Diversity and composition of methanotroph communities in caves

Kevin D. Webster[a,,b]# Arndt Schimmelmann[c], Agnieszka Drobniak[d], Maria Mastalerz[d], Laura Rosales Lagarde[e], Penelope J. Boston[f], Jay T. Lennon[g]#

[a]Diné College, School of Science, Technology, Engineering and Math, Tsaile, Arizona, USA 86556

[b]Planetary Science Institute, Tucson, Arizona, USA 85719

[c]Department of Earth and Atmospheric Sciences, Indiana University, Bloomington, Indiana, USA 47405

[d]Indiana Geological and Water Survey, Indiana University, Bloomington, Indiana, USA 47405

[e]Department of Physical and Life Sciences, Nevada State College, School of Liberal Arts and Sciences, Henderson, Nevada, USA.

[f]NASA Astrobiology Institute, NASA Ames Research Center, Moffett Field, California, USA.

[g]Department of Biology, Indiana University, Bloomington, Indiana, USA 47405

#Address correspondence to Kevin D. Webster, webster@psi.edu, or Jay T. Lennon, lennonj@indiana.edu.

ABSTRACT Methane oxidizing microorganisms (methanotrophs) are ubiquitous in the environment and represent a major sink for the greenhouse gas methane (CH_4). Recent studies have demonstrated that methanotrophs are abundant and contribute to CH_4 dynamics in caves. However, very little is known about what controls the distribution and abundance of methanotrophs in subterranean ecosystems. Here, we report a survey of soils collected from > 20 caves in North America to elucidate the factors shaping cave methanotroph communities. Using 16S rRNA sequencing, we recovered methanotrophs from nearly all (98 %) of the samples, including cave sites where CH_4 concentrations were at or below detection limits (≤ 0.3 ppmv). We identified a core methanotroph community among caves that was comprised of high-affinity methanotrophs. Although associated with local-scale mineralogy, methanotroph composition did not systematically vary between the entrances and interior of caves, where CH_4 concentrations varied. We also observed that methanotrophs are able to disperse readily between cave systems showing these organisms have low barriers to dispersal. Last, the relative abundance of methanotrophs was positively correlated with cave-air CH_4 concentrations suggesting that these microorganisms contribute to CH_4 flux in subterranean ecosystems.

IMPORTANCE Recent observations have shown that the atmospheric greenhouse gas methane (CH_4) is consumed by microorganisms (methanotrophs) in caves at rates comparable to CH_4 oxidation in surface soils. Caves are abundant in karst landscapes that comprise 14 % of Earth's land surface area, and therefore may represent a potentially important, but overlooked CH_4 sink. We sampled cave soils to gain a better understand the community composition and structure of cave methanotrophs. Our results show that the members of the USC- γ clade are dominant in cave communities and can easily disperse through the environment, that methanotroph relative

abundance was correlated with local scale mineralogy of soils, and that the relative abundance of methanotrophs was positively correlated with CH_4 concentrations in cave air.

KEYWORDS: Biogeography; Cave; Greenhouse Gas; Karst; Methane; Methanotroph

INTRODUCTION

Methane (CH_4) is as major greenhouse gas in Earth's atmosphere, and an energy source for humanity (1, 2). Because of this, it is important to identify the pathways that produce and consume CH_4. Two major processes are responsible for the removal of CH_4 from Earth's atmosphere, oxidation by atmospheric hydroxyl radicals, and by methane-consuming microorganisms (methanotrophs) in surface soils and waters. Top-down and bottom up estimates of yearly atmospheric CH_4 removal indicate hydroxyl radicals account for 90 % of the sink while methanotrophs consume roughly 5 % (3). Despite the influence that methanotrophs have on regulating Earth's climate, gaps remain in our understanding of methanotrophs, like how they respond to changing CH_4 concentrations or how they disperse through the environment (4, 5).

Rocks that host caves are widespread on Earth, and are emerging as an environment that consumes atmospheric CH_4 (6–11) (Figure 1). Recent studies of cave microbial communities have revealed the presence of atmospheric-CH_4 consuming methanotrophs in caves (8, 12–15). The methanotrophs that are thought to be responsible for this are termed high-affinity methanotrophs, and have been previously observed from upland soils. These groups are referred to as the Upland Soil Cluster (USC) -α and -γ clades. Intermittent or unsustained atmospheric

63 CH_4 oxidation has also been observed in methanotrophs among the families of the
64 Alphaproteobacteria (the *Methylocystaceae* and *Beijernickiaceae*), but it is generally thought that
65 these organisms contribute little to atmospheric CH_4 consumption (16). Both high-affinity and
66 low-affinity alphaproteobacterial methanotrophs tend to be the most prominent members of the
67 methanotrophic community in environments with low CH_4 concentrations and high O_2
68 concentrations. They are less common in environments with high CH_4 and low O_2
69 concentrations, like lakes or geological seeps, which are dominated by low-affinity
70 methanotrophs found among families belonging to the Gammaproteobacteria (Methylococcaceae
71 and Methylothermaceae) (17, 18).

72 The factors that control the distribution and abundance of methanotrophs in caves are not
73 well understood. Previous work has shown that methanotrophs are frequently related to CH_4
74 concentrations, soil texture, and the abundances of other microorganisms (19), but it is unclear
75 how these factors, or the local mineralogy influences methanotroph communities in caves. For
76 example, low-affinity methanotrophs exhibit niche partitioning between environments with high
77 and low CH_4 concentrations (18, 20). However, it is unknown if methanotrophic communities
78 exhibit similar changes in community composition along atmospheric to subatmospheric
79 gradients of CH_4.

