Title: Community structure of phototrophic co-cultures from extreme environments

Charles Brooke[1], Morgan P. Connolly[2], Javier A. Garcia[3], Miranda Harmon-Smith[4], Nicole Shapiro[4], Erik Hawley[5], Michael Barton[4], Susannah G. Tringe[4], Tijana Glavina del Rio[4], David E. Culley[6], Richard Castenholz[7] and *Matthias Hess[1, 4]

[1]Systems Microbiology & Natural Products Laboratory, University of California, Davis, CA
[2]Microbiology Graduate Group, University of California, Davis, CA
[3]Biochemistry, Molecular, Cellular, and Developmental Biology Graduate Group, University of California, Davis, CA
[4]DOE Joint Genome Institute, Walnut Creek, CA
[5]Bayer, Pittsburg, PA
[6]LifeMine, Cambridge, MA
[7]University of Oregon, Eugene, OR

***Corresponding author:**

Matthias Hess
University of California, Davis
1 Shields Ave
Davis, CA 95616, USA
P (530) 530-752-8809
F (530) 752-0175
mhess@ucdavis.edu

ABSTRACT

Cyanobacteria are found in most illuminated environments and are key players in global carbon and nitrogen cycling. Although significant efforts have been made to advance our understanding of this important phylum, still little is known about how members of the cyanobacteria affect and respond to changes in complex biological systems. This lack of knowledge is in part due to the reliance on our ability to maintain pure cultures when determining the metabolism and function of a microorganism. To fill this knowledge-gap, we selected 26 photosynthetic co-cultures from the Culture Collection of Microorganisms from Extreme Environments (CCMEE) for 16S rRNA gene sequencing. We assessed if samples readily available from the CCMEE could contribute valuable insights to advance applied and fundamental science in the areas of global and local carbon and nitrogen cycling, without growing individual members of these co-cultures axenically. Results from this work will aid in determining whether culture depositories in general hold the potential to advance fundamental and applied research. Since maintaining culture depositories is resource intensive, such an assessment will be of great value in guiding future funding decisions.

INTRODUCTION

Cyanobacteria are photosynthetic prokaryotes that are found in the majority of illuminated habitats and are known to be some of the most morphologically diverse prokaryotes on our planet [1]. The global cyanobacterial biomass is estimated to total ~$3x10^{14}$ g of carbon [2] and cyanobacteria may account for 20–30% of Earth's primary photosynthetic productivity [3]. The efficient photosynthetic machinery of cyanobacteria has inspired

49 growing interest in the utilization of axenic cyanobacteria as well as cyanobacteria
50 containing co-cultures in microbial fuel cells [4,5]. In addition to having a global effect on the
51 carbon cycle, cyanobacteria-mediated nitrogen fixation has been estimated to supply 20–
52 50% of the nitrogen input in some marine environments [6]. A detailed comprehension of
53 cyanobacteria and their contribution to global carbon and nitrogen cycling is therefore
54 necessary for a multi-scalar understanding of these globally important nutrient cycles and
55 ultimately for our ability to build accurate models to predict future climate patterns.

56

57 Besides their ecological relevance, cyanobacteria have potential applications in
58 biotechnology. The photosynthetic metabolism of cyanobacteria facilitates the assimilation
59 of carbon dioxide, a cheap and abundant substrate, to synthesize a variety of value-added
60 compounds with industrial relevance [7]. Numerous strains of cyanobacteria have been
61 investigated for their potential to produce bioactive compounds, biofertilizer, biofuels, and
62 bioplastics [8]; and interactions of cyanobacterial strains with other bacteria have been
63 found to improve desirable cyanobacterial phenotypes [9]. Genes encoding enzymes
64 capable of catalyzing reactions that result in unique products, such as modified trichamide,
65 a cyclic peptide suggested to protect the bloom-forming *Trichodesmium erythraeum*
66 against predation [10], and prochlorosins, a family of lanthipeptides with diverse functions
67 that are synthesized by various strains of *Prochlorococcus* and *Synechococcus* [11,12], have
68 been identified from cyanobacterial genomes [13,14]. It is very likely that *de novo* genome
69 assembly from metagenomic data will facilitate the discovery of novel enzymes from
70 cyanobacteria for which we currently lack the appropriate isolation and cultivation
71 techniques. Although metagenome-derived genomes hold great potential to enhance our
72 knowledge about genomic dark matter, ultimately, improved techniques to isolate and
73 enable axenic culturing of microorganisms that are currently characterized as

74 "unculturable", as well as new genetic tools to work with non-axenic cultures will be
75 necessary in order to fully access the biotechnological potential of cyanobacteria.

76

77 Culture collections provide the possibility of preserving microbial isolates over extended
78 periods of time without introducing significant genetic changes [15] and they facilitate open
79 access to these isolates and their associated metadata [16]. Hence co-culture repositories
80 represent a promising starting point for developing and testing techniques to study and
81 manipulate uncultivated microbes. Although culture collections hold enormous potential
82 for capturing and preserving microbial biodiversity and for improving cultivation
83 techniques, there are numerous challenges in maintaining these biological depositories
84 and the individual samples they contain. A detailed understanding of the make-up of
85 individual co-cultures is essential to assess their true value and ultimately to develop
86 strategies that will be suitable to address challenges associated with long-term sample
87 maintenance. With recent advances in DNA sequencing technologies and the accessibility
88 of 16S rRNA gene-based microbial community profiling, we are now well positioned to re-
89 inventory and evaluate existing culture collections. Standardized inventories will facilitate
90 sample documentation for deposits maintained by individual laboratories and large culture
91 collections alike, which will be essential for cataloguing, preserving and surveying the
92 planet's microbial biodiversity.

93

94 To explore the potential of culture collections, specifically those that maintain samples of
95 microbial co-cultures, to provide reference genomes from environmentally or industrially
96 relevant microorganisms, we reexamined the biodiversity of 26 historical phototrophic
97 samples from the Culture Collection of Microorganisms from Extreme Environments

(CCMEE). While some of the samples, and their dominant phototrophs were studied previously using 16S rRNA profiling and morphological characterization [17-24] the diversity of the photosynthetic and non-photosynthetic organisms and the overall community assemblage of these co-cultures have not yet been characterized. To add further value to this study, we selected samples that originated from diverse extreme environments across the globe; with properties suggesting each co-culture would yield a unique microbial consortium. An enhanced understanding of the microbial diversity that is preserved within environmental co-cultures available through public culture collections will contribute to a better understanding of global microbial biodiversity.

Materials and Methods

Sample collection & sample description

Co-cultures selected for this study are part of a larger culture collection and were collected from different locations (Table 1) between 1988 and 2002. Isolates were collected using sterile techniques, kept in the dark and stored on ice as quickly as possible. Samples were transported to the laboratory where aliquots were prepared for cultivation and preservation at -80°C. For this study, co-cultures were selected from the CCMEE to cover a variety of geographical locations (Figure 1) as well as a range of different ecosystems (Table 1). Due to the lack of a consistent terminology historically used to describe the sampling sites, we categorized co-cultures according to the geographical location (e.g. Antarctica, Bermuda, Denmark, Mexico and Spain) and based on a general description of the habitat (i.e. creek, crust, freshwater, hot spring, marine, saline pond, terrestrial, travertine, and tree bark) from where the co-cultures

were collected. In addition, we used the growth medium and temperature (i.e. 12°C, 23°C, 40°C, 45°C, 55°C) at which available co-cultures have been maintained historically in the laboratory to categorize the co-cultures used in this study.

FECB1 (CCMEE ID 5011) and FECB3 (CCMEE ID 5034) were collected from saline and brackish melt ponds in Antarctica respectively and were dominated by phototrophic cyanobacteria classified as *Oscillatoria* sp [18]. FECB2 (CCMEE ID 5019) was collected from a small freshwater pond (Pinnacle Pond) in the Ice Pinnacle area near Bratina Island, Antarctica, and was phylogenetically uncharacterized prior to our efforts. FECB4 (CCMEE ID 5047; AP1) and FECB5 (CCMEE ID 5049; AO21) were isolated from Lake Arcas, Spain and the dominant photosynthetic organisms within these samples were classified by 16S rRNA sequence analysis as being related to *Pseudanabaena limnetica* and *Oscillatoria* cf. *tenuis,* respectively [24]. FECB6 (CCMEE ID 5051), FECB14 (CCMEE ID 5093; WT-97 Cal), FECB15 (CCMEE ID 5083), and FECB19 (CCMEE ID 5091; Y-97) were collected from diverse hot springs and a warm hot spring outflow (warm creek) within Yellowstone National Park (YNP) (Table 1). FECB10 (CCMEE ID 5056; M88-VD (1)) was collected as epiliths from the Viscaino Desert in Mexico [23]. FECB17 (CCMEE ID 5085; RC-97 Cal) and FECB36 (CCMEE ID 6076) were isolated from Rabbit Creek and a crust in the Sentinel Spring Meadows in YNP respectively and dominant phototrophs of these co-cultures were characterized previously as *Calothrix* spp. [22]. FECB22 (CCMEE ID 5097; HW-91) and FECB26 (CCMEE ID 5099; B77-scy,j,) were collected from a tree trunk in Hawaii and a wooded fence in Bermuda respectively. FECB24 (CCMEE ID 5098; AN-90) was obtained from a shallow melt pond (~10 m^2) in the Victoria Valley, Antarctica, whereas FECB28 (CCMEE ID 5102) was collected from a saline melt pond on Bratina Island, Antarctica [19]. FECB32 (CCMEE ID 6031), FECB34 (CCMEE ID 6069) and FECB38 (CCMEE ID 6083) were endoliths collected from

147 subsurface (1-5 mm depths) travertine deposits in YNP [20]. FECB53 (CCMEE ID 5610)
148 was collected from Sylvan Springs in YNP. Temperature and pH at the sampling site of
149 FECB53 were determined to be 40°C and pH4. The dominant phototropic strain in
150 FECB53 was identified previously as the thermo-acidophilic *Cyanidioschyzon* [21].
151 FECB58 (CCMEE ID 5216; OH-9-45C) and FECB68 (CCMEE ID 5240; OH-2-55C) were
152 collected from Hunter's Hot Spring in Oregon and the phototroph dominating these
153 samples was determined to be a thermophilic member belonging to the genus
154 *Synechococcus* [17].

