area around the central orifice fully
748 covered in *Oasisia sp.* tubeworms. Microbial mat and shallow flange structures are visible on

749 the sides of the chimney, indicative of diffuse fluid discharge through the chimney wall. **D)**
 750 Carbonate platform covered in white and grey microbial mat with *Oasisia sp.* tubeworms (~20cm
 751 tall) clustered around a localized spot of hydrothermal fluid discharge. This is a representative
 752 example of the unlabeled elevated mounds shown in panel A. **E)** Sediment covered in microbial
 753 mat (at Diane's vent) with heterogeneity of colors and textures indicating centimeter scale
 754 spatial heterogeneity in fluid diffusion through the sediment. Pushcore locations in panel A
 755 correspond to areas with prevalent microbial mats.

756

Figure 2. Concatenated marker gene phylogeny of the 325 Auka MAGs.

Phylogeny of the 325 MAGs recovered from Auka vent field sediments, based on 25 concatenated marker genes. The scale bar indicates 1 substitution per site. From inside to outside, the concentric circles around the phylogeny indicate: the MAG ID, the average nucleotide identity (ANI) with MAGs previously retrieved from Guaymas Basin, phylum level taxonomy, MAG predicted optimal growth temperature (OGT), MAG abundance in the surface (0-7cm) and deep (7+cm) section of core DR750-PC67, and MAG abundance in the the abundance in the surface (0-7cm) and deep (7+cm) section of core DR750-PC80. Numbers in parentheses indicate the number of MAGs belonging to that lineage in the dataset (All MAGs are shown in the figure). The Tharpobacteria branch (see text) is highlighted in pink. The predicted OGT of PBMBMC_261 (111 °C) was outside the scale, and is indicated with an

769 asterisk. Phyla corresponding to abbreviated groups in the taxonomy legend: *1 UBP7 (1),
770 Ratteibacteria (2), Omnitrophota (6), Calescibacterota (2), Aerophobota (7); **Fermentibacterota
771 (1), Krumholzibacteriota (1), Cloacimonadota (4), Latescibacterota (3), Zixibacteria (2) KSB1 (1)
772 SM23-31 (1), Calditrichota (1), Marinisomatota (6); ***Proteobacteria (2), Myxococcota (1),
773 Desulfuromonadota (3), Desulfobacterota_A (5), Desulfobacterota (23); ****UBP3 (1),
774 Sumerlaeota (1), RBG-13-66-14 (1), Poribacteria (1), Hydrogenedentota (1), Eremiobacterota
775 (1), Firmicutes (1), Firmicutes_A (2); *****Caldatribacteriota (2), Synergistota (3), Caldisericota
776 (3), Bipolaricaulota (6), Thermotogota (11).

Figure 3. 16 rRNA gene phylogeny and abundance in metagenome and amplicon data. Phylogeny of 284 16S rRNA genes reconstructed from the metagenomic data. From inside to outside the concentric circles around the phylogenetic tree indicate: the taxonomy of the major clades in the phylogeny, as assigned by Silva138, the recovery of the sequence through phyloFlash and/or annotation of the retrieved MAGs, the abundance in the surface (0-7cm) and deep (7+cm) section of the cores used for metagenomic sequencing, the abundance in the surface (0-7cm) and deep (7+cm) sections of the cores retrieved from Z-vent through amplicon sequencing, and the abundance in the shallow (0-7cm) and deep (7+cm) sections of the cores retrieved from Diane's vent through amplicon sequencing. Circled numbers highlight abundant taxa discussed in the main text.

795 DR750-PC80 surface horizon (0-7 cm) and DR750-PC80 deep horizon (7+ cm). **C)** Comparing
 796 MAG abundance between the surface horizons of both cores. **D)** Comparing MAG abundance
 797 between the deep horizons of both cores.

798

799

Figure 5. Community similarity between Pescadero Basin and other hydrothermal sites.

A) Histograms of the Auka MAGs with average nucleotide identity (ANI) greater than 75% with MAGs obtained from Guaymas Basin (161), Cayman Rise (12), and Juan de Fuca ridge (9), indicating a high community similarity between Auka and Guaymas Basin. **B)** Scatterplot of Auka vs Guaymas Basin (GB) MAG ANI and predicted optimal growth temperature (OGT), showing no correlation between OGT and ANI. **C)** Scatterplot of Auka vs GB MAG ANI and MAG abundance in the surface horizons, indicating no correlation between surface abundance

807 and ANI. D) Scatterplot of Auka vs GB MAG ANI and MAG abundance in the deep horizons,
808 indicating no correlation between abundance in deeper horizons and ANI.

Figure 6. HAO family protein as proposed candidates for reduction of the terminal electron acceptor in Tharpobacteria clade.