80 In addition, very little is known about the spatial distribution of methanotrophs, even
81 though such information could shed light on the biogeography of microbes within and among
82 cave ecosystems. While the community composition of methanotrophs may be influenced by
83 environmental conditions of caves, their abundances may also be influenced by their capacity to
84 disperse. Caves are unique and insular habitats that have a patchy distribution in the terrestrial
85 landscape. However, cells may move through cracks and pores from surface environments or

86 aeolian transport into nearby cave systems. These potentially stochastic modes of dispersal could
87 explain be important in structuring cave methanotroph communities and may help explain the
88 global distribution of methanotrophs. For example, similar methanotrophs have been observed in
89 both Hawai'i and the Arctic (4) suggesting that the movement of these organisms may be
90 relatively unhindered. However, biota in caves are typically unique due to their isolation from
91 other environments (21, 22). Thus, caves provide an ideal environment to test whether
92 methanotrophs are able to disperse readily across the environment or are limited by their ability
93 to disperse.

94 In this study, we sampled 42 cave soils from 21 limestone caves in North America to
95 examine the factors regulating the community composition of cave methanotrophs (Figure 1).
96 First, we hypothesized that the methanotroph composition should be similar to what has been
97 reported for methanotrophic communities in other low CH_4 environments, and that the relative
98 abundance of methanotrophs should be correlated with cave-air CH_4 concentrations.
99 Additionally, we hypothesized that, on the local scale, methanotrophs would be affected by the
100 soil texture and mineral composition of cave soils since these features influence the diffusion of
101 air and micronutrient abundances. Finally, on the regional scale we hypothesized that high-
102 affinity methanotrophs should be able to disperse easily between caves.
103

104 **RESULTS**

105 **Methanotroph community structure.** Methanotrophs were present in nearly all caves.
106 Based on 16s rRNA sequencing we identified methanotrophs in 98 % of the cave-soil samples,
107 including locations where CH_4 concentrations were at or below the analytical detection limit (≤
108 0.3 ppmv). The maximum relative abundance of methanotrophs in the cave soil samples was

6.25 % and the median relative abundance was 0.88 %. The average site contained 6 OTUs belonging to methanotrophs. High-affinity methanotrophs comprised a median of 99.0 % of the total methanotrophic community at locations where methanotrophs were observed. Members of the USC-γ clade represented 47.7 % of the community on average (median), while the members of the USC-α clade represented 43.1 % of the community on average (median). Low-affinity methanotrophs represented 0.09 % of the observed methanotroph community on average (median). Members of the USC-γ and -α were represented by 22 and nine OTUs, repectively, whereas low-affinity methanotrophs belonging to the Alphaproteobacteria and Gammaproteobacteria were represented by 18 and 22 unique OTUs, respectively. Members of the USC-γ clade were dominant in the sampled caves (Figure 2). The members of methanotrophic community from Cave 38 (Mexico) were solely high-affinity methanotophs.

Core Methanotroph Community. We defined taxa that were present in at least 60 % of samples as the core methanotroph community (23, 24). Based on this, the core methanotroph community was made up of three OTUs. Two of these OTUs were from members of the USC-α and one was from the USC-γ. These taxa were found together in 79 % of samples, and the most widely distributed OTU was observed in 93 % of samples. The core methanotroph community accounted for 97 % of the total methanotroph sequences recovered across all cave samples.

Environmental influences on community structure. Methanotroph composition was related to the mineralogy of cave soils (Figure 3). The total methanotrophic community was related to the abundance of clinochlorite (multiple regression, $P = 0.008$), muscovite (multiple regression , $P = 0.03$), and microcline (multiple regression, $P = 0.007$). High-affinity

methanotrophs showed relationships with the abundance of clinochlorite (multiple regression, $P = 0.03$), and microcline (multiple regression, $P = 0.02$), while the low-affinity methanotrophs were not related to the abundance of any observed mineral ($P > 0.05$). In contrast to mineralogical associations, the composition of the methanotrophic community was not related to the distributions of grain size that comprised a soil (regression, $P > 0.05$).

Biogeography. In ecology, the distance decay relationship (DDR) describes how community similarity changes with increasing geographic distance and environmental similarity (25, 26). Briefly, as the space between two points increases, both biological communities and the environment tend to become less similar. The degree to which species can disperse through the environment can alter how quickly the community assemblage changes across space. Communities that are hindered by dispersal tend to change more quickly than measures of environmental similarity, while communities that have have a high capacity to disperse tend to change as quickly as measures of environmental similarity. Within the methanotroph assemblages, we found that the composition of high-affinity taxa changed at the same rate as geographic distance and measures of environmental similarity ($m_{high\text{-}afinnity} = -0.09$; $m_{env} = -0.05$; Difference of slopes test, $P = 0.10$) (Figure 4). Meanwhile, the composition of low-affinity methanotrophs and the measures of environmental similarity changed little with geographic distance ($m_{low\text{-}affinity} = 0.10 \pm 0.06$; Linear regression, $P = 0.09$; $m_{env} = -0.04 \pm 0.05$; Linear regression, $P = 0.43$). At smaller spatial scales, the relative abundance of methanotrophs was not related to the distance from a cave entrance (High-affinity: Spearman's rank correlations, $\rho = -0.15$, S = 4195, $P = 0.45$; Low-affinity: Spearman's rank correlations, $\rho = 0.31$, S = 2488, $P = 0.09$).

155

156 **Relationships with cave air CH_4 concentrations.** The CH_4 concentration in cave air
157 from caves in the USA was positively correlated with the relative abundance of methanotrophs
158 (Figure 5, Spearman's test, $\rho = 0.40$, S = 5451, $P < 0.01$). This relationship was largely driven by
159 high-affinity methanotrophs (Spearman's test, $\rho = 0.42$, S = 5778, $P < 0.01$). The relative
160 abundances of low-affinity methanotrophs were not correlated with the CH_4 concentration (Low-
161 affinity methanotrophs: Spearman's test, $\rho = 0.03$, S = 9601, $P = 0.86$). The highest relative
162 abundance of methanotrophs at any site occurs in a sample from Cave 38 which has internal
163 sources of CH_4 and hydrogen sulfide (H_2S) (27), however samples from this cave also contained
164 low relative abundances of methanotrophs.