155

156 *Growth of co-cultures*

157 To obtain sufficient biomass for subsequent DNA analysis, 100 µL of each co-culture
158 were transferred to 25 mL of sterile BG11 media [25]. For FECB52 and FECB53 BG11
159 was substituted by Cyanidium medium [26]. Co-cultures were subjected to a 12 hr diurnal
160 light/dark cycle while grown at the temperatures indicated in Table 1.
161

162 *DNA extraction and 16S rRNA gene amplification*

163 Total microbial DNA was extracted from 500 µL of each photosynthetic co-culture using
164 the FastDNA SPIN Kit for Soil (MP Biomedical, Solon, OH) according to the
165 manufacturer's instructions. Extracted DNA was quantified via fluorescence (Qubit;
166 Thermo Scientific, USA) and the hypervariable V4 region of the 16S rRNA gene was
167 amplified from extracted DNA using the primer set 515F/805R (515F: 5′-
168 GTGCCAGCMGCCGCGGTAA-3′ and 805R: 5′-GGACTACHVGGGTWTCTAAT-3′). The
169 forward primer included an 11 bp barcode to allow multiplexing of samples during
170 sequencing. The barcode sequence for each sample is listed in Supplemental Table S1.

Subsequent PCR reactions were performed using the 5PRIME HotMasterMix amplification mix (QIAGEN, Beverly, MA) with the following PCR conditions: initial denaturation for 90 sec at 94°C, followed by 30 amplification cycles (45 sec at 94°C, 60 sec at 60°C, and 90 sec at 72°C) followed by a final extension step of 72°C for 10 min. Amplification products were cooled to 4°C. Samples were sequenced at the Department of Energy's Joint Genome Institute (JGI; http://www.jgi.doe.gov) according to JGI's standard operating procedure using Illumina's MiSeq platform and v3 chemistry.

Sequence data analysis

Raw sequencing data were downloaded from the JGI's Genome Portal (http://genome.jgi.doe.gov/) under the project ID 1032475. Data were decompressed and de-interleaved using the 7-zip software (www.7-zip.org) and an in-house script, respectively. De-interleaved files were subsequently processed using MOTHUR version 1.38.1 [27,28]. Paired-end reads were combined using the *make.contigs* command. Sequences with ambiguous base calls and sequences longer than 325 bp were removed using *screen.seqs*. Duplicate sequences were merged using *unique.seqs*, and the resulting unique sequences were aligned to the V4 region of the SILVA database (v123) [29]. Chimeras were removed using UCHIME [30] and quality filtered sequences were taxonomically classified at 80% confidence to the GreenGenes reference taxonomy (release gg_13_5_99) [31]. Non-prokaryotic sequences were removed and the *dist.seqs* command was used to calculate pairwise distances between the aligned sequences. The resulting pairwise distance matrix was used to cluster sequences into operational taxonomic units (OTUs) with a 97% sequence identity cut-off using UCLUST [32]. The most abundant sequence of each OTU was picked as the representative sequence. OTUs were taxonomically classified using the *classify.otu* command using the

196 GreenGenes reference taxonomy (release gg_13_5_99). Shannon, Simpson, and
197 Chao1 estimators were calculated in MOTHUR [27].

198

199 In order to visualize the overall compositional differences between the co-cultures, an
200 uncorrected pairwise distance matrix was generated using the *dist.seqs* command in
201 MOTHUR and a tree was generated using *Clearcut* (version 1.0.9*)* [33]. A cladogram from
202 the resulting tree file was constructed and visualized using iTOL (https://itol.embl.de;
203 accessed on October 16th, 2016; [34]). Cluster designations were assigned at a branch
204 length of 0.05. Samples whose branches split at a distance >0.05 were considered as
205 part of the same cluster (Figure 2).

206

207 **Availability of data and material**
208 Co-cultures subject to this study are publicly available through the CCMEE and the
209 UTEX Culture Collection of Algae at the University of Texas at Austin upon request
210 using the corresponding FECB ID (Table 1). Co-cultures can also be obtained from the
211 Hess Lab at UC Davis. Sequences generated during this project have been deposited
212 and are publicly available at NCBI's SRA under the BioProject ID PRJNA401502. All
213 other data is included in this published article and its supplementary information files.

214

215 **Results & Discussion**

216

217 A total of 3,357,905 raw reads (mean (SD) = 129,150 (±15,845) reads per sample) were
218 generated from the V4 region of the 16S rRNA gene (Table 2). Quality filtering removed
219 ~3.8% (±0.57%) of the raw reads from each sample due to insufficient quality. The

220 remaining reads were assigned to a total of 5,785 distinct Operational Taxonomic Units
221 (OTUs) based on 97% sequence identity (Table S2).

222

223 To estimate the microbial diversity within each sample, rarefaction analyses were
224 performed (Supplemental Figure S1) and diversity indices were calculated (Table 2). The
225 inverse Simpson index of the samples ranged between 1.52 and 9.24 with the lowest and
226 highest indices calculated for FECB3 and FECB32 respectively (Table 2). Not surprisingly,
227 the diversity in the co-cultures under investigation appeared to be negatively correlated
228 with the proportion of reads recruited by the dominant OTU of each sample (Pearson r =-
229 0.8806; p <0.01). Although samples ranked slightly differently based on their diversity
230 when Chao1 or Shannon indices were calculated, the overall trend remained the same
231 (Table 2).

232

233

234 *The McMurdo Dry Valley Lake System, a physically highly stable lacustrine system*

235 The McMurdo Dry Valley (MDV) is one of the most extreme deserts on Earth, and although
236 the importance of the microbial communities for the biogeochemical cycles of this region
237 is widely accepted, the microbial ecology of the MDV remains poorly understood [35].
238 FECB3, originating from a brackish pond on Bratina Island, was dominated by
239 OTU000003, which recruited 80.3% of all reads (Supplemental Table S2). OTU000003
240 was classified as the cyanobacterium *Phormidium pseudopriestleyi*, previously reported
241 to dominate microbial mats of the anoxic zone of Lake Fryxell, Antarctica [36]. The second
242 and third most abundant OTUs in FECB3 were OTU000015 and OTU000061 respectively
243 (Supplemental Table S2). Both OTU000015 and OTU000061 were classified as

Rhodobacteriaceae and recruited 9.2% and 8.2% of the reads generated for FECB3. Whereas a taxonomic classification of OTU000015 was not possible at a resolution higher than the family level, OTU000061 was classified as *Paracoccus marcusii,* a Gram negative organism that displays a bright orange color due to the synthesis of carotenoids such as astaxanthin [37].

While the microbial ecology of melt ponds and lakes in the MDV, habitats covered year-round with an ice sheet, have been studied in great detail; most of the insights regarding the microbial community assemblage in these waters are based primarily on microscopy [36]. Molecular data, like those presented here and those that could theoretically be generated from other MDV samples that are readily available from the CCMEE and other culture collections, will be of great value to extend our knowledge framework of the microbial ecology of this unique ecosystem.

Omnipresence of Cyanobacteria and Proteobacteria within photosynthetic co-cultures

While the microbial communities of the co-cultures under investigation varied greatly, cyanobacteria and proteobacteria co-occurred in all 26 of the community assemblages. Community composition analysis revealed that each of the co-cultures contained at least one OTU (mean (SD) = 2 ($\pm$1.23)) that recruited >0.1% of the co-culture specific reads and that was classified as *Cyanobacteria* (Table 3). The only other phylum present in each of the individual 26 co-cultures and represented by at least one OTU recruiting >0.1% of the reads was the *Proteobacteria* phylum (Table 3). In contrast, only three samples, namely FECB5, FECB30 and FECB68, contained OTUs that recruited >0.1% of the sample specific reads and that could not be classified at the phylum level or at a higher taxonomic resolution (Table 3). It is possible that the relatively high abundance of

non-classified phyla might contribute to the separation of these samples into distinct clusters (i.e. cluster XII, IX, and IV) (Figure 2). In addition to their omnipresence, *Cyanobacteria* and *Proteobacteria* also recruited the majority of the reads in all but four (i.e. FECB2, FECB12, FECB58, and FECB68) of the samples under investigation (Figure 3 and Supplemental Table S3). In FECB2 and FECB12 the majority of the reads were recruited by OTUs classified as members of the phylum *Bacteroidetes* (recruiting 50.6% and 72% of the reads respectively), whereas within FECB58 and FECB68, *Armatimonadetes* (38.3%) and *Chloroflexi* (25.9%) were identified as the most abundant phyla (Figure 3 and Supplemental Table S3). The fact that these samples were dominated by phyla other than the *Cyanobacteria* or *Proteobacteria* may also help to explain why these samples (Figure 2) form distinct clusters (cluster I, XIV and V, IV respectively).