A) Schematic overview of components of the electron transport chain detected in Tharpobacteria genomes, with the octaheme cytochrome c proposed to be involved in terminal electron reduction in Tharpobacteria indicated in red. **B)** t-distributed stochastic neighborhood embedding (tSNE) representation of alignment score matrix of 5855 proteins of the HAO/OTR/ONR/nrfA/MCC/c554 structural fold family. Each point represents a protein sequence, colored by taxonomic affiliation at the phylum level. Dashed ellipse indicates sequences included in the phylogeny in panel C. HAO: hydroxylamine oxidoreductase, OTR: octaheme tetrathionate reductase, ONR: octaheme nitrite reductase, nrfA: pentaheme nitrite reductase, MCC: octaheme sulfite reductase, c554: tetraheme cytochrome c554. **C)** Approximate maximum likelihood phylogeny of 1502 HAO family protein sequences. The outer ring indicates taxonomic affiliation of the 5 most represented phyla, white space indicates other

823 phyla. Sequences known to be involved in nitrogen cycling are indicated with a grey arc.
 824 Discussed Tharpobacteria clade sequences are indicated with an arrow, sequences of
 825 structures included in the promals3D alignment are indicated with asterisks, known clades
 826 involved in nitrogen cycling are highlighted with 3 letter abbreviations. The function of several
 827 HAO family members in anammox *Planctomycetota* is unknown. AOB: ammonia oxidizing
 828 bacteria, MOB: methane oxidizing bacteria, CMX: comammox *Nitrospirota*, AMX: anammox
 829 *Planctomycetota*.

830

831 References

- 832 Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. "Basic Local Alignment
833 Search Tool." *Journal of Molecular Biology* 215 (3): 403–10.
- 834 Anantharaman, Karthik, John A. Breier, and Gregory J. Dick. 2016. "Metagenomic Resolution of
835 Microbial Functions in Deep-Sea Hydrothermal Plumes across the Eastern Lau Spreading
836 Center." *The ISME Journal* 10 (1): 225–39.
- 837 Anderson, Rika E., Mónica Torres Beltrán, Steven J. Hallam, and John A. Baross. 2013.
838 "Microbial Community Structure across Fluid Gradients in the Juan de Fuca Ridge
839 Hydrothermal System." *FEMS Microbiology Ecology* 83 (2): 324–39.
- 840 Anderson, Rika E., Julie Reveillaud, Emily Reddington, Tom O. Delmont, A. Murat Eren, Jill M.
841 McDermott, Jeff S. Seewald, and Julie A. Huber. 2017. "Genomic Variation in Microbial
842 Populations Inhabiting the Marine Subseafloor at Deep-Sea Hydrothermal Vents." *Nature*
843 *Communications* 8 (1): 1114.
- 844 Anderson, Rika E., Mitchell L. Sogin, and John A. Baross. 2015. "Biogeography and Ecology of
845 the Rare and Abundant Microbial Lineages in Deep-Sea Hydrothermal Vents." *FEMS*
846 *Microbiology Ecology* 91 (1): 1–11.
- 847 Bankevich, Anton, Sergey Nurk, Dmitry Antipov, Alexey A. Gurevich, Mikhail Dvorkin, Alexander
848 S. Kulikov, Valery M. Lesin, et al. 2012. "SPAdes: A New Genome Assembly Algorithm and
849 Its Applications to Single-Cell Sequencing." *Journal of Computational Biology: A Journal of*
850 *Computational Molecular Cell Biology* 19 (5): 455–77.
- 851 Biddle, Jennifer F., Zena Cardman, Howard Mendlovitz, Daniel B. Albert, Karen G. Lloyd, Antje
852 Boetius, and Andreas Teske. 2012. "Anaerobic Oxidation of Methane at Different
853 Temperature Regimes in Guaymas Basin Hydrothermal Sediments." *The ISME Journal* 6
854 (5): 1018–31.

855 Blöchl, E., R. Rachel, S. Burggraf, D. Hafenbradl, H. W. Jannasch, and K. O. Stetter. 1997.
856 “Pyrolobus Fumarii, Gen. and Sp. Nov., Represents a Novel Group of Archaea, Extending
857 the Upper Temperature Limit for Life to 113 Degrees C.” *Extremophiles: Life under Extreme*
858 *Conditions* 1 (1): 14–21.

859 Boll, M., and G. Fuchs. 1995. “Benzoyl-Coenzyme A Reductase (dearomatizing), a Key Enzyme
860 of Anaerobic Aromatic Metabolism. ATP Dependence of the Reaction, Purification and
861 Some Properties of the Enzyme from Thauera Aromatica Strain K172.” *European Journal*
862 *of Biochemistry / FEBS* 234 (3): 921–33.

863 Bolyen, Evan, Jai Ram Rideout, Matthew R. Dillon, Nicholas A. Bokulich, Christian C. Abnet,
864 Gabriel A. Al-Ghalith, Harriet Alexander, et al. 2019. “Reproducible, Interactive, Scalable
865 and Extensible Microbiome Data Science Using QIIME 2.” *Nature Biotechnology* 37 (8):
866 852–57.

867 Buchfink, Benjamin, Chao Xie, and Daniel H. Huson. 2015. “Fast and Sensitive Protein
868 Alignment Using DIAMOND.” *Nature Methods* 12 (1): 59–60.

869 Bushnell, Brian. 2014. “BBMap: A Fast, Accurate, Splice-Aware Aligner.” Lawrence Berkeley
870 National Lab.(LBNL), Berkeley, CA (United States). <https://www.osti.gov/biblio/1241166>.

871 Callahan, Benjamin J., Paul J. McMurdie, Michael J. Rosen, Andrew W. Han, Amy Jo A.
872 Johnson, and Susan P. Holmes. 2016. “DADA2: High-Resolution Sample Inference from
873 Illumina Amplicon Data.” *Nature Methods* 13 (7): 581–83.