165

166 **DISCUSSION**

167 **Methanotroph community composition and diversity in caves.** Methanotrophs are
168 widespread in cave environments, with high-affinity methanotrophs dominating these
169 ecosystems. The dominance of methanotrophs from the high-affinity clades in the sampled caves
170 is consistent with observations that have been made in other low CH_4 environments. For
171 example, high-affinity taxa dominate methanotroph assemblages in forest soils, Arctic soils, and
172 soils associated with basaltic rocks (4, 27, 28). Additionally, the high relative abundance of
173 methanotrophs in the cave soils from this study, and in particular of the USC-γ clade, agrees with
174 observations from other caves where relative abundances of high-affinity methanotrophs can
175 exceed 10 % of the entire bacterial community (8, 12, 14, 15). The relative abundances of
176 methanotrophs observed in our study are higher than the observations of methanotrophs from
177 caves in Vietnam (7) and contrasts with the relatively depauperate assemblage found in a
178 Spanish cave, which reportedly consisted of only three species (29). The highest relative

179 abundance of methanotrophs at any site occurs in a sample from Cave 38 which has an internal
180 CH_4 source (30).

181 Although less abundant, low-affinity methanotrophs from the Alphaproteobacteria and
182 Gammproteobacteria were also commonly observed in our cave survey. Low-affinity
183 alphaproteobacterial methanotrophs are typically more abundant in aerated soils where they have
184 access to oxygen, whereas low-affinity gammaproteobacterial methanotrophs are more often
185 found in environments with high CH_4 concentrations and low oxygen availability (17, 18). Cave
186 soils are typically nutrient poor (31, 32), implying that anoxic conditions created by the decay of
187 organic matter are not likely to be widespread. Thus, the presence of low-affinity
188 gammaproteobacterial methanotrophs in cave soils may be limited to local pockets of anoxia
189 which have been suggested based on observations of limited methanogenesis in cave systems
190 (10, 11).

191 We identified a core group of three OTUs that were present in the studied caves. This
192 core community was derived entirely of high-affinity methanotrophs, suggesting that these taxa
193 are well attuned to subterranean environments that they are able to disperse between cave
194 ecosystems. This interpretation is consistent with previous observations suggesting that high-
195 affinity methanotrophs have high-dispersal capabilities. For example, methanotrophs from a
196 newly formed soil in Hawai'i showed close taxonomic affiliations with Arctic methanotrophs
197 (4). Given that microbes are known to travel intercontinental distances on windblown dust (33),
198 it is possible that surface dwelling high-affinity methanotrophs (27, 34) may be carried into
199 caves on windblown dust or other passive vectors.

200 It is important to note that our survey presents a fairly conservative estimate for diversity
201 and abundance of methanotrophs in cave ecosystems. In the bacteria, methanotrophy is carried

202 out through the presence of the enzyme methane monooxygenase. Methanotrophs are viewed as
203 specialists with regard to their metabolism (35), and, as a result, once the enzymatic machinery
204 has been established in a lineage, it is likely to be conserved since there are barriers to evolving
205 back into generalist species (36). Recent findings suggest that some high-affinity methanotrophs
206 may be able to use other trace gases for carbon and energy, however (37). Often, diversity
207 inventories of methanotrophs focus on sequences of the alpha subunit of the particulate methane
208 monooxygenase enzyme (34). 16S rRNA analysis of methanotrophs may therefore miss some
209 individuals in a community. However, in some studies, it has been shown that the two
210 approaches yield similar estimates of diversity (38).
211

212 **Microenvironmental influences on methanotroph communities.** Our results indicate
213 that soil properties can play an important role in structuring methanotroph communities. The
214 relationship between soil mineralogy and methanotrophs is most likely related to the abundance
215 of trace metals. For example, copper is the trace metal responsible for the oxidation of CH_4 in
216 particulate methane monooxygenase enzyme. The methanotroph community also showed
217 relationships with the abundance of clinochlorite, muscovite, and microcline. These minerals are
218 associated with felsic metamorphic and igneous rocks, and metals like copper are associated with
219 these kinds of rocks as well (39).

220 In contrast, the grain size of the cave soils did not influence the abundance or
221 composition of methanotrophs, even though these edaphic characteristics strongly affect the
222 physical and chemical properties that influence metabolism. Previous research observed that
223 methanotrophs tend to be more abundant in silt and clay fractions of soils because these fractions
224 made up the bulk of the soils themselves (40, 41). Despite some previous observations showing

225 relationships between methanotrophs and soil characteristics, other studies have also shown
226 absences of relationships between methanotroph community composition and soil characteristics
227 (42). The importance of soil grain size in structuring methanotroph communities remains an open
228 question.

229

230 **Biogeography of methanotroph communities.** We observed regional changes in
231 methanotrophic communities in caves. The community composition of high-affinity
232 methanotrophs changed at the same rate as measures of environmental similarity compared to
233 changes in the distance between the samples. This suggests that process like mass effects—
234 which are dominant when organisms are able to actively disperse through their environment—
235 are operating on this community (43). This observation aligns with past work indicating that high
236 affinity methanotrophs are able to disperse through the environment (4), and the observation that
237 three OTU accounted for 97 % of the total methanotrophic community. Dispersal through
238 surface soils (27, 34) to caves in drip waters, and transport into caves on windblown dust (30)
239 represent two possible ways for methanotrophs to disperse into cave systems.

240 The community composition of low-affinity methanotrophs did not change appreciably
241 with geographic distance nor did measures of environmental similarity for these samples. This
242 suggests that neutral processes are operating on these communities (43) and that random birth-
243 death processes, or stochastic dispersal event are important in determining the distribution and
244 abundance of methanotroph popoulations. These processes are thought to operate more strongly
245 on communities that are not strongly influenced by the environment and are the strongest in

small populations (44). This suggests that these communities may be being passively dispersed into the caves.