Firmicutes dominate photosynthetic co-cultures from hot springs

Firmicutes abundances calculated for co-cultures from hot spring samples were higher compared to those calculated for co-cultures from other environments studied during this project. OTUs assigned to the *Firmicutes* phylum were detected above the applied cut-off level of 0.1% in only four of the twenty-six co-cultures under investigation (Table 3). Interestingly, these samples (i.e. FECB34, FECB52, FECB58 and FECB68) are co-cultures collected from hot springs or deposits within hot springs, with FECB52, FECB58 and FECB68 being maintained in culture at temperatures >40°C. OTU000073 (classified as *Alicyclobacillus tolerans*), OTU00082 (classified as members of the genus *Paenibacillus*), OTU000154 (classified as *Geobacillus vulcani*), and OTU000158 (classified as a member of the *Bacillaceae* family) recruited 5.9%, 3.4% , 0.5% and 0.4%

294 of the reads generated from FECB52, FECB34, FECB68 and FECB58 respectively
295 (Supplemental Table S2). *Alicyclobacillus tolerans* and *Geobacillus vulcani* have been
296 described previously as aerobic spore-forming thermophiles and have been isolated
297 from lead–zinc ores [38] and hot springs [39] in Russia, respectively. Members of the genus
298 *Paenibacillus* have been isolated from a wide variety of environments and some
299 *Paenibacillus* species have been found to promote crop growth directly via biological
300 nitrogen fixation, phosphate solubilization, production of the phytohormone indole-3-
301 acetic acid and they have been identified as a potential source of novel antimicrobial
302 agents [40]. Although it is difficult to make a reliable prediction of the metabolic capacities
303 of the organism associated with OTU000082 solely based on 16S rRNA data, it is
304 certainly possible that this organism might possess the ability to promote or inhibit plant
305 and microbial growth respectively.

306

307 *Photosynthetic co-cultures from Antarctica and YNP to study adaptation to increased*
308 *radiation, low temperatures and oligotrophic growth conditions*

309 Microbial adaptation to extreme environments and the molecular framework that allows
310 microorganisms to survive and thrive in the presence of increased rates of radiation, low
311 temperatures and in the absence of nutrients has fascinated the scientific community for
312 decades and remains poorly understood. In an attempt to provide a better basis of the
313 taxonomic make-up of co-cultures that were collected from ecosystems that are
314 characterized by these extremes we included co-cultures from Antarctica and YNP in
315 this study (Table 1). OTU-based comparison of Antarctica and YNP co-cultures revealed
316 between 197 (FECB2) and 549 (FECB6) distinct OTUs (mean (SD) = 342 ($\pm$87.2)
317 OTUs), based on 97% sequence similarity (Table 2). The number of OTUs that recruited
318 >0.1% of all reads ranged from 3 to 29 OTUs, with FECB2 and FECB32 having the

lowest and highest OTU count respectively (Table 2). FECB2 was dominated by an OTU classified as *Hymenobacter,* which recruited all *Bacteroidetes*-specific reads generated from this sample (Tables 3 & 4). The genus *Hymenobacter* contains several pigmented bacteria that have been isolated from Antarctica and have been reported to possess increased resistance to radiation [41,42], which might explain their increased abundance in FECB2, a co-culture isolated from an environment known to possess increased levels of UV radiation. Taking this into consideration, FECB2 and its individual community members could be a potential target for future studies to enhance our understanding of processes that infer resistance to radiation and DNA damage. The second most abundant OTU in FECB2, recruiting 48% of the generated samples, was classified as *Phormidium* sp (Supplemental Table S2), a cyanobacterial genus that has been reported to dominate aquatic microbial mats from Antarctica [43,44]. Representative isolates from this genus have been proposed previously as cost-effective options for industrial carotenoid production [45], suggesting that FECB2 may hold the potential for industrial carotenoid production.

FECB32 is a mixed culture isolated from an ancient travertine at Mammoth in YNP. Our analysis indicated that FECB32 contained 29 OTUs that each accounted for >0.1% of all the generated reads (Table 2). Fifteen of these OTUs recruited >1% of all reads and 4 OTUs collectively accounted for ~60% of the reads generated from this co-culture (Supplemental Table S4). These 4 OTUs were classified as *Sphingopyxis alaskensis, Chelativorans sp.* and as members of the *Chitinophagaceae* and *Comamonadaceae families,* recruiting ~19%, 13%, 17%, and 11% of the reads respectively (Supplemental Tables S2 & S4). *Sphingopyxis alaskensis* is a Gram negative bacterium found in relatively high abundance in oligotrophic regions of the ocean [46,47] and it has been

344 studied in great detail as a model marine bacterium, specifically to understand microbial
345 adaptation to cold or oligotrophic environments [48,49]. The *Chitinophagaceae* family
346 contains a wide phylogenetic diversity with many of its members being mesophilic.
347 However, *Chitinophagaceae* have been reported to grow optimally at temperatures of
348 55°C and higher [50,51].

349

350 *Photosynthetic co-cultures containing the deep-branching candidate phylum*
351 *Melainabacteria*

352 Extreme environments, similar to those on early Earth, are often proposed to hold critical
353 information about the historical progression of life on our planet and one niche that
354 encompasses those physical stresses is the endolithic environment of rocks [20].
355 Phylogenetic analysis of the heterotrophic population associated with FECB32, which
356 was isolated from travertine deposited by hot springs in YNP, found that sequences from
357 MLE-12 (OTU000109) recruited ~2% of the sample specific sequences (Supplemental
358 Table S2). This rendered MLE-12, previously assigned to the deep-branching candidate
359 phylum *Melainabacteria* [52], as the eleventh most abundant organism in this
360 photosynthetic co-culture. It has been proposed previously that *Melainabacteria*, which is
361 commonly found in aquatic habitats, separated from the cyanobacteria before the latter
362 acquired photosynthetic capabilities [52]. Hence FECB32 might be a particularly valuable
363 co-culture to generate new insights into the evolution of and relationship between the
364 phylogenetically closely related *Cyanobacteria* and *Melainabacteria*. In addition, this
365 sample might provide the opportunity to enhance our understanding of the origin of
366 oxygenic photosynthesis and aerobic respiration in *Cyanobacteria*, an area that is
367 currently still poorly understood [53].

368

369 Interestingly, OTU000109 was also detected in FECB36 and FECB38 (Supplemental
370 Table S2), although at significantly lower abundance (<0.001%). FECB36 and FECB38
371 were similar to FECB32 in that they were isolated from sites in YNP. Interestingly,
372 FECB32 and FECB38 cluster together (cluster IX) suggesting a similar overall microbial
373 community profiles, but separately from FECB36 (Figure 2). The only additional samples
374 that contained OTUs classified as *Melainabacteria,* recruiting >0.1% of the generated
375 reads, were FECB58 and FECB68 with ~0.9% and ~0.2% of their reads to this deeply
376 branched phylum, respectively (Supplemental Table S2). It seems noteworthy that
377 FECB58 and FECB68 were also isolated from hot springs and clustered closely together
378 based on their overall microbiome composition (Clusters V and IV respectively; Figure
379 2).

380

381 *Photosynthetic co-cultures from Hunter's Hot Spring, Oregon*

382 Hunters Hot Spring continues to be a source of interesting microbial activity. FECB58
383 and FECB68 were both isolated from Hunters Hot Spring in Oregon, USA and they
384 shared similar microbial community members. Despite their similar community profile,
385 abundances of the dominant OTUs associated with these two hot spring co-cultures
386 were remarkably different. FECB58 was dominated by 3 OTUs (OTU000014,
387 OTU000024, and OTU000033). OTU000014 was classified as OS-L, an uncultured
388 representative of the phylum *Armatimonadetes*, OTU000024 which was classified as
389 belonging to the *Bacteroidetes* phylum, and OTU000033 which was classified as
390 *Thermosynechococcus*. These OTUs contributed 38%, 29% and 20% of the reads
391 generated from FECB58 respectively. Whereas OTU000014 recruited ~4.9% of all reads

392 generated from FECB68, representing the sixth most abundant OTU in the FECB68
393 community, OTU000024 and OTU000033 were only present at an abundance <0.0001%
394 in FECB68 (Supplemental Table S2).

395

396 FECB68 was dominated by 6 OTUs (i.e. OTU000028, OTU000030, OTU000036,
397 OTU000049, OTU000065, and OTU000014) recruiting ~25.7%, 23.1%, 20.4%, 14.3%,
398 7.6%, and 4.9% of the reads respectively. OTU000028 was classified as belonging to
399 the genus *Chloroflexus*, whereas OTU000030 and OTU000036 were classified as
400 representative of the genus *Meiothermus* and *Gloeobacter*, respectively. *Chloroflexus* is
401 an anoxygenic phototrophic bacterium that grows at temperatures up to 70°C [54] and
402 forms yellow-orange-greenish mats in association with cyanobacteria [55]. Members of the
403 cyanobacterial genus *Gloeobacter* lack thylakoids, and have been proposed to host the
404 earliest ancestors, or a missing link, in the cyanobacteria lineage [56]. Thus, FECB68
405 offers a unique opportunity to investigate interspecies interaction between a member of
406 these basal cyanobacteria and the thermophilic phototroph *Chloroflexus*, represented by
407 OTU000028 in this co-culture. As outlined in a recent review [54], Hunter's Hot Spring
408 located in Oregon is one of the most studied hot springs in the world, and has a large
409 repertoire of work conducted over the last 40 years [54]. However, most of this work was
410 performed prior to the advent of recent molecular techniques. Hence, the sequencing
411 data generated from FECB58 and FECB68 during this study will complement previous
412 work performed using traditional microbiology techniques and facilitate new insights into
413 the microbiology of this unique ecosystem.