874 Chaumeil, Pierre-Alain, Aaron J. Mussig, Philip Hugenholtz, and Donovan H. Parks. 2019.
875 “GTDB-Tk: A Toolkit to Classify Genomes with the Genome Taxonomy Database.”
876 *Bioinformatics* , November. <https://doi.org/10.1093/bioinformatics/btz848>.

877 Chuvochina, Maria, Christian Rinke, Donovan H. Parks, Michael S. Rappé, Gene W. Tyson,
878 Pelin Yilmaz, William B. Whitman, and Philip Hugenholtz. 2019. “The Importance of

879 Designating Type Material for Uncultured Taxa.” *Systematic and Applied Microbiology* 42
880 (1): 15–21.

881 Cline, Joel D. 1969. “SPECTROPHOTOMETRIC DETERMINATION OF HYDROGEN SULFIDE
882 IN NATURAL WATERS1.” *Limnology and Oceanography*.
883 <https://doi.org/10.4319/lo.1969.14.3.0454>.

884 Dick, Gregory J. 2019. “The Microbiomes of Deep-Sea Hydrothermal Vents: Distributed
885 Globally, Shaped Locally.” *Nature Reviews. Microbiology* 17 (5): 271–83.

886 Dick, Gregory J., and Bradley M. Tebo. 2010. “Microbial Diversity and Biogeochemistry of the
887 Guaymas Basin Deep-Sea Hydrothermal Plume.” *Environmental Microbiology* 12 (5):
888 1334–47.

889 Ding, Jian, Yu Zhang, Han Wang, Huahua Jian, Hao Leng, and Xiang Xiao. 2017. “Microbial
890 Community Structure of Deep-Sea Hydrothermal Vents on the Ultraslow Spreading
891 Southwest Indian Ridge.” *Frontiers in Microbiology* 8 (June): 1012.

892 Dombrowski, Nina, Kiley W. Seitz, Andreas P. Teske, and Brett J. Baker. 2017. “Genomic
893 Insights into Potential Interdependencies in Microbial Hydrocarbon and Nutrient Cycling in
894 Hydrothermal Sediments.” *Microbiome* 5 (1): 106.

895 Dombrowski, Nina, Andreas P. Teske, and Brett J. Baker. 2018. “Expansive Microbial Metabolic
896 Versatility and Biodiversity in Dynamic Guaymas Basin Hydrothermal Sediments.” *Nature*
897 *Communications* 9 (1): 4999.

898 Dong, Xiyang, Chris Greening, Jayne E. Rattray, Anirban Chakraborty, Maria Chuvochina,
899 Daisuke Mayumi, Jan Dolfing, et al. 2019. “Metabolic Potential of Uncultured Bacteria and
900 Archaea Associated with Petroleum Seepage in Deep-Sea Sediments.” *Nature*
901 *Communications* 10 (1): 1816.

902 Dowell, Frederick, Zena Cardman, Srishti Dasarathy, Matthias Y. Kellermann, Julius S. Lipp, S.

903 Emil Ruff, Jennifer F. Biddle, et al. 2016. "Microbial Communities in Methane- and Short
904 Chain Alkane-Rich Hydrothermal Sediments of Guaymas Basin." *Frontiers in Microbiology*
905 7 (January): 17.

906 Edgar, Robert C. 2004. "MUSCLE: A Multiple Sequence Alignment Method with Reduced Time
907 and Space Complexity." *BMC Bioinformatics* 5 (August): 113.

908 ———. 2010. "Search and Clustering Orders of Magnitude Faster than BLAST." *Bioinformatics*
909 26 (19): 2460–61.

910 El-Gebali, Sara, Jaina Mistry, Alex Bateman, Sean R. Eddy, Aurélien Luciani, Simon C. Potter,
911 Matloob Qureshi, et al. 2019. "The Pfam Protein Families Database in 2019." *Nucleic Acids*
912 *Research* 47 (D1): D427–32.

913 Espinosa-Asuar, Laura, Luis A. Soto, Diana L. Salcedo, Abril Hernández-Monroy, Luis E.
914 Eguiarte, Valeria Souza, and Patricia Velez. 2020. "Bacterial Communities from Deep
915 Hydrothermal Systems: The Southern Gulf of California as an Example of Primeval
916 Environments." In *Astrobiology and Cuatro Ciénegas Basin as an Analog of Early Earth*,
917 edited by Valeria Souza, Antígona Segura, and Jamie S. Foster, 149–66. Cham: Springer
918 International Publishing.

919 Goffredi, Shana K., Shannon Johnson, Verena Tunnicliffe, David Caress, David Clague, Elva
920 Escobar, Lonny Lundsten, et al. 2017. "Hydrothermal Vent Fields Discovered in the
921 Southern Gulf of California Clarify Role of Habitat in Augmenting Regional Diversity."
922 *Proceedings. Biological Sciences / The Royal Society* 284 (1859).
923 <https://doi.org/10.1098/rspb.2017.0817>.

924 Gonnella, Giorgio, Stefanie Böhnke, Daniela Indenbirken, Dieter Garbe-Schönberg, Richard
925 Seifert, Christian Mertens, Stefan Kurtz, and Mirjam Perner. 2016. "Endemic Hydrothermal
926 Vent Species Identified in the Open Ocean Seed Bank." *Nature Microbiology* 1 (8): 16086.