Biogeographic patterns at the local scale were also relatively weak. The composition and relative abundance of methanotrophic communities at the entrance versus the interior of caves were statistically indistinguishable. This suggests that environmental conditions at the entrance and interior of the cave are similar in ways the select for similar methanotroph communities. Alternatively, similarity of methanotrophs at the relatively small spatial scale may reflect repeated dispersal events in and out of the cave owing to air exchange that accompanies temperature fluctuations over daily and seasonal periods. Cave size and shape may also contribute to the exchange of microbial propagules between above ground and subterranean habitats.

Methanotroph relative abundance and cave air CH_4 concentrations. The relative abundance of methanotrophs was positively correlated with CH_4 concentrations in cave air. This pattern was driven by the high-affinity methanotrophs. High-affinity methanotrophs are thought to derive a significant portion of their energy from CH_4 at atmospheric concentrations and less energy is available for these organisms at locations in caves with sub-atmospheric CH_4 concentrations. Some of the variability observed in the relationship between high-affinity methanotrophs and cave-air CH_4 concentrations may be explained by daily and seasonal variation in cave air flow patterns. Many caves are known to exhibit faster air flow when the external atmospheric temperature is lower than the temperature of cave air. This leads to higher CH_4 concentrations in caves during the winter (10, 45). One result of this is that the measured CH_4 concentration at a

268 particular point in a cave may or may not be representative of the CH_4 flux at that location, and
269 the CH_4 flux at a particular site is likely a strong driver of methanotroph relative abundance.

270

271 **CONCLUSIONS**

272 The structure of cave methanotroph communities appears to be best understood in terms
273 of how cave methanotrophs access CH_4. Cave methanotrophic communities showed relationships
274 with cave-air CH_4 concentrations. Additionally, high-affinity methanotrophs were numerically
275 dominant. This community structure mirrors that of other cave communities, and other
276 methanotroph communities that appear to consume atmospheric CH_4 (8, 12, 14, 15, 38). High-
277 affinity methanotrophs also comprised a core methanotrophic community with three members
278 present together in 79 % of samples. Cave methanotroph communities showed relationships to
279 minerals in the environment that are known to be related to the abundance of Cu, an essential
280 trace metal for methane oxidation. These lines of evidence suggest these communities are
281 actively consuming CH_4 within the cave environment. The cave methanotrophic community
282 assemblage appears to be actively dispersing across cave environments, suggesting that the
283 dispersal barrier for high-affinity methanotrophs from cave to cave may be low. Questions
284 remain about the relative abundance of methanotrophs in cave communities and the CH_4 flux at a
285 particular site. Further research may be inclined to investigate how USC-γ and USC-α
286 methanotrophs partition themselves across low CH_4 availability niches.

287

288 **MATERIALS AND METHODS**

289 **Data and software availability.** This Targeted Locus Study project has been deposited
290 at DDBJ/EMBL/GenBank under the accession KCRG00000000. The version described in this
291 paper is the first version, KCRG00000000.1. All code and data used in this study can be found in
292 a public GitHub repository (https://github.com/websterkgd/CaveMethanotrophs).
293

294 **Microbial sampling.** We sampled microbial communities from limestone caves along
295 the western front of the Appalachian fold and thrust belt, in intracratonic settings of the USA,
296 and in the Sierra Madre of Mexico. We obtained 42 cave soil samples along transects from their
297 entrances to interiors (Data available at: https://github.com/websterkgd/CaveMethanotrophs).
298 Cave soil samples were only collected from locations that had accumulations of sediments and
299 included walls and floors. Samples were collected using a spatula that had been sterilized with 70
300 volume % ethanol in water from locations that were 0.1 m^2 in area (46). Samples were stored on
301 ice until they could be transferred to a −80 °C freezer.

302

303 **Environmental variables.** We measured multiple environmental variables to assess
304 factors that potentially influence the composition of cave methanotrophic communities. We
305 measured CH_4 concentrations *in-situ* with Fourier Transform Infrared spectroscopy (FTIR)
306 **(**Gasmet DX4030 – Milton Keynes, United Kingdom) or we collected discrete air samples
307 collected in the field for measurement the laboratory with FTIR spectroscopy or gas-
308 chromatography (Varian – Agilent Technologies, Palo Alto, California). Some of the CH_4
309 concentrations listed in this study have been previously published (11, 30), and new data were
310 collected according to methods reported in the same publications. In cases where cave maps
311 were available, the distance from the nearest entrance to the sampling location was calculated

312 along the length of the cave passages. We measured cave soil grain-size distributions with a
313 Malvern Mastersizer 3000 (Malvern Instruments Inc., Westborough Massachusetts). Raw data
314 from the Mastersizer were converted to % gravel, % sand, % silt, and % clay sized particles by
315 volume using the GRADISTAT software package (47). Data used for statistical analyses are
316 available in a public GitHub repository (https://github.com/websterkgd/CaveMethanotrophs).
317 Latitude and longitude information, however, have been removed from the data to protect the
318 individual caves. This information can be obtained by contacting the authors directly.

319

320 **Cave soil mineralogy.** We randomly selected 14 samples for mineralogical analysis
321 using X-ray powder diffraction. Soil sediments were powdered to < 5 μm using a mortar, pestle,
322 and acetone. Following the methods of Furmann et al. (48), we used Bruker D8 Advance X-ray
323 diffractometer with a Sol-X solid-state detector and a W X-ray tube operated at 30 mA and 40
324 kV to identify mineral phases. We placed cave sediment powders in an aluminum large
325 frontpacked mount. Packed mounts were scanned from 2° to 70° using a count time of 2 s per
326 0.02° step. Rietveld refinements were used to determine the abundances of the minerals present
327 in the sample and were quantified with TOPAS software. Proportional abundances of the
328 minerals from the cave soils are reported in Table 1.