414

Photosynthetic co-cultures from lignocellulosic surfaces with potential to fix nitrogen and degrade aromatic compounds.

FECB22 and FECB26 are mesophilic co-cultures collected from similar habitats (i.e. from tree bark and a wooden fence) from two locations (i.e. Hawaii and Bermuda) approximately 9,000 kilometers apart from each other (Figure 1 & Table 1). Diversity index calculation placed these two samples in the mid-range of the diversity spectrum of the 26 co-cultures analyzed for this study. The inverse Simpson, Chao1, and Shannon index was calculated at 4.46, 937.11 and 2.08 for FECB22 and 2.29, 722.94, and 1.32 for FECB26, respectively (Table 2). Within FECB22, 23 OTUs were identified as individually recruiting more than 0.1% of the generated reads. In contrast, FECB26 contained only 16 OTUs that recruited more than 0.1% of the reads each (Table S2). FECB22, scraped from tree bark in Hawaii, was dominated by 11 OTUs, each recruiting >1% of the reads. The most abundant OTU (OTU000017) was classified as a member of the *Mycoplana*, a genus that contains bacteria capable of aromatic compound degradation [57], and it recruited 40.2% of the reads. OTU000042 (classified as *Rhizobium leguminosarum*), OTU000045 (classified as *Acetobacteraceae*), and OTU000072 (classified as *Cyanobacteria*), were the next most abundant OTUs, recruiting 17.1%, 16.3%, and 5.5% of the reads generated from FECB22 respectively. *Rhizobium leguminosarum* is a well-studied α-proteobacterium capable of N_2-fixation and "rhizobia" have been suggested repeatedly to facilitate more sustainable agricultural practices through their symbiosis with legumes, reducing the need for nitrogen fertilizer [58]. It remains to be seen if OTU000042 provides N_2 to the other organisms in this co-culture or if it consumes all of the fixed N_2 itself. *Acetobacteraceae* are α-proteobacteria often associated with low pH environments and are known for their ability to efficiently synthesize biological cellulose [59,60]. Furthermore, *Acetobacteraceae* have been reported

before as some of the dominant players in photosynthetic consortia during soil formation [61]. It would be interesting to explore the agricultural and chemical potential of a minimalistic co-culture composed of the 4 OTUs (i.e. OTU000017, OTU000042, OTU000045 and OTU000072) that dominated FECB22, as they may combine the ability to degrade aromatic compounds and synthesize cellulose while removing nitrogen from the atmosphere. FECB26, on the other hand, was dominated by OTU000010, which recruited 63.2% of the reads generated and classified as an unclassified member of the *Nostocales*; a phylogenetic group known for their functional and morphological diversity. Members of the *Sphingomonadaceae* (i.e. OTU000041 and OTU000062), phototropic α-proteobacteria often found in high abundance in environments previously thought to support mostly the growth of cyanobacteria [62], contributed to a total of 25.6% of the generated reads. Most interestingly, OTU000017 was also identified within FECB26 recruiting ~1.6% of the reads. It is possible that OTU000017 facilitates a metabolic reaction, in which aromatic compounds typically associated with the decomposition of woody material under aerobic conditions are utilized. Further characterization of this organism in co-culture and eventually in axenic culture might provide further clarity if this is the case.

Conclusion

Culture collections can provide easy access to biological samples without the need for extensive resources by the requesting individual, subsequently facilitating new studies and ultimately advancing our understanding and appreciation of phylogenetic and functional biodiversity. The 16S rRNA based community fingerprints of the 26

464 photosynthetic co-cultures described here provide us with a first glimpse into the
465 taxonomic and functional diversity of communities from extreme environments that were
466 considered for a long time as too harsh to support the growth of complex microbial
467 communities. The extreme conditions that are associated with the habitats from where
468 these co-cultures were collected offer the unique opportunity to study the molecular
469 mechanisms that support the growth of these extremophilic co-cultures and their role in
470 global carbon and nitrogen cycling. Furthermore, an in-depth understanding of these
471 extreme co-cultures holds the potential to discover novel microbial proteins that might
472 render current agricultural, industrial and medical processes more economical and
473 sustainable. The relatively low diversity and complexity of these co-cultures make them
474 ideal subjects to investigate symbiotic relationships. By determining the chemical and
475 physical requirements of individual community members in the low complexity co-cultures
476 described here, the goal of enabling their axenic growth is promising. Advancing our ability
477 to infer the metabolic requirements of individual hitherto uncultivable microorganisms is of
478 great importance as has been shown by a large body of work including the “Genomic
479 Encyclopedia of Bacteria and Archaea” (GEBA) and CyanoGEBA projects, multi-
480 investigator projects spearheaded by the Kerfeld and Eisen group in collaboration with
481 Drs. Krypides and Woyke from the DOE’s Joint Genome Institute [63-65]. However,
482 bioinformatics and wet-lab tools to dissect complex microbial communities and processes
483 into their individual components are still in their infancy and obtaining pure isolates from
484 complex microbial communities still represents a major challenge. The photosynthetic co-
485 cultures utilized in this work represent excellent model systems for tool development and
486 verification due to their relatively low community complexity and their public availability via
487 culture collections such as the CCMEE and UTEX.

488

489 This study highlights a major challenge (i.e. standardization of protocols) associated with
490 environmental samples and sample data obtained during independent sampling efforts.
491 Fortunately, with recent advances in data technologies, the task of data acquisition and
492 dissemination has become less of a challenge. Recording standardized geographical and
493 environmental data, such as latitude, longitude, elevation, and temperature, can now be
494 performed with relatively high accuracy on a range of electronic devices - from most cell
495 phones, to inexpensive handheld devices developed specifically for accurate data
496 acquisition under field conditions. In order to make the best use of these technologies and
497 of biological samples that will be collected, defining a set of minimal information
498 parameters to be recorded during the collection of an environmental sample is of great
499 importance. Similar efforts have been successfully implemented by the Genomic
500 Standards Consortium (GSC) for microbial genomes and metagenomes in the form of the
501 “minimum information about a genome sequence” (MIGS) [66] and are enforced when
502 describing a novel microbial species [67]. It is understandable that characteristics suitable
503 as “minimal information” for environmental samples might differ from those that have been
504 established for axenic isolates and their genomes. However, as long as there is
505 consistency, the reported data will provide a valuable starting point for future efforts to
506 retroactively study archived samples.

507

508 16S rRNA gene profiling has become a commodity and easily generated phylogenetic
509 fingerprints provide a useful starting point to classify and categorize environmental
510 samples of microbial co-cultures. Due to the wide availability and continuing decline in
511 cost, this technique provides an ideal approach to re-examine the phylogenetic makeup
512 of legacy samples before they disappear due to continuous decline in funding for
513 maintaining small and non-centralized culture collections. In combination with biochemical

514 measurements, such as carbon and nitrogen utilization capabilities, the scientific
515 community would have immediate access to the phylogenetic and functional diversity
516 available through the existing culture collections. The identification of **M**inimum
517 **I**nformation about a **C**o-**C**ulture **S**ample (MICCS) would be a significant step in
518 standardizing sample acquisition and maintenance, increasing the value of current and
519 future microbial samples collected from the environment. Developing MICCS and applying
520 them to co-cultures currently available from existing culture depositories is beyond the
521 scope of the work presented here, but we hope that the results presented here will
522 contribute to the initiation of this process and stimulate broad involvement and support
523 from the scientific community and various funding agencies.

524

525 Another noteworthy aspect of samples readily available through existing culture
526 collections, including the consortia discussed in this work, is their educational value. More
527 specifically, samples that can be acquired and maintained without the need of significant
528 resources and for which basic phylogenetic and functional information is available. These
529 co-cultures provide a unique opportunity for exciting undergraduate research, in
530 combining microbial diversity, microbial ecology and biotechnology. Techniques for basic
531 biochemical and physiological characterizations of these samples could be learned and
532 conducted by dedicated undergraduate students within a few weeks. A research program
533 based on these co-cultures would provide students with the unique opportunity to develop
534 laboratory skills and to learn firsthand about biogeochemical processes that shape our
535 environment and climate. Additional publicly available *omics* data, such as metagenomics
536 and metatranscriptomics generated from individual samples, would extend the scope of
537 these undergraduate research programs, in providing students the opportunity to learn

538 various *omics* analysis techniques using web-based tools or standalone scripts,
539 depending on the educational level and interest of each student.

540

541 In summary, culture collections that provide access to and standardized information
542 about microorganisms and microbial consortia provide opportunities for educational and
543 scientific progress. Therefore, it is of high importance that culture collections continue to
544 obtain the financial support necessary to provide this invaluable service to our society

545

546 **References**

547 1 Whitton, B. A. & Potts, M. *The ecology of cyanobacteria: their diversity in time*
548 *and space*., 3-4 (Kluwer Academic, 2000).