927 Graham, E. D., J. F. Heidelberg, and B. J. Tully. 2018. "Potential for Primary Productivity in a
928 Globally-Distributed Bacterial Phototroph." *The ISME Journal* 12 (7): 1861–66.

929 Green-Saxena, A., A. E. Dekas, N. F. Dalleska, and V. J. Orphan. 2014. "Nitrate-Based Niche
930 Differentiation by Distinct Sulfate-Reducing Bacteria Involved in the Anaerobic Oxidation of
931 Methane." *The ISME Journal* 8 (1): 150–63.

932 Grimaldi, Stéphane, Barbara Schoepp-Cothenet, Pierre Ceccaldi, Bruno Guigliarelli, and Axel
933 Magalon. 2013. "The Prokaryotic Mo/W-bisPGD Enzymes Family: A Catalytic Workhorse in
934 Bioenergetic." *Biochimica et Biophysica Acta* 1827 (8-9): 1048–85.

935 Gruber-Vodicka, Harald R., Brandon K. B. Seah, and Elmar Priesse. 2020. "phyloFlash: Rapid
936 Small-Subunit rRNA Profiling and Targeted Assembly from Metagenomes." *mSystems* 5
937 (5). <https://doi.org/10.1128/mSystems.00920-20>.

938 Haase, Doreen, Bianca Hermann, Oliver Einsle, and Jörg Simon. 2017. "Epsilonproteobacterial
939 Hydroxylamine Oxidoreductase (ϵ Hao): Characterization of a 'Missing Link' in the
940 Multihaem Cytochrome c Family." *Molecular Microbiology* 105 (1): 127–38.

941 Haberstroh, P. R., and D. M. Karl. 1989. "Dissolved Free Amino Acids in Hydrothermal Vent
942 Habitats of the Guaymas Basin." *Geochimica et Cosmochimica Acta* 53 (11): 2937–45.

943 Hao, Liping, Thomas Yssing Michaelsen, Caitlin Margaret Singleton, Giulia Dottorini, Rasmus
944 Hansen Kirkegaard, Mads Albertsen, Per Halkjær Nielsen, and Morten Simonsen Dueholm.
945 2020. "Novel Syntrophic Bacteria in Full-Scale Anaerobic Digesters Revealed by
946 Genome-Centric Metatranscriptomics." *The ISME Journal* 14 (4): 906–18.

947 Hermann, Bianca, Melanie Kern, Luigi La Pietra, Jörg Simon, and Oliver Einsle. 2015. "The
948 Octahaem MccA Is a Haem c-Copper Sulfite Reductase." *Nature* 520 (7549): 706–9.

949 Holler, Thomas, Friedrich Widdel, Katrin Knittel, Rudolf Amann, Matthias Y. Kellermann,
950 Kai-Uwe Hinrichs, Andreas Teske, Antje Boetius, and Gunter Wegener. 2011. "Thermophilic

951 Anaerobic Oxidation of Methane by Marine Microbial Consortia.” *The ISME Journal* 5 (12):
952 1946–56.

953 Huang, Jiao-Mei, Brett J. Baker, Jiang-Tao Li, and Yong Wang. 2019. “New Microbial Lineages
954 Capable of Carbon Fixation and Nutrient Cycling in Deep-Sea Sediments of the Northern
955 South China Sea.” *Applied and Environmental Microbiology* 85 (15).
956 <https://doi.org/10.1128/AEM.00523-19>.

957 Hubert, Casey, Alexander Loy, Maren Nickel, Carol Arnosti, Christian Baranyi, Volker Brüchert,
958 Timothy Ferdelman, et al. 2009. “A Constant Flux of Diverse Thermophilic Bacteria into the
959 Cold Arctic Seabed.” *Science* 325 (5947): 1541–44.

960 Huerta-Cepas, Jaime, Kristoffer Forslund, Luis Pedro Coelho, Damian Szklarczyk, Lars Juhl
961 Jensen, Christian von Mering, and Peer Bork. 2017. “Fast Genome-Wide Functional
962 Annotation through Orthology Assignment by eggNOG-Mapper.” *Molecular Biology and
963 Evolution* 34 (8): 2115–22.

964 Huerta-Cepas, Jaime, Damian Szklarczyk, Davide Heller, Ana Hernández-Plaza, Sofia K.
965 Forslund, Helen Cook, Daniel R. Mende, et al. 2019. “eggNOG 5.0: A Hierarchical,
966 Functionally and Phylogenetically Annotated Orthology Resource Based on 5090
967 Organisms and 2502 Viruses.” *Nucleic Acids Research* 47 (D1): D309–14.

968 Hyatt, Doug, Gwo-Liang Chen, Philip F. Locascio, Miriam L. Land, Frank W. Larimer, and Loren
969 J. Hauser. 2010. “Prodigal: Prokaryotic Gene Recognition and Translation Initiation Site
970 Identification.” *BMC Bioinformatics* 11 (March): 119.

971 Igarashi, N., H. Moriyama, T. Fujiwara, Y. Fukumori, and N. Tanaka. 1997. “The 2.8 Å Structure
972 of Hydroxylamine Oxidoreductase from a Nitrifying Chemoautotrophic Bacterium,
973 *Nitrosomonas Europaea*.” *Nature Structural Biology* 4 (4): 276–84.