329

330 **Molecular techniques.** We extracted genomic DNA from cave soil samples with a
331 MoBio PowerSoilTM extraction kit (MoBio, Carlsbad, California USA). About 10 ng of
332 extracted DNA was used as a template for amplification by polymerase chain reaction (PCR).
333 The V4 region of the 16S rRNA gene was amplified using 515F and 806R with barcoded primers
334 designed to work with the Illumina MiSeq platform (49). PCR reactions were as follows: 3 min

335 denaturing step at 94 °C, followed by 30 cycles of 94 °C for 45 s, 50 °C for 60 s, and 72 °C for
336 90 s. A final 10-minute extension was carried out at 72 °C. Quality of the PCR products was
337 checked by gel electrophoresis. Amplified DNA was cleaned using a commercial kit (Beckman
338 Coulter Agencourt AMPure XP, Indianapolis, Indiana, USA) before quantification using the
339 QuanIt PicoGreen kit (Invitrogen) and after which aliquots were pooled at a final concentration
340 of 20 ng per sample.

341 We sequenced the pooled sample containing PCR products using Illumina MiSeq
342 technology (Illumina Reagent Kit v2, 500 reaction kit) at the Center for Bioinformatics and
343 Genomics at Indiana University. Data quality and unique sequences obtained from the PCR
344 amplifications were analyzed using mothur (50). DNA sequence data were aligned using the
345 Needleman algorithm and read lengths were limited to 197 base pairs. Sequences matching
346 chimeras were removed using VSEARCH (51). Sequences that matched chloroplasts, Archaea,
347 and other non-bacterial sequences were also removed. OTUs were created by binning the data at
348 97 % sequence similarity using the opticlust algorithm (52). OTUs were identified using the
349 SILVA reference database (version 132).

350

351 **Classification of methanotrophs.** We identified methanotrophs from the recovered 16S
352 rRNA sequences using a three step process. First, we compiled a database of organisms known
353 methanotrophs through literature surveys. This included assembling lists of low-affinity
354 methanotrophs from the Methylococcaceae, Methylocystaceae, and Beijernickiaceae (34, 53).
355 We used the genera *Methylacidiphilum*, *Methylacidimicrobium* to identify methanotrophs
356 belonging to the Verrucomicrobia (34, 54). High-affinity methanotrophs were defined as
357 organisms belonging to the USC-*α* and -*γ* clades, and *Methylocapsa gorgona* MG08 (13, 34,

358 53). Second, organisms from this database were checked against the identifications assigned to
359 the OTUs during the initial analysis from the SILVA database (version 132). Third, we compiled
360 a FASTA file of the 16S rRNA gene of 71 known methanotrophs. 16S rRNA sequences for
361 USC-α and USC-γ methanotrophs were obtained using published 16S rRNA sequences for USC-
362 α from (12), for USC-γ from (13), and for *Methylocapsa gorgona* MG08 from (37).. This file
363 was compared against the FASTA file of all organisms in the dataset using the *usearch_global*
364 command to identify putative methanotrophs that may have been missed through the SILVA
365 analysis (usearch v.11.0.667) (55). This list of putative methanotophs was then compared against
366 the NCBI database. If the sequences did not match an organism capable of the function of
367 interest at 98 % similarity they were excluded from downstream analysis. The process just
368 described was intended to prevent the false classification of an organism as a methanotroph.
369 The compiled methanotroph FASTA file used in our analysis is available in a public GitHub
370 repository (https://github.com/websterkgd/CaveMethanotrophs).

371

372 **Statistical analyses.** We assessed the relationships between methanotrophs and
373 environmental conditions through a series of statistical analyses. First, to avoid complication
374 arising from unavoidable variation in the total number of reads between samples, we calculated
375 the relative abundance for each taxon within a sample. Next, to quantify among cave variation,
376 we calculated pairwise dissimilarities using the Bray-Curtis metric. The core methanotroph
377 community was characterized by including taxa that were found in 60 % or more of our samples.
378 We used this cutoff because it is in the middle of the range of what has been used in other studies
379 (23, 24). We evaluated the effect of sediment mineralogy and grain size on the multivariate
380 composition of methanotroph communities using canonical correspondence analysis (CCA) with

the *vegan* package in R (56). If a sample did not have associated mineralogical data, it was not included in the CCA analysis.

To gain insight into the biogeography of cave methanotrophs, we examined distance decay relationships (DDR) while also comparing community composition along gradients from cave exteriors to cave interiors. We analyzed DDR by plotting Bray-Curtis similarity of methanotroph composition against environmental and geographic distance calculated as standardized Euclidean distances. Samples from Cave 38 (Cueva de Villa Luz) were excluded in DDR analyses since they were roughly 2000 km away from the nearest sample. Slopes for geographic and environmental DDRs were calculated using ordinary least squares (23). We then tested for differences in the slopes using the statistical test for a difference of slope between regression lines (57). We used Spearman's test to examine how methanotrophic communities changed along gradients from cave exteriors to interiors. Finally, we tested for a relationship between methanotroph relative abundance and CH_4 concentrations using Spearman's test.

ACKNOWLEDGMENTS

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Award Number DE-SC0006978 to AS. Additional support was provided by a National Speleological Society research grant awarded to KDW, the National Geographic Society Expeditions Council (Grant #EC0644-13 to PJB), the National Science Foundation (1442246 to JTL), the US Army Research Office (W911NF-14-1-0411 to JTL) and the National Aeronautics and Space Administration (80NSSC20K0618 to JTL). We thank Brent Lehmkuhl for assistance in the laboratory. Members of the Bloomington Indiana Grotto facilitated access to caves in Indiana and Kentucky. We thank Rodolfo Gómez

Cruz for his support at Cueva de Villa Luz (Cave 38). ML Larsen, MM Muscarella, and KJ Locey aided in data analysis and interpretation. KDW thanks A Barberán for discussions on genetic analysis that resulted in development of the methods used in this paper. KDW thanks the Planetary Science Institute Mars Laboratory for providing computing resources. The authors declare no conflicts of interest.

REFERENCES

1. Ciais P, Sabine C, Bala G. 2013. Carbon and other biogeochemical cycles, p. 465–570. *In* Stocker, TF, Qin, D, Pattner, G-K, Tignor, M, Allen, SK, Boschung, J, Nauels, A, Xia, X, Bex, V, Midgley, PM (eds.), Climate Change 2013: The Physcial Science Basis. Contribution of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambride, United Kindom and New York, NY, USA.