549 2 Garcia-Pichel, F., Belnap, J., Neuer, S. & Schanz, F. Estimates of global
550 cyanobacterial biomass and its distribution. *Algological Studies* **109**, 213-227,
551 doi:10.1127/1864-1318/2003/0109-0213 (2003).

552 3 Pisciotta, J. M., Zou, Y. & Baskakov, I. V. Light-dependent electrogenic activity of
553 cyanobacteria. *PLoS One* **5**, e10821, doi:10.1371/journal.pone.0010821 (2010).

554 4 Gajda, I., Greenman, J., Melhuish, C. & Ieropoulos, I. Self-sustainable electricity
555 production from algae grown in a microbial fuel cell system. *Biomass &*
556 *Bioenergy* **82**, 87-93, doi:10.1016/j.biombioe.2015.05.017 (2015).

557 5 Zhao, J., Li, X. F., Ren, Y. P., Wang, X. H. & Jian, C. Electricity generation from
558 Taihu Lake cyanobacteria by sediment microbial fuel cells. *Journal of Chemical*
559 *Technology and Biotechnology* **87**, 1567-1573, doi:10.1002/jctb.3794 (2012).

560 6 Karl, D. *et al.* The role of nitrogen fixation in biogeochemical cycling in the
561 subtropical North Pacific Ocean. *Nature* **388**, 533-538 (1997).

562 7 Al-Haj, L., Lui, Y. T., Abed, R. M., Gomaa, M. A. & Purton, S. Cyanobacteria as
563 Chassis for Industrial Biotechnology: Progress and Prospects. *Life (Basel)* **6**,
564 doi:10.3390/life6040042 (2016).

565 8 Abed, R. M., Dobretsov, S. & Sudesh, K. Applications of cyanobacteria in
566 biotechnology. *J Appl Microbiol* **106**, 1-12, doi:10.1111/j.1365-2672.2008.03918.x
567 (2009).

9 Wang, R. *et al.* Construction and characteristics of artificial consortia of Scenedesmus obliquus-bacteria for S. obliquus growth and lipid production. *Algal Research* **12**, 436-445, doi:http://doi.org/10.1016/j.algal.2015.10.002 (2015).

10 Sudek, S., Haygood, M. G., Youssef, D. T. & Schmidt, E. W. Structure of trichamide, a cyclic peptide from the bloom-forming cyanobacterium Trichodesmium erythraeum, predicted from the genome sequence. *Appl Environ Microbiol* **72**, 4382-4387, doi:10.1128/AEM.00380-06 (2006).

11 Li, B. *et al.* Catalytic promiscuity in the biosynthesis of cyclic peptide secondary metabolites in planktonic marine cyanobacteria. *Proc Natl Acad Sci U S A* **107**, 10430-10435, doi:10.1073/pnas.0913677107 (2010).

12 Cubillos-Ruiz, A., Berta-Thompson, J. W., Becker, J. W., van der Donk, W. A. & Chisholm, S. W. Evolutionary radiation of lanthipeptides in marine cyanobacteria. *Proceedings of the National Academy of Sciences* **114**, E5424-E5433, doi:10.1073/pnas.1700990114 (2017).

13 Kleigrewe, K., Gerwick, L., Sherman, D. H. & Gerwick, W. H. Unique marine derived cyanobacterial biosynthetic genes for chemical diversity. *Nat Prod Rep* **33**, 348-364, doi:10.1039/c5np00097a (2016).

14 Zarzycki, J., Axen, S. D., Kinney, J. N. & Kerfeld, C. A. Cyanobacterial-based approaches to improving photosynthesis in plants. *J Exp Bot* **64**, 787-798, doi:10.1093/jxb/ers294 (2013).

15 McCluskey, K. A Review of Living Collections with Special Emphasis on Sustainability and Its Impact on Research Across Multiple Disciplines. *Biopreserv Biobank* **15**, 20-30, doi:10.1089/bio.2016.0066 (2017).

16 Boundy-Mills, K. *et al.* The United States Culture Collection Network (USCCN): Enhancing Microbial Genomics Research through Living Microbe Culture Collections. *Appl Environ Microbiol* **81**, 5671-5674, doi:10.1128/AEM.01176-15 (2015).

17 Miller, S. R. & Castenholz, R. W. Evolution of thermotolerance in hot spring cyanobacteria of the genus Synechococcus. *Appl Environ Microbiol* **66**, 4222-4229 (2000).

18 Nadeau, T.-L., Milbrandt, E. C. & Castenholz, R. W. EVOLUTIONARY RELATIONSHIPS OF CULTIVATED ANTARCTIC OSCILLATORIANS (CYANOBACTERIA). *Journal of Phycology* **37**, 650-654, doi:10.1046/j.1529-8817.2001.037004650.x (2001).

19 Nadeau, T.-L. & Castenholz, R. W. CHARACTERIZATION OF PSYCHROPHILIC OSCILLATORIANS (CYANOBACTERIA) FROM ANTARCTIC MELTWATER PONDS. *Journal of Phycology* **36**, 914-923, doi:10.1046/j.1529-8817.2000.99201.x (2000).

606 20 Norris, T. B. & Castenholz, R. W. Endolithic photosynthetic communities within
607 ancient and recent travertine deposits in Yellowstone National Park. *FEMS*
608 *Microbiol Ecol* **57**, 470-483, doi:10.1111/j.1574-6941.2006.00134.x (2006).

609 21 Toplin, J. A., Norris, T. B., Lehr, C. R., McDermott, T. R. & Castenholz, R. W.
610 Biogeographic and phylogenetic diversity of thermoacidophilic cyanidiales in
611 Yellowstone National Park, Japan, and New Zealand. *Appl Environ Microbiol* **74**,
612 2822-2833, doi:10.1128/AEM.02741-07 (2008).

613 22 Dillon, J. G. & Castenholz, R. W. The synthesis of the UV-screening pigment,
614 scytonemin, and photosynthetic performance in isolates from closely related
615 natural populations of cyanobacteria (Calothrix sp.). *Environmental Microbiology*
616 **5**, 484-491, doi:10.1046/j.1462-2920.2003.00436.x (2003).

617 23 Dillon, J. G., Tatsumi, C. M., Tandingan, P. G. & Castenholz, R. W. Effect of
618 environmental factors on the synthesis of scytonemin, a UV-screening pigment,
619 in a cyanobacterium (Chroococcidiopsis sp.). *Arch Microbiol* **177**, 322-331,
620 doi:10.1007/s00203-001-0395-x (2002).

621 24 Camacho, A., Garcia-Pichel, F., Vicente, E. & Castenholz, R. W. Adaptation to
622 sulfide and to the underwater light field in three cyanobacterial isolates from Lake
623 Arcas (Spain). *FEMS Microbiology Ecology* **21**, 293-301, doi:10.1111/j.1574-
624 6941.1996.tb00126.x (1996).

625 25 Allen, M. M. & Stanier, R. Y. Selective Isolation of Blue-green Algae from Water
626 and Soil. *Microbiology* **51**, 203-209, doi:doi:10.1099/00221287-51-2-203 (1968).

627 26 Castenholz, R. W. in *The Prokaryotes: A Handbook on Habitats, Isolation, and*
628 *Identification of Bacteria* (eds Mortimer P. Starr *et al.*) 236-246 (Springer Berlin
629 Heidelberg, 1981).

630 27 Schloss, P. D. *et al.* Introducing mothur: open-source, platform-independent,
631 community-supported software for describing and comparing microbial
632 communities. *Applied and environmental microbiology* **75**, 7537-7541,
633 doi:10.1128/AEM.01541-09 (2009).

634 28 Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D.
635 Development of a dual-index sequencing strategy and curation pipeline for
636 analyzing amplicon sequence data on the MiSeq Illumina sequencing platform.
637 *Appl Environ Microbiol* **79**, 5112-5120, doi:10.1128/AEM.01043-13 (2013).

638 29 Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data
639 processing and web-based tools. *Nucleic Acids Res* **41**, D590-596,
640 doi:10.1093/nar/gks1219 (2013).

641 30 Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C. & Knight, R. UCHIME
642 improves sensitivity and speed of chimera detection. *Bioinformatics* **27**, 2194-
643 2200, doi:10.1093/bioinformatics/btr381 (2011).

644 31 McDonald, D. *et al.* An improved Greengenes taxonomy with explicit ranks for
645 ecological and evolutionary analyses of bacteria and archaea. *ISME J* **6**, 610-
646 618, doi:10.1038/ismej.2011.139 (2012).

647 32 Edgar, R. C. Search and clustering orders of magnitude faster than BLAST.
648 *Bioinformatics* **26**, 2460-2461, doi:10.1093/bioinformatics/btq461 (2010).

649 33 Evans, J., Sheneman, L. & Foster, J. Relaxed neighbor joining: a fast distance-
650 based phylogenetic tree construction method. *J Mol Evol* **62**, 785-792,
651 doi:10.1007/s00239-005-0176-2 (2006).

652 34 Letunic, I. & Bork, P. Interactive tree of life (iTOL) v3: an online tool for the
653 display and annotation of phylogenetic and other trees. *Nucleic Acids Res* **44**,
654 W242-245, doi:10.1093/nar/gkw290 (2016).

655 35 Chan, Y., Van Nostrand, J. D., Zhou, J., Pointing, S. B. & Farrell, R. L. Functional
656 ecology of an Antarctic Dry Valley. *Proc Natl Acad Sci U S A* **110**, 8990-8995,
657 doi:10.1073/pnas.1300643110 (2013).