974 Imachi, Hiroyuki, Masaru K. Nobu, Nozomi Nakahara, Yuki Morono, Miyuki Ogawara, Yoshihiro

975 Takaki, Yoshinori Takano, et al. 2020. "Isolation of an Archaeon at the
976 Prokaryote-Eukaryote Interface." *Nature* 577 (7791): 519–25.

977 Jain, Chirag, Luis M. Rodriguez-R, Adam M. Phillippy, Konstantinos T. Konstantinidis, and
978 Srinivas Aluru. 2018. "High Throughput ANI Analysis of 90K Prokaryotic Genomes Reveals
979 Clear Species Boundaries." *Nature Communications* 9 (1): 5114.

980 Jungbluth, Sean P., Jan P. Amend, and Michael S. Rappé. 2017. "Metagenome Sequencing and
981 98 Microbial Genomes from Juan de Fuca Ridge Flank Subsurface Fluids." *Scientific Data*
982 4 (March): 170037.

983 Kanehisa, M., and S. Goto. 2000. "KEGG: Kyoto Encyclopedia of Genes and Genomes."
984 *Nucleic Acids Research* 28 (1): 27–30.

985 Kang, Dongwan D., Feng Li, Edward Kirton, Ashleigh Thomas, Rob Egan, Hong An, and Zhong
986 Wang. 2019. "MetaBAT 2: An Adaptive Binning Algorithm for Robust and Efficient Genome
987 Reconstruction from Metagenome Assemblies." *PeerJ* 7 (July): e7359.

988 Kartal, Boran, and Jan T. Keltjens. 2016. "Anammox Biochemistry: A Tale of Heme c Proteins."
989 *Trends in Biochemical Sciences* 41 (12): 998–1011.

990 Kashefi, Kazem, and Derek R. Lovley. 2003. "Extending the Upper Temperature Limit for Life."
991 *Science* 301 (5635): 934.

992 Kniemeyer, Olaf, Florin Musat, Stefan M. Sievert, Katrin Knittel, Heinz Wilkes, Martin
993 Blumenberg, Walter Michaelis, et al. 2007. "Anaerobic Oxidation of Short-Chain
994 Hydrocarbons by Marine Sulphate-Reducing Bacteria." *Nature* 449 (7164): 898–901.

995 Krukenberg, Viola, Katie Harding, Michael Richter, Frank Oliver Glöckner, Harald R.
996 Gruber-Vodicka, Birgit Adam, Jasmine S. Berg, et al. 2016. "Candidatus Desulfotomaculum
997 Auxilii, a Hydrogenotrophic Sulfate-Reducing Bacterium Involved in the Thermophilic
998 Anaerobic Oxidation of Methane." *Environmental Microbiology* 18 (9): 3073–91.

999 Kurr, Margit, Robert Huber, Helmut König, Holger W. Jannasch, Hans Fricke, Antonio Trincone,
1000 Jakob K. Kristjansson, and Karl O. Stetter. 1991. "Methanopyrus Kandleri, Gen. and Sp.
1001 Nov. Represents a Novel Group of Hyperthermophilic Methanogens, Growing at 110°C."
1002 *Archives of Microbiology* 156 (4): 239–47.

1003 Langwig, Marguerite V., Valerie De Anda, Nina Dombrowski, Kiley W. Seitz, Ian M. Rambo,
1004 Chris Greening, Andreas P. Teske, and Brett J. Baker. 2021. "Large-Scale Protein Level
1005 Comparison of Deltaproteobacteria Reveals Cohesive Metabolic Groups." *The ISME*
1006 *Journal*, July. <https://doi.org/10.1038/s41396-021-01057-y>.

1007 Laso-Pérez, Rafael, Gunter Wegener, Katrin Knittel, Friedrich Widdel, Katie J. Harding, Viola
1008 Krukenberg, Dimitri V. Meier, et al. 2016. "Thermophilic Archaea Activate Butane via
1009 Alkyl-Coenzyme M Formation." *Nature* 539 (7629): 396–401.

1010 Le, Si Quang, Cuong Cao Dang, and Olivier Gascuel. 2012. "Modeling Protein Evolution with
1011 Several Amino Acid Replacement Matrices Depending on Site Rates." *Molecular Biology*
1012 *and Evolution* 29 (10): 2921–36.

1013 Lewis, Tony E., Ian Sillitoe, Natalie Dawson, Su Datt Lam, Tristan Clarke, David Lee, Christine
1014 Orengo, and Jonathan Lees. 2018. "Gene3D: Extensive Prediction of Globular Domains in
1015 Proteins." *Nucleic Acids Research* 46 (D1): D435–39.

1016 Li, Dinghua, Chi-Man Liu, Ruibang Luo, Kunihiro Sadakane, and Tak-Wah Lam. 2015.
1017 "MEGAHIT: An Ultra-Fast Single-Node Solution for Large and Complex Metagenomics
1018 Assembly via Succinct de Bruijn Graph." *Bioinformatics* 31 (10): 1674–76.

1019 Li, Heng, and Richard Durbin. 2009. "Fast and Accurate Short Read Alignment with
1020 Burrows-Wheeler Transform." *Bioinformatics* 25 (14): 1754–60.