2. Schimmelmann A, Ensminger SA, Drobniak A, Mastalerz M, Etiope G, Jacobi RD, Frankenberg C. 2018. Natural geological seepage of hydrocarbon gas in the Appalachian Basin and Midwest USA in relation to shale tectonic fracturing and past industrial hydrocarbon production. Sci Total Environ 644:982–993.

3. Kirschke S, Bousquet P, Ciais P, Saunois M, Canadell JG, Dlugokencky EJ, Bergamaschi P, Bergmann D, Blake DR, Bruhwiler L, Cameron-Smith P, Castaldi S, Chevallier F, Feng L, Fraser A, Heimann M, Hodson EL, Houweling S, Josse B, Fraser PJ, Krummel PB, Lamarque J-F, Langenfelds RL, Le Quéré C, Naik V, O'Doherty S, Palmer PI, Pison I, Plummer D, Poulter B, Prinn RG, Rigby M, Ringeval B, Santini M, Schmidt M, Shindell DT, Simpson IJ, Spahni R, Steele LP, Strode SA, Sudo K, Szopa S, van der Werf GR, Voulgarakis A, van Weele M, Weiss RF, Williams JE, Zeng G. 2013. Three decades of global methane sources and sinks. Nat Geosci 6:813–823.

4. King GM, Nanba K. 2008. Distribution of atmospheric methane oxidation and methanotrophic communities on hawaiian volcanic deposits and soils. Microbes Environ 23:326–330.

5. Degelmann DM, Borken W, Drake HL, Kolb S. 2010. Different atmospheric methane-oxidizing communities in european beech and norway spruce soils. Appl Environ Microbiol 76:3228–3235.

6. Ford, Derek C; Williams P. 2013. Karst Hydrogeology and Geomorphology. Jon Wiley & Sons, London, England.

7. Lennon JT, Nguyễn-Thùy D, Phạm TM, Drobniak A, Tạ PH, Phạm NĐ, Streil T, Webster KD, Schimmelmann A. 2017. Microbial contributions to subterranean methane sinks. Geobiology 15:254–258.

8. Waring CL, Hankin SI, Griffith DWT, Kertesz MA, Kobylski V, Wilson NL, Coleman N V., Kettlewell G, Zlot R, Bosse M, Bell G. 2017. Seasonal total methane depletion in limestone caves. Sci Rep 7.

9. Mattey DP, Fisher R, Atkinson TC, Latin J-P, Durrell R, Ainsworth M, Lowry D, Fairchild IJ. 2013. Methane in underground air in Gibraltar karst. Earth Planet Sci Lett 374:71–80.

10. Webster KD, Mirza A, Deli JM, Sauer PE, Schimmelmann A. 2016. Consumption of atmospheric methane in a limestone cave in Indiana, USA. Chem Geol 443:1–9.

11. Webster KD, Drobniak A, Etiope G, Mastalerz M, Sauer PE, Schimmelmann A. 2018. Subterranean karst environments as a global sink for atmospheric methane. Earth Planet Sci Lett 485:9–18.

12. Pratscher J, Vollmers J, Wiegand S, Dumont MG, Kaster A-K. 2018. Unravelling the identity, metabolic potential and global biogeography of the atmospheric methane-oxidizing upland soil cluster α. Environ Microbiol https://doi.org/10.1111/1462-2920.14036.

13. Edwards CR, Onstott TC, Miller JM, Wiggins JB, Wang W, Lee CK, Cary SC, Pointing SB, Lau MCY. 2017. Draft genome sequence of uncultured upland soil cluster *Gammaproteobacteria* gives molecular insights into high-affinity methanotrophy. Genome Announc 5:e00047-17.

14. Zhao R, Wang H, Cheng X, Yun Y, Qiu X. 2018. Upland soil cluster γ dominates the methanotroph communities in the karst Heshang Cave. FEMS Microbiol Ecol 94:1–13.

15. Cheng X, Liu X, Lu X, Wang H, Su C, Zhao R, Bodelier PLE, Wang W. 2021. USCγ Dominated Community Composition and Cooccurrence Network of Methanotrophs and Bacteria in Subterranean Karst. Microbiol Spectr 9:e00820-21.

16. Knief C, Dunfield PF. 2005. Response and adaptation of different methanotrophic bacteria to low methane mixing ratios. Environ Microbiol 7:1307–1317.

17. Bowman JP. 2005. Methylococcales *ord. nov.*, p. 248–270. *In* Brenner, D, Krieg, N, Staley, J, Garrity, G, Boone, D, De Vos, P, Goodfellow, M, Rainey, F, Schleifer, K-H (eds.), Bergey's Manual of Systematic Bacteriology. Springer, Boston, MA.

468 18. Mills CT, Slater GF, Dias RF, Carr SA, Reddy CM, Schmidt R, Mandernack KW. 2013.
469 The relative contribution of methanotrophs to microbial communities and carbon cycling
470 in soil overlying a coal-bed methane seep. FEMS Microbiol Ecol 84:474–494.

471 19. He R, Wooller MJ, Pohlman JW, Quensen J, Tiedje JM, Leigh MB. 2012. Shifts in
472 identity and activity of methanotrophs in Arctic Lake sediments in response to
473 temperature changes. Appl Environ Microbiol 78:4715–4723.

474 20. Krause S, Meima-Franke M, Hefting MM, Bodelier PLE. 2013. Spatial patterns of
475 methanotrophic communities along a hydrological gradient in a riparian wetland. FEMS
476 Microbiol Ecol 86:59–70.

477 21. Gordon MS, Rosen DE. 1962. A cavernicolous form of the poeciliid fish *Poecilia*
478 *sphenops* from Tabasco, Mexico. Copeia 1962:360–368.

479 22. Howarth FG. 1993. High-stress subterranean habitats and evolutionary change in cave-
480 inhabing arthropods. Am Nat 142:S65–S77.