658 36 Jungblut, A. D. *et al.* Microbial Mat Communities along an Oxygen Gradient in a
659 Perennially Ice-Covered Antarctic Lake. *Appl Environ Microbiol* **82**, 620-630,
660 doi:10.1128/AEM.02699-15 (2015).

661 37 Harker, M., Hirschberg, J. & Oren, A. Paracoccus marcusii sp. nov., an orange
662 gram-negative coccus. *Int J Syst Bacteriol* **48 Pt 2**, 543-548,
663 doi:10.1099/00207713-48-2-543 (1998).

664 38 Karavaiko, G. I. *et al.* Reclassification of 'Sulfobacillus thermosulfidooxidans
665 subsp. thermotolerans' strain K1 as Alicyclobacillus tolerans sp. nov. and
666 Sulfobacillus disulfidooxidans Dufresne et al. 1996 as Alicyclobacillus
667 disulfidooxidans comb. nov., and emended description of the genus
668 Alicyclobacillus. *Int J Syst Evol Microbiol* **55**, 941-947, doi:10.1099/ijs.0.63300-0
669 (2005).

670 39 Nazina, T. N. *et al.* Geobacillus gargensis sp. nov., a novel thermophile from a
671 hot spring, and the reclassification of Bacillus vulcani as Geobacillus vulcani
672 comb. nov. *Int J Syst Evol Microbiol* **54**, 2019-2024, doi:10.1099/ijs.0.02932-0
673 (2004).

674 40 Grady, E. N., MacDonald, J., Liu, L., Richman, A. & Yuan, Z. C. Current
675 knowledge and perspectives of Paenibacillus: a review. *Microb Cell Fact* **15**, 203,
676 doi:10.1186/s12934-016-0603-7 (2016).

677 41 Marizcurrena, J. J. *et al.* Searching for novel photolyases in UVC-resistant
678 Antarctic bacteria. *Extremophiles* **21**, 409-418, doi:10.1007/s00792-016-0914-y
679 (2017).

680 42 Oh, T. J., Han, S. R., Ahn, D. H., Park, H. & Kim, A. Y. Complete genome
681 sequence of Hymenobacter sp. strain PAMC26554, an ionizing radiation-
682 resistant bacterium isolated from an Antarctic lichen. *J Biotechnol* **227**, 19-20,
683 doi:10.1016/j.jbiotec.2016.04.011 (2016).

43 Strunecky, O., Elster, J. & Komarek, J. Molecular clock evidence for survival of Antarctic cyanobacteria (Oscillatoriales, Phormidium autumnale) from Paleozoic times. *FEMS Microbiol Ecol* **82**, 482-490, doi:10.1111/j.1574-6941.2012.01426.x (2012).

44 Jungblut, A.-D. *et al.* Diversity within cyanobacterial mat communities in variable salinity meltwater ponds of McMurdo Ice Shelf, Antarctica. *Environmental Microbiology* **7**, 519-529, doi:10.1111/j.1462-2920.2005.00717.x (2005).

45 Shukla, S. P. & Kashyap, A. K. An assessment of biopotential of three cyanobacterial isolates from Antarctic for carotenoid production. *Indian J Biochem Biophys* **40**, 362-366 (2003).

46 Vancanneyt, M. *et al.* Sphingomonas alaskensis sp. nov., a dominant bacterium from a marine oligotrophic environment. *Int J Syst Evol Microbiol* **51**, 73-79, doi:10.1099/00207713-51-1-73 (2001).

47 Cavicchioli, R., Ostrowski, M., Fegatella, F., Goodchild, A. & Guixa-Boixereu, N. Life under nutrient limitation in oligotrophic marine environments: an eco/physiological perspective of Sphingopyxis alaskensis (formerly Sphingomonas alaskensis). *Microb Ecol* **45**, 203-217, doi:10.1007/s00248-002-3008-6 (2003).

48 Lauro, F. M. *et al.* The genomic basis of trophic strategy in marine bacteria. *Proc Natl Acad Sci U S A* **106**, 15527-15533, doi:10.1073/pnas.0903507106 (2009).

49 Ting, L. *et al.* Cold adaptation in the marine bacterium, Sphingopyxis alaskensis, assessed using quantitative proteomics. *Environ Microbiol* **12**, 2658-2676, doi:10.1111/j.1462-2920.2010.02235.x (2010).

50 Hanada, S., Tamaki, H., Nakamura, K. & Kamagata, Y. Crenotalea thermophila gen. nov., sp. nov., a member of the family Chitinophagaceae isolated from a hot spring. *Int J Syst Evol Microbiol* **64**, 1359-1364, doi:10.1099/ijs.0.058594-0 (2014).

51 Anders, H. *et al.* Thermoflavifilum aggregans gen. nov., sp. nov., a thermophilic and slightly halophilic filamentous bacterium from the phylum Bacteroidetes. *Int J Syst Evol Microbiol* **64**, 1264-1270, doi:10.1099/ijs.0.057463-0 (2014).

52 Di Rienzi, S. C. *et al.* The human gut and groundwater harbor non-photosynthetic bacteria belonging to a new candidate phylum sibling to Cyanobacteria. *Elife* **2**, e01102, doi:10.7554/eLife.01102 (2013).

53 Soo, R. M., Hemp, J., Parks, D. H., Fischer, W. W. & Hugenholtz, P. On the origins of oxygenic photosynthesis and aerobic respiration in Cyanobacteria. *Science* **355**, 1436-1440, doi:10.1126/science.aal3794 (2017).

54 Castenholz, R. W. Portrait of a Geothermal Spring, Hunter's Hot Springs, Oregon. *Life (Basel)* **5**, 332-347, doi:10.3390/life5010332 (2015).

722 55 Hanada, S. in *The Prokaryotes: Other Major Lineages of Bacteria and The*
723 *Archaea* (eds Eugene Rosenberg *et al.*) 515-532 (Springer Berlin Heidelberg,
724 2014).

725 56 Saw, J. H. *et al.* Cultivation and complete genome sequencing of Gloeobacter
726 kilaueensis sp. nov., from a lava cave in Kilauea Caldera, Hawai'i. *PLoS One* **8**,
727 e76376, doi:10.1371/journal.pone.0076376 (2013).

728 57 Urakami, T., Oyanagi, H., Araki, H., Suzuki, K.-I. & Komagata, K.
729 Recharacterization and Emended Description of the Genus *Mycoplan*a and
730 Description of Two New Species, *Mycoplana ramosa* and *Mycoplana segnis*.
731 *International Journal of Systematic and Evolutionary Microbiology* **40**, 434-442,
732 doi:doi:10.1099/00207713-40-4-434 (1990).

733 58 Marek-Kozaczuk, M. *et al.* Host-dependent symbiotic efficiency of Rhizobium
734 leguminosarum bv. trifolii strains isolated from nodules of Trifolium rubens.
735 *Antonie Van Leeuwenhoek*, doi:10.1007/s10482-017-0922-7 (2017).

736 59 Semjonovs, P. *et al.* Cellulose synthesis by Komagataeibacter rhaeticus strain P
737 1463 isolated from Kombucha. *Appl Microbiol Biotechnol* **101**, 1003-1012,
738 doi:10.1007/s00253-016-7761-8 (2017).

739 60 Rozenberga, L. *et al.* Characterisation of films and nanopaper obtained from
740 cellulose synthesised by acetic acid bacteria. *Carbohydr Polym* **144**, 33-40,
741 doi:10.1016/j.carbpol.2016.02.025 (2016).

742 61 Mapelli, F. *et al.* Bacterial communities involved in soil formation and plant
743 establishment triggered by pyrite bioweathering on arctic moraines. *Microb Ecol*
744 **61**, 438-447, doi:10.1007/s00248-010-9758-7 (2011).

745 62 Tahon, G. & Willems, A. Isolation and characterization of aerobic anoxygenic
746 phototrophs from exposed soils from the Sor Rondane Mountains, East
747 Antarctica. *Syst Appl Microbiol* **40**, 357-369, doi:10.1016/j.syapm.2017.05.007
748 (2017).

749 63 Shih, P. M. *et al.* Improving the coverage of the cyanobacterial phylum using
750 diversity-driven genome sequencing. *Proceedings of the National Academy of*
751 *Sciences of the United States of America* **110**, 1053-1058,
752 doi:10.1073/pnas.1217107110 (2013).

753 64 Mukherjee, S. *et al.* 1,003 reference genomes of bacterial and archaeal isolates
754 expand coverage of the tree of life. *Nat Biotechnol* **35**, 676-683,
755 doi:10.1038/nbt.3886 (2017).

756 65 Kyrpides, N. C. *et al.* Genomic encyclopedia of bacteria and archaea:
757 sequencing a myriad of type strains. *PLoS Biol* **12**, e1001920,
758 doi:10.1371/journal.pbio.1001920 (2014).

759 66 Field, D. *et al.* The minimum information about a genome sequence (MIGS)
760 specification. *Nat Biotechnol* **26**, 541-547, doi:10.1038/nbt1360 (2008).

67 Kampfer, P., Buczolits, S., Albrecht, A., Busse, H. J. & Stackebrandt, E. Towards a standardized format for the description of a novel species (of an established genus): Ochrobactrum gallinifaecis sp. nov. *Int J Syst Evol Microbiol* **53**, 893-896, doi:10.1099/ijs.0.02710-0 (2003).