1021 Lizarralde, Daniel, Gary J. Axen, Hillary E. Brown, John M. Fletcher, Antonio
1022 González-Fernández, Alistair J. Harding, W. Steven Holbrook, et al. 2007. "Variation in

- 1023   Styles of Rifting in the Gulf of California.” *Nature* 448 (7152): 466–69.
- 1024 Lutz, Richard A., Timothy M. Shank, George W. Luther, Costantino Vetriani, Maya Tolstoy,
1025   Donald B. Nuzzio, Tommy S. Moore, et al. 2008. “Interrelationships Between Vent Fluid
1026   Chemistry, Temperature, Seismic Activity, and Biological Community Structure at a
1027   Mussel-Dominated, Deep-Sea Hydrothermal Vent Along the East Pacific Rise.” *Journal of*
1028   *Shellfish Research* 27 (1): 177–90.
- 1029 Marchler-Bauer, Aron, Myra K. Derbyshire, Noreen R. Gonzales, Shennan Lu, Farideh Chitsaz,
1030   Lewis Y. Geer, Renata C. Geer, et al. 2015. “CDD: NCBI’s Conserved Domain Database.”
1031   *Nucleic Acids Research* 43 (Database issue): D222–26.
- 1032 Martin, Marcel. 2011. “Cutadapt Removes Adapter Sequences from High-Throughput
1033   Sequencing Reads.” *EMBnet.journal* 17 (1): 10–12.
- 1034 McKay, Luke, Vincent W. Klokman, Howard P. Mendlovitz, Douglas E. LaRowe, Daniel R. Hoer,
1035   Daniel Albert, Jan P. Amend, and Andreas Teske. 2016. “Thermal and Geochemical
1036   Influences on Microbial Biogeography in the Hydrothermal Sediments of Guaymas Basin,
1037   Gulf of California.” *Environmental Microbiology Reports* 8 (1): 150–61.
- 1038 Merewether, Ray, Mark S. Olsson, and Peter Lonsdale. 1985. “Acoustically Detected
1039   Hydrocarbon Plumes Rising from 2-Km Depths in Guaymas Basin, Gulf of California.”
1040   *Journal of Geophysical Research* 90 (B4): 3075.
- 1041 Metcalfe, Kyle S., Ranjani Murali, Sean W. Mullin, Stephanie A. Cannon, and Victoria J. Orphan.
1042   2021. “Experimentally-Validated Correlation Analysis Reveals New Anaerobic Methane
1043   Oxidation Partnerships with Consortium-Level Heterogeneity in Diazotrophy.” *The ISME*
1044   *Journal* 15 (2): 377–96.
- 1045 Mowat, Christopher G., Emma Rothery, Caroline S. Miles, Lisa McIver, Mary K. Doherty, Katy
1046   Drewette, Paul Taylor, Malcolm D. Walkinshaw, Stephen K. Chapman, and Graeme A.

1047 Reid. 2004. "Octaheme Tetrathionate Reductase Is a Respiratory Enzyme with Novel Heme
1048 Ligation." *Nature Structural & Molecular Biology* 11 (10): 1023–24.

1049 Müller, Albert Leopold, Júlia Rosa de Rezende, Casey R. J. Hubert, Kasper Urup Kjeldsen, Ilias
1050 Lagkouravdos, David Berry, Bo Barker Jørgensen, and Alexander Loy. 2014. "Endospores
1051 of Thermophilic Bacteria as Tracers of Microbial Dispersal by Ocean Currents." *The ISME*
1052 *Journal* 8 (6): 1153–65.

1053 Murat Eren, A., Özcan C. Esen, Christopher Quince, Joseph H. Vineis, Hilary G. Morrison,
1054 Mitchell L. Sogin, and Tom O. Delmont. 2015. "Anvi'o: An Advanced Analysis and
1055 Visualization Platform for 'omics Data." *PeerJ* 3 (October): e1319.

1056 Nemergut, Diana R., Elizabeth K. Costello, Micah Hamady, Catherine Lozupone, Lin Jiang,
1057 Steven K. Schmidt, Noah Fierer, et al. 2011. "Global Patterns in the Biogeography of
1058 Bacterial Taxa." *Environmental Microbiology* 13 (1): 135–44.

1059 Olm, Matthew R., Alexander Crits-Christoph, Spencer Diamond, Adi Lavy, Paula B. Matheus
1060 Carnevali, and Jillian F. Banfield. 2020. "Consistent Metagenome-Derived Metrics Verify
1061 and Delineate Bacterial Species Boundaries." *mSystems* 5 (1).
1062 <https://doi.org/10.1128/mSystems.00731-19>.

1063 Orphan, Victoria J., Christopher H. House, Kai-Uwe Hinrichs, Kevin D. McKeegan, and Edward
1064 F. DeLong. 2002. "Multiple Archaeal Groups Mediate Methane Oxidation in Anoxic Cold
1065 Seep Sediments." *Proceedings of the National Academy of Sciences of the United States*
1066 *of America* 99 (11): 7663–68.

1067 Paduan, Jennifer B., Robert A. Zierenberg, David A. Clague, Ronald M. Spelz, David W.
1068 Caress, Giancarlo Troni, Hans Thomas, et al. 2018. "Discovery of Hydrothermal Vent Fields
1069 on Alarcón Rise and in Southern Pescadero Basin, Gulf of California." *Geochemistry,*
1070 *Geophysics, Geosystems* 19 (12): 4788–4819.