481 23. Björk JR, O'Hara RB, Ribes M, Coma R, Björk JR, Montoya JM. 2017. The dynamic core
482 microbiome: Structure, dynamics and stability. bioRxiv 13.

483 24. Hernandez-Agreda A, Gates RD, Ainsworth TD. 2017. Defining the Core Microbiome in
484 Corals' Microbial Soup. Trends Microbiol 25:125–140.

485 25. Nekola JC, White PS. 1999. The distance decay of similarity in biogeography and
486 ecology. J Biogeogr 26:867–878.

487 26. Locey KJ, Muscarella ME, Larsen ML, Bray SR, Jones SE, Lennon JT. 2020. Dormancy
488 dampens the microbial distance-decay relationship. Philos Trans R Soc B Biol Sci 375.

489 27. Kolb S. 2009. The quest for atmospheric methane oxidizers in forest soils. Environ
490 Microbiol Rep 1:336–346.

491 28. Lau MCY, Stackhouse BT, Layton AC, Chauhan A, Vishnivetskaya TA, Chourey K,
492 Ronholm J, Mykytczuk NCS, Bennett PC, Lamarche-Gagnon G, Burton N, Pollard WH,
493 Omelon CR, Medvigy DM, Hettich RL, Pfiffner SM, Whyte LG, Onstott TC. 2015. An
494 active atmospheric methane sink in high Arctic mineral cryosols. ISME J 9:1880–1891.

495 29. Fernandez-Cortes A, Cuezva S, Alvarez-Gallego M, Garcia-Anton E, Pla C, Benavente D,
496 Jurado V, Saiz-Jimenez C, Sanchez-Moral S. 2015. Subterranean atmospheres may act as
497 daily methane sinks. Nat Commun 6:7003.

498 30. Webster KD, Rosales Lagarde L, Sauer PE, Schimmelmann A, Lennon JT, Boston PJ.
499 2017. Isotopic evidence for the migration of thermogenic methane into a sulfidic cave,
500 Cueva de Villa Luz, Tabasco, Mexico. J Cave Karst Stud 79:24–34.

31. Levy DB. 2007. Oxidation-reduction chemistry of Lechuguilla Cave seepage. J Cave Karst Stud 69:351–358.

32. Laiz L, Groth I, Gonzalez I, Saiz-Jimenez C. 1999. Microbiological study of the dripping water in Altamira cave (Santillana del Mar, Spain). Jounal Microbiol Methods 36:129–138.

33. Barberán A, Ladau J, Leff JW, Pollard KS, Menninger HL, Dunn RR, Fierer N. 2015. Continental-scale distributions of dust-associated bacteria and fungi. Proc Natl Acad Sci 112:5756–5761.

34. Knief C. 2015. Diversity and habitat preferences of cultivated and uncultivated aerobic methanotrophic bacteria evaluated based on pmoA as molecular marker. Front Microbiol 6.

35. Wood AP, Aurikko JP, Kelly DP. 2004. A challenge for 21st century molecular biology and biochemistry: What are the causes of obligate autotrophy and methanotrophy? FEMS Microbiol Rev 28:335–352.

36. Sriswasdi S, Yang C, Iwasaki W. 2017. Generalist species drive microbial dispersion and evolution. Nat Commun 8.

37. Tveit AT, Hestnes AG, Robinson SL, Schintlmeister A, Dedysh SN, Jehmlich N, von Bergen M, Herbold C, Wagner M, Richter A, Svenning MM. 2019. Widespread soil bacterium that oxidizes atmospheric methane. Proc Natl Acad Sci U S A 116:8515–8524.

38. Cai Y, Zhou X, Shi L, Jia Z. 2020. Atmospheric Methane Oxidizers Are Dominated by Upland Soil Cluster Alpha in 20 Forest Soils of China. Microb Ecol 80:859–871.

39. Müller D, Groves DI. 2019. Tectonic Settings of Potassic Igneous Rocks, p. 31–71. *In* Mineral Resource Reviews. Springer International Publishing AG.

40. Bender M, Conrad R. 1994. Methane oxidation activity in various soils and freshwater sediments: Occurrence, characteristics, vertical profiles, and distribution on grain size fractions. J Geophys Res 99:16531–16540.

41. Jensen S, Øvreås L, Daae FL, Torsvik V. 1998. Diversity in methane enrichments from agricultural soil revealed by DGGE separation of PCR amplified 16S rDNA fragments. FEMS Microbiol Ecol 26:17–26.

42. Conrad R. 1995. Soil microbial processes involved in production and consumption of atmospheric trace gases, p. 207–250. *In* Advances in Microbial Ecology.

43. Tornero I, Boix D, Bagella S, Pinto-Cruz C, Caria MC, Belo A, Lumbreras A, Sala J, Compte J, Gascón S. 2018. Dispersal mode and spatial extent influence distance-decay patterns in pond metacommunities. PLoS One 13:e0203119.

535 44. Vellend M. 2010. Synthesis in community Ecology. Q Rev Biol 85:183–206.

536 45. Gregorič A, Zidanšek A, Vaupotič J. 2011. Dependence of radon levels in Postojna Cave
537 on outside air temperature. Nat Hazards Earth Syst Sci 11:1523–1528.

538 46. Pašić L, Kovče B, Sket B, Herzog-Velikonja B. 2010. Diversity of microbial communities
539 colonizing the walls of a Karstic cave in Slovenia. FEMS Microbiol Ecol 71:50–60.

540 47. Blott SJ, Pye K. 2001. Gradistat: A grain size Ddstribution and statistics package for the
541 analysis of unconcolidated sediments. Earth Surf Process Landforms 26:1237–1248.

542 48. Furmann A, Mastalerz M, Schimmelmann A, Pedersen PK, Bish D. 2014. Relationships
543 between porosity, organic matter, and mineral matter in mature organic-rich marine
544 mudstones of the Belle Fourche and Second White Specks formations in Alberta, Canada.
545 Mar Pet Geol 54:65–81.

546 49. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM,
547 Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R. 2012. Ultra-
548 high-throughput microbial community analysis on the Illumina HiSeq and MiSeq
549 platforms. ISME J 6:1621–1624.