Acknowledgements: This work was funded by the College of Agricultural and Environmental Science and the Microbiology & Biochemistry, Molecular, Cellular, and Developmental Biology Graduate Group at University of California Davis (Davis, CA) and the U.S. Department of Energy (DOE) Joint Genome Institute (JGI) in Walnut Creek, CA. Work conducted by the JGI, a DOE User Facility, is supported by DOE's Office of Science under Contract No. DE-AC02-05CH11231. We would also thank Drs. Jorge Rodrigues and John Meeks from UC Davis for providing valuable comments and suggestions on how to improve this manuscript.

We would like to dedicate this publication to Professor Dr. Richard Castenholz who passed away during the completion of this work after a long and satisfying journey in the world of Cyanobacteria. He was, and will remain, a great inspiration to many of us.

Conflict of Interest: The authors declare no conflicts of interest.

Author Contributions

Charles Brooke, Richard Castenholz, David E. Culley, Matthias Hess, and Susannah G. Tringe wrote the manuscript. Richard Castenholz and Matthias Hess designed the experiment. Erik Hawley and Matthias Hess performed experiment. Michael Barton, David E. Culley, Tijana Glavina del Rio, Miranda Harmon-Smith, Erik Hawley, Matthias Hess, Nicole Shapiro, and Susannah G. Tringe generated the data. Michael Barton, Charles Brooke, Morgan P. Connolly, David E. Culley, Javier A. Garcia, Tijana Glavina

787 del Rio, Miranda Harmon-Smith, Erik Hawley, Matthias Hess, and Nicole Shapiro

788 analyzed the data.

Figures & Tables Legends

Figure 1: Geographical locations of co-cultures analyzed. 1.) Antarctica. McMurdo Ice Shelf; Bratina Island; 2.) Spain. Lake Arcas; 3.) USA. Yellowstone National Park; 4.) Mexico. Vizcaino Desert; 5.) USA. Eugene, Oregon. 6.) USA. Hawaii; 7.) Bermuda, Somerset; 8.) Denmark. Limfjord Shallows; 9.) USA. Hunter's Hot Spring, Oregon. (Map downloaded and adapted from https://commons.wikimedia.org/wiki/File:ColoredBlankMap-World-10E.svg#file)

Figure 2: Cladogram of 16S rRNA based community composition of co-cultures under investigation. FECB ID are provided for each co-culture. Sample location is indicated on the branch. Roman numerals on the right indicate the clusters identified at a branch cutoff of 0.05. Symbols (i.e. circles and squares) next to the sample ID indicate habitat type and their color indicates the temperatures at which samples were historically maintained in the CCMEE.

Figure 3: Relative abundance of phyla associated with phototrophic co-cultures. 16S rRNA based community profile. Only phyla recruiting >1% of the reads in at least one of the co-cultures are shown.

808 **Table 1: Summary of photosynthetic co-cultures for which 16S rRNA gene profiles**
809 **were generated.**

810

811 **Table 2: Sequencing statistics and diversity indices for co-cultures investigated in**
812 **this study.**

813

814 **Table 3: Count and phylogenetic classification of identified OTUs at the phylum**
815 **level.** Only OTUs recruiting >0.1% of the co-culture specific reads are shown.

816

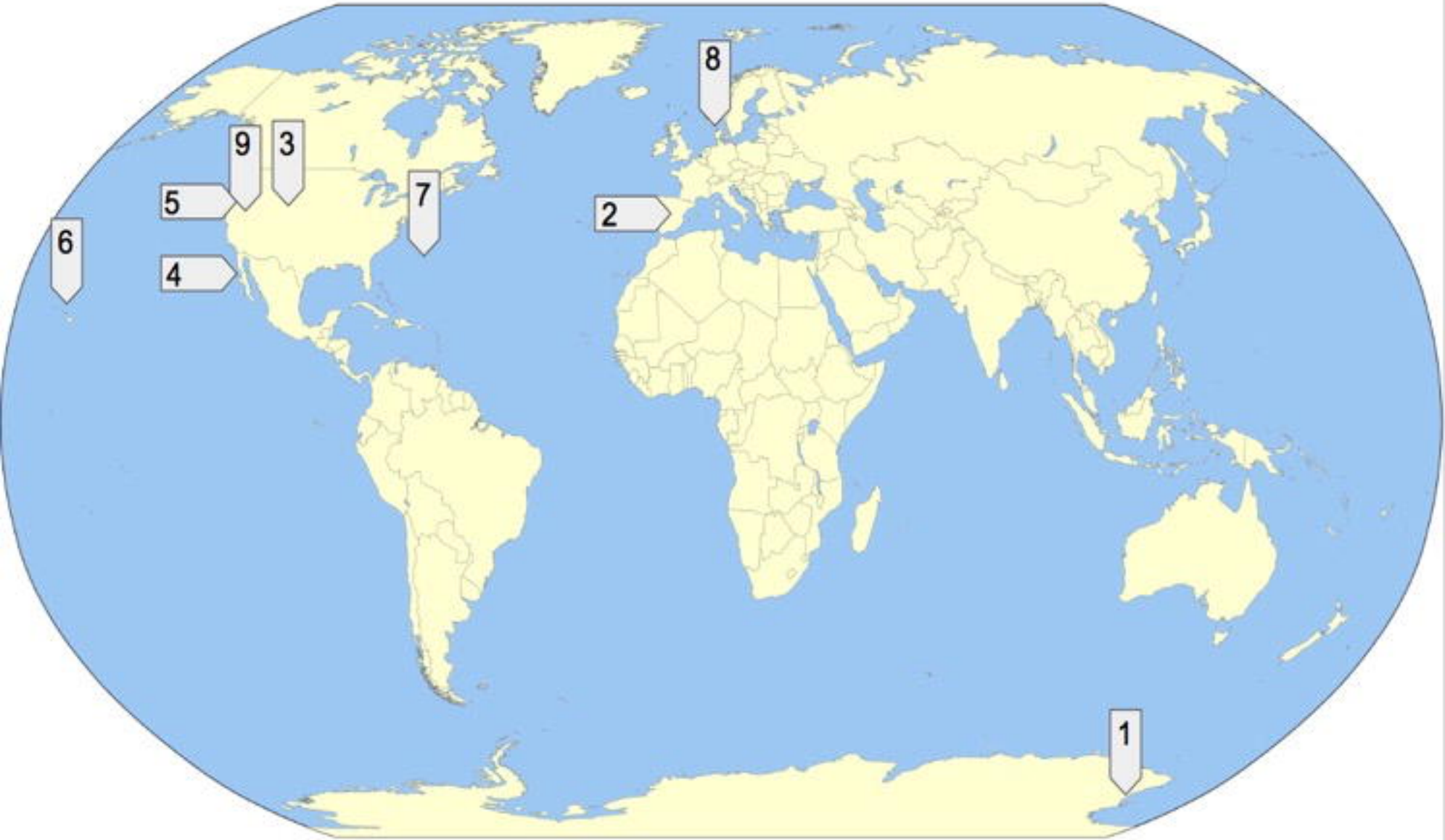
817 **Table 4: Taxonomy relative abundance of dominant OTU identified in each co-**
818 **culture.**