1071 Parks, Donovan H., Maria Chuvoshina, Pierre-Alain Chaumeil, Christian Rinke, Aaron J.
1072 Mussig, and Philip Hugenholtz. 2020. "A Complete Domain-to-Species Taxonomy for
1073 Bacteria and Archaea." *Nature Biotechnology* 38 (9): 1079–86.

1074 Pattengale, Nicholas D., Masoud Alipour, Olaf R. P. Bininda-Emonds, Bernard M. E. Moret, and
1075 Alexandros Stamatakis. 2009. "How Many Bootstrap Replicates Are Necessary?" In
1076 *Research in Computational Molecular Biology*, 184–200. Springer Berlin Heidelberg.

1077 Pley, Ursula, Jutta Schipka, Agata Gambacorta, Holger W. Jannasch, Hans Fricke, Reinhard
1078 Rachel, and Karl O. Stetter. 1991. "Pyrodictium Abyssii Sp. Nov. Represents a Novel
1079 Heterotrophic Marine Archaeal Hyperthermophile Growing at 110°C." *Systematic and*
1080 *Applied Microbiology* 14 (3): 245–53.

1081 Price, Morgan N., Paramvir S. Dehal, and Adam P. Arkin. 2010. "FastTree 2--Approximately
1082 Maximum-Likelihood Trees for Large Alignments." *PloS One* 5 (3): e9490.

1083 Procesi, M., G. Ciotoli, A. Mazzini, and G. Etiope. 2019. "Sediment-Hosted Geothermal
1084 Systems: Review and First Global Mapping." *Earth-Science Reviews* 192 (May): 529–44.

1085 Rasko, David A., Garry S. A. Myers, and Jacques Ravel. 2005. "Visualization of Comparative
1086 Genomic Analyses by BLAST Score Ratio." *BMC Bioinformatics* 6 (January): 2.

1087 Resing, Joseph A., Peter N. Sedwick, Christopher R. German, William J. Jenkins, James W.
1088 Moffett, Bettina M. Sohst, and Alessandro Tagliabue. 2015. "Basin-Scale Transport of
1089 Hydrothermal Dissolved Metals across the South Pacific Ocean." *Nature* 523 (7559):
1090 200–203.

1091 Reveillaud, Julie, Emily Reddington, Jill McDermott, Christopher Algar, Julie L. Meyer, Sean
1092 Sylva, Jeffrey Seewald, Christopher R. German, and Julie A. Huber. 2016. "Subseafloor
1093 Microbial Communities in Hydrogen-Rich Vent Fluids from Hydrothermal Systems along the
1094 Mid-Cayman Rise." *Environmental Microbiology* 18 (6): 1970–87.

1095 Ruff, S. Emil, Jennifer F. Biddle, Andreas P. Teske, Katrin Knittel, Antje Boetius, and Alban
1096 Ramette. 2015. "Global Dispersion and Local Diversification of the Methane Seep
1097 Microbiome." *Proceedings of the National Academy of Sciences of the United States of*
1098 *America* 112 (13): 4015–20.

1099 Sauer, David B., and Da-Neng Wang. 2019. "Predicting the Optimal Growth Temperatures of
1100 Prokaryotes Using Only Genome Derived Features." *Bioinformatics* 35 (18): 3224–31.

1101 Schouten, Stefan, Stuart G. Wakeham, Ellen C. Hopmans, and Jaap S. Sinninghe Damsté.
1102 2003. "Biogeochemical Evidence That Thermophilic Archaea Mediate the Anaerobic
1103 Oxidation of Methane." *Applied and Environmental Microbiology* 69 (3): 1680–86.

1104 Seitz, Kiley W., Nina Dombrowski, Laura Eme, Anja Spang, Jonathan Lombard, Jessica R.
1105 Sieber, Andreas P. Teske, Thijs J. G. Ettema, and Brett J. Baker. 2019. "Asgard Archaea
1106 Capable of Anaerobic Hydrocarbon Cycling." *Nature Communications* 10 (1): 1822.

1107 Shaiber, Alon, Amy D. Willis, Tom O. Delmont, Simon Roux, Lin-Xing Chen, Abigail C. Schmid,
1108 Mahmoud Yousef, et al. 2020. "Functional and Genetic Markers of Niche Partitioning
1109 among Enigmatic Members of the Human Oral Microbiome." *Genome Biology* 21 (1): 292.

1110 Sillitoe, Ian, Natalie Dawson, Tony E. Lewis, Sayoni Das, Jonathan G. Lees, Paul Ashford,
1111 Adeyelu Tolulope, et al. 2019. "CATH: Expanding the Horizons of Structure-Based
1112 Functional Annotations for Genome Sequences." *Nucleic Acids Research* 47 (D1):
1113 D280–84.

1114 Simon, Jörg, Melanie Kern, Bianca Hermann, Oliver Einsle, and Julea N. Butt. 2011.
1115 "Physiological Function and Catalytic Versatility of Bacterial Multiheme Cytochromes c
1116 Involved in Nitrogen and Sulfur Cycling." *Biochemical Society Transactions* 39 (6):
1117 1864–70.

1118 Skennerton, Connor T., Mohamed F. Haroon, Ariane Briegel, Jian Shi, Grant J. Jensen, Gene

1119 W. Tyson, and Victoria J. Orphan. 2016. “Phylogenomic Analysis of Candidatus
1120 ‘Izimaplasma’ Species: Free-Living Representatives from a Tenericutes Clade Found in
1121 Methane Seeps.” *The ISME Journal* 10 (11): 2679–92.

1122 Speth, Daan R., Michiel H. In ’t Zandt, Simon Guerrero-Cruz, Bas E. Dutilh, and Mike S. M.
1123 Jetten. 2016. “Genome-Based Microbial Ecology of Anammox Granules in a Full-Scale
1124 Wastewater Treatment System.” *Nature Communications* 7 (March): 11172.

1125 Speth, Daan R., and Victoria J. Orphan. 2019. “ASM-Clust: Classifying Functionally Diverse
1126 Protein Families Using Alignment Score Matrices.” *Cold Spring Harbor Laboratory*.
1127 <https://doi.org/10.1101/792739>.

1128 Stamatakis, Alexandros. 2014. “RAxML Version 8: A Tool for Phylogenetic Analysis and
1129 Post-Analysis of Large Phylogenies.” *Bioinformatics* 30 (9): 1312–13.

1130 Stetter, K. O., R. Huber, E. Blöchl, M. Kurr, R. D. Eden, M. Fielder, H. Cash, and I. Vance. 1993.
1131 “Hyperthermophilic Archaea Are Thriving in Deep North Sea and Alaskan Oil Reservoirs.”
1132 *Nature* 365 (6448): 743–45.

1133 Takai, Ken, Kentaro Nakamura, Tomohiro Toki, Urumu Tsunogai, Masayuki Miyazaki, Junichi
1134 Miyazaki, Hisako Hirayama, Satoshi Nakagawa, Takuro Nunoura, and Koki Horikoshi. 2008.
1135 “Cell Proliferation at 122 Degrees C and Isotopically Heavy CH₄ Production by a
1136 Hyperthermophilic Methanogen under High-Pressure Cultivation.” *Proceedings of the*
1137 *National Academy of Sciences of the United States of America* 105 (31): 10949–54.

1138 Tatusov, Roman L., Natalie D. Fedorova, John D. Jackson, Aviva R. Jacobs, Boris Kiryutin,
1139 Eugene V. Koonin, Dmitri M. Krylov, et al. 2003. “The COG Database: An Updated Version
1140 Includes Eukaryotes.” *BMC Bioinformatics* 4 (September): 41.

1141 Teske, Andreas. 2020. “Guaymas Basin, a Hydrothermal Hydrocarbon Seep Ecosystem.” In
1142 *Marine Hydrocarbon Seeps: Microbiology and Biogeochemistry of a Global Marine Habitat*,

1143 edited by Andreas Teske and Verena Carvalho, 43–68. Cham: Springer International
1144 Publishing.

1145 Teske, Andreas, Dirk de Beer, Luke J. McKay, Margaret K. Tivey, Jennifer F. Biddle, Daniel
1146 Hoer, Karen G. Lloyd, et al. 2016. “The Guaymas Basin Hiking Guide to Hydrothermal
1147 Mounds, Chimneys, and Microbial Mats: Complex Seafloor Expressions of Subsurface
1148 Hydrothermal Circulation.” *Frontiers in Microbiology* 7 (February): 75.

1149 Teske, Andreas, Kai-Uwe Hinrichs, Virginia Edgcomb, Alvin de Vera Gomez, David Kysela,
1150 Sean P. Sylva, Mitchell L. Sogin, and Holger W. Jannasch. 2002. “Microbial Diversity of
1151 Hydrothermal Sediments in the Guaymas Basin: Evidence for Anaerobic Methanotrophic
1152 Communities.” *Applied and Environmental Microbiology* 68 (4): 1994–2007.

1153 Von Damm, K. L., J. M. Edmond, C. I. Measures, and B. Grant. 1985. “Chemistry of Submarine
1154 Hydrothermal Solutions at Guaymas Basin, Gulf of California.” *Geochimica et*
1155 *Cosmochimica Acta* 49 (11): 2221–37.

1156 Walters, William, Embriette R. Hyde, Donna Berg-Lyons, Gail Ackermann, Greg Humphrey,
1157 Alma Parada, Jack A. Gilbert, et al. 2016. “Improved Bacterial 16S rRNA Gene (V4 and
1158 V4-5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial
1159 Community Surveys.” *mSystems* 1 (1). <https://doi.org/10.1128/mSystems.00009-15>.

1160 Welhan, John A. 1988. “Origins of Methane in Hydrothermal Systems.” *Chemical Geology* 71
1161 (1): 183–98.

1162 Zheng, Rikuan, Rui Liu, Yeqi Shan, Ruining Cai, Ge Liu, and Chaomin Sun. 2021.
1163 “Characterization of the First Cultured Free-Living Representative of Candidatus
1164 Izemoplasma Uncovers Its Unique Biology.” *The ISME Journal*, March.
1165 <https://doi.org/10.1038/s41396-021-00961-7>.

1166 Zhou, J., M. A. Bruns, and J. M. Tiedje. 1996. “DNA Recovery from Soils of Diverse

1167   Composition.” *Applied and Environmental Microbiology* 62 (2): 316–22.

1168