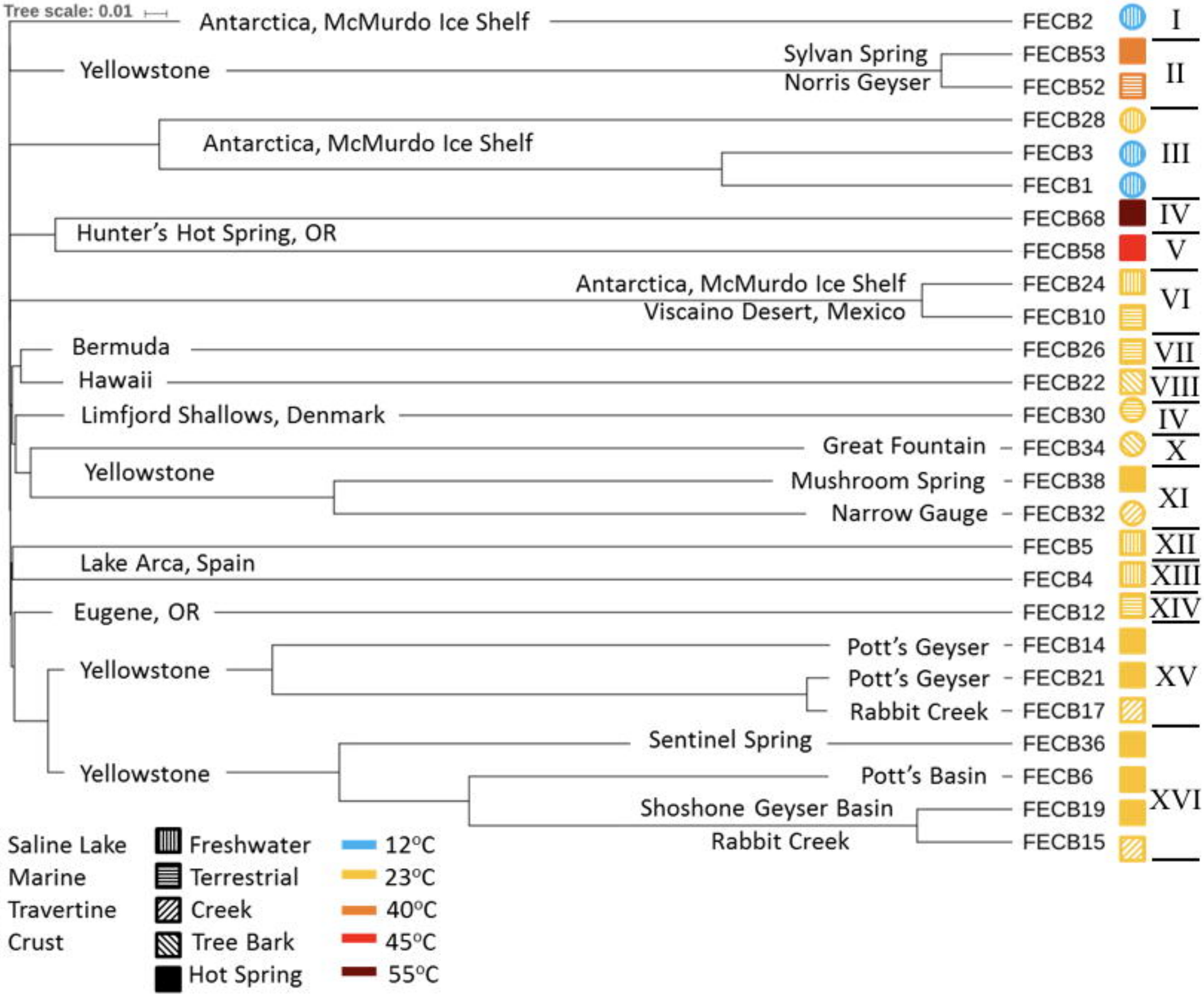
819

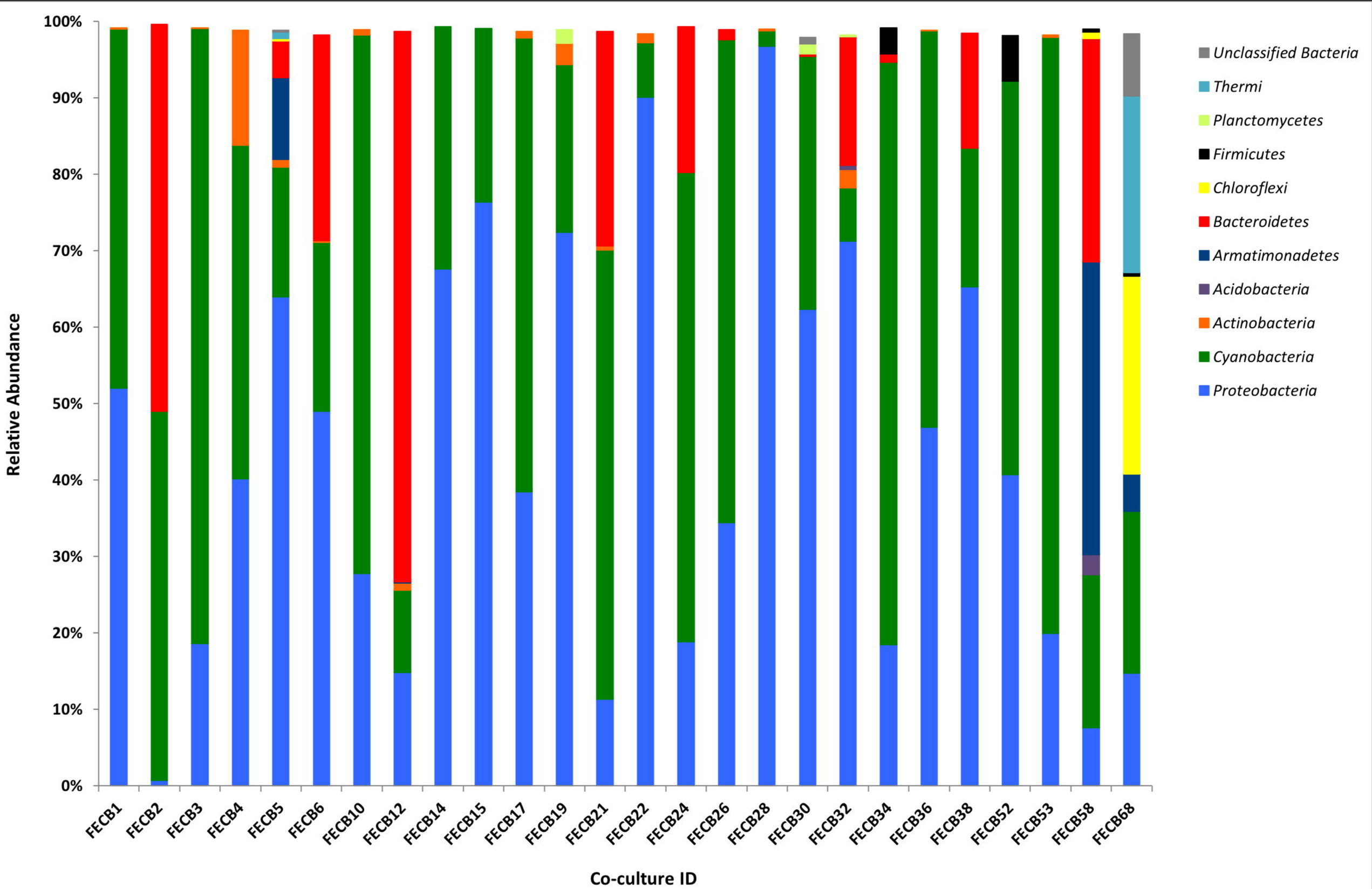
820

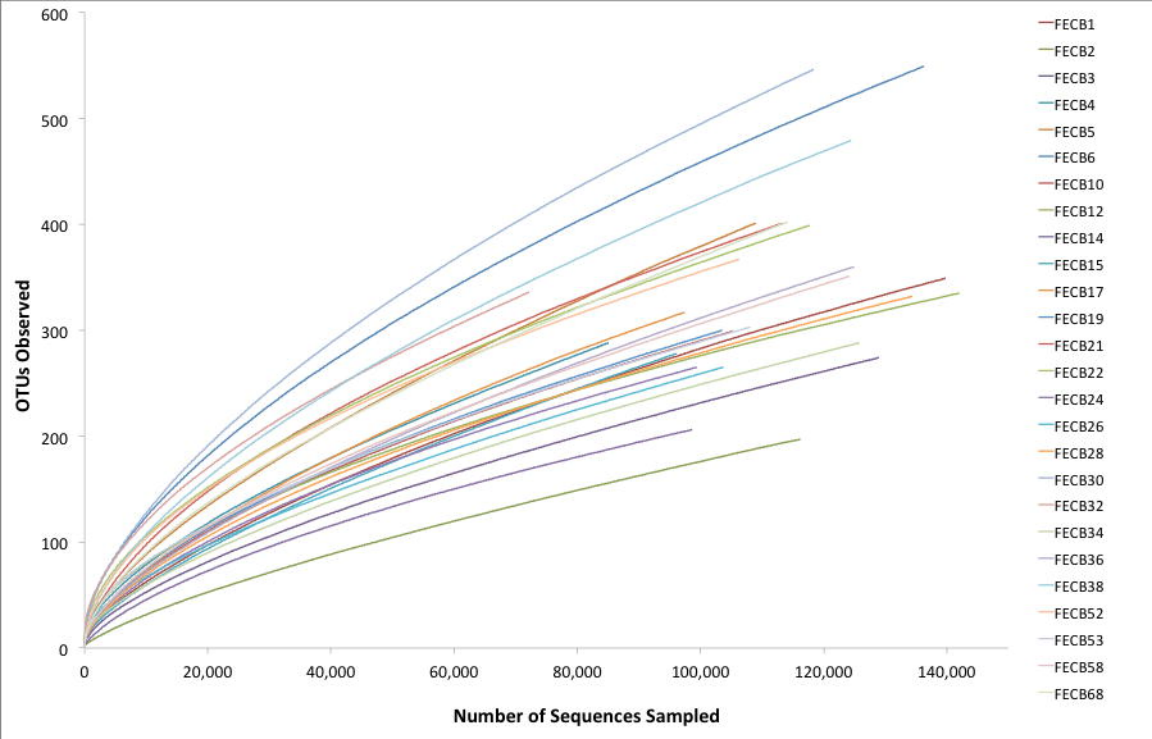
821

822


8
9
3
5
7
2
6
4
1




100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Relative Abundance
FECB1
FECB2
FECB3
FECB4
FECB5
FECB6
FECB10
FECB12
FECB14
FECB15
FECB17
FECB19
FECB21
FECB22
FECB24
FECB26
FECB28
FECB30
FECB32
FECB34
FECB36
FECB38
FECB52
FECB53
FECB58
FECB68
Co-culture ID
Unclassified Bacteria
Thermi
Planctomycetes
Firmicutes
Chloroflexi
Bacteroidetes
Armatimonadetes
Acidobacteria
Actinobacteria
Cyanobacteria
Proteobacteria


600
500
400
300
200
100
0
OTUs Observed
0
20,000
40,000
60,000
80,000
100,000
120,000
140,000
Number of Sequences Sampled
FECB1
FECB2
FECB3
FECB4
FECB5
FECB6
FECB10
FECB12
FECB14
FECB15
FECB17
FECB19
FECB21
FECB22
FECB24
FECB26
FECB28
FECB30
FECB32
FECB34
FECB36
FECB38
FECB52
FECB53
FECB58
FECB68