550 50. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA,
551 Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ,
552 Weber CF. 2009. Introducing mothur: Open-source, platform-independent, community-
553 supported software for describing and comparing microbial communities. Appl Environ
554 Microbiol 75:7537–7541.

555 51. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. 2016. VSEARCH: A versatile open
556 source tool for metagenomics. PeerJ 2016:4:e2584.

557 52. Westcott SL, Schloss PD. 2017. OptiClust, an Improved Method for Assigning Amplicon-
558 Based Sequence Data to Operational Taxonomic Units. mSphere 2.

559 53. Nazaries L, Murrell JC, Millard P, Baggs L, Singh BK. 2013. Methane, microbes and
560 models: fundamental understanding of the soil methane cycle for future predictions.
561 Environ Microbiol 15:2395–2417.

562 54. Carere CR, Hards K, Houghton KM, Power JF, McDonald B, Collet C, Gapes DJ,
563 Sparling R, Boyd ES, Cook GM, Greening C, Stott MB. 2017. Mixotrophy drives niche
564 expansion of verrucomicrobial methanotrophs. ISME J 11:2599–2610.

565 55. Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST.
566 Bioinformatics 26:2460–2461.

56. Oksanen, J.; Blanchet, F.; Kindt, R.; Legendere, P.; Minchin, P.R.; O'Hara, R.B.O.; Simpson, G.L; Solymos, P.; Stevens, H.H.; Wagner H. 2018. Vegan: Community ecology package. R package version v 2.4-6.

57. Paternoster R, Brame R, Mazerolle P, Piquero A. 1998. Using the correct statistical test for the equality of regression coefficients. Criminology 36:859–866.

58. Weary DJ, Doctor DH. 2014. Karst in the United States: a digital map compilation and database. United States Geol Surv Open-File Rep 1156:1–23.

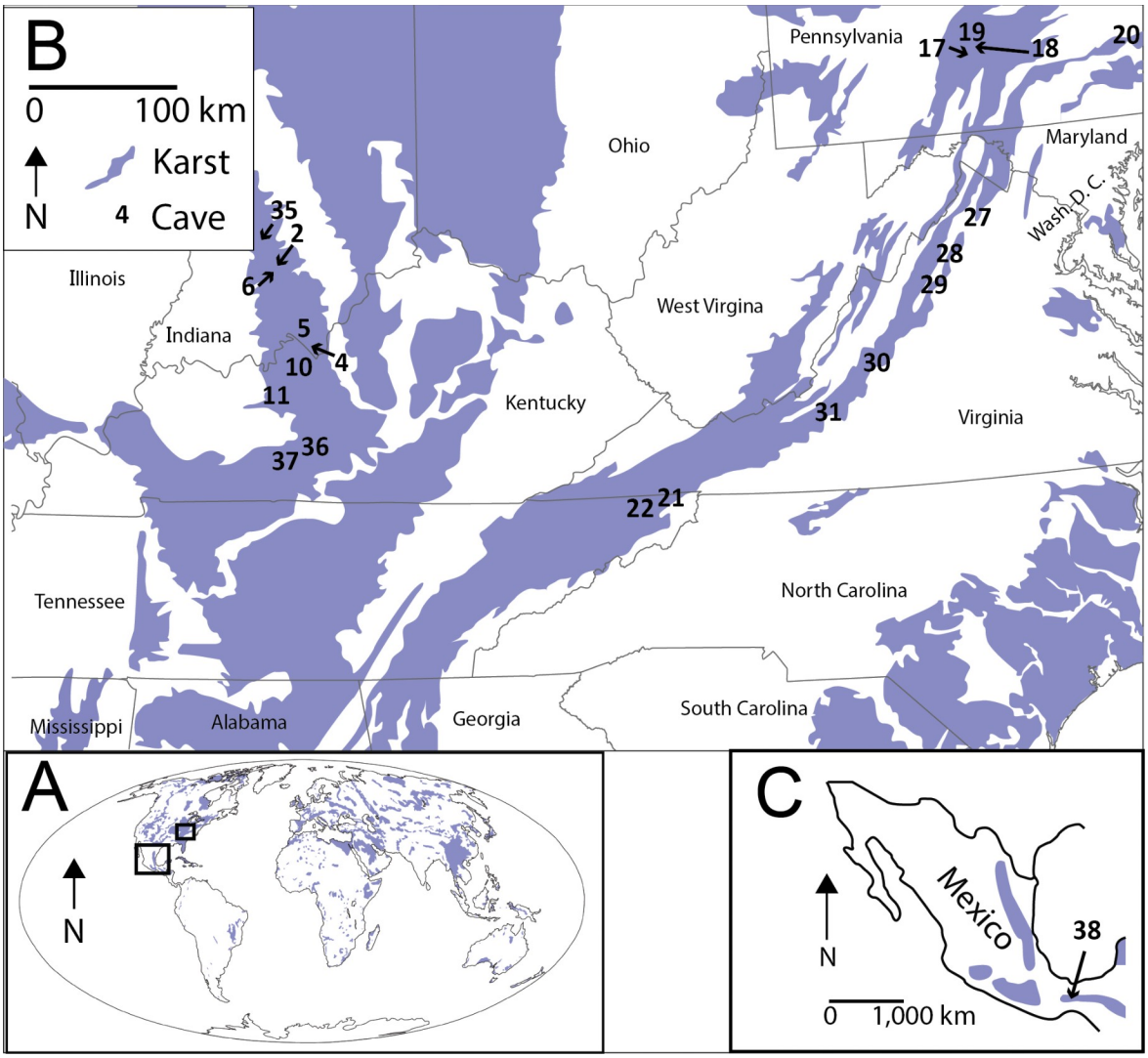

Figure 1: (A) The distributions of karst landscapes at global scale. Inserts show the occurrence of karst (B) in the eastern United States (C) and in Mexico. Numbers represent the locations of sampled caves in this study. Karst land cover data were obtained from (6, 58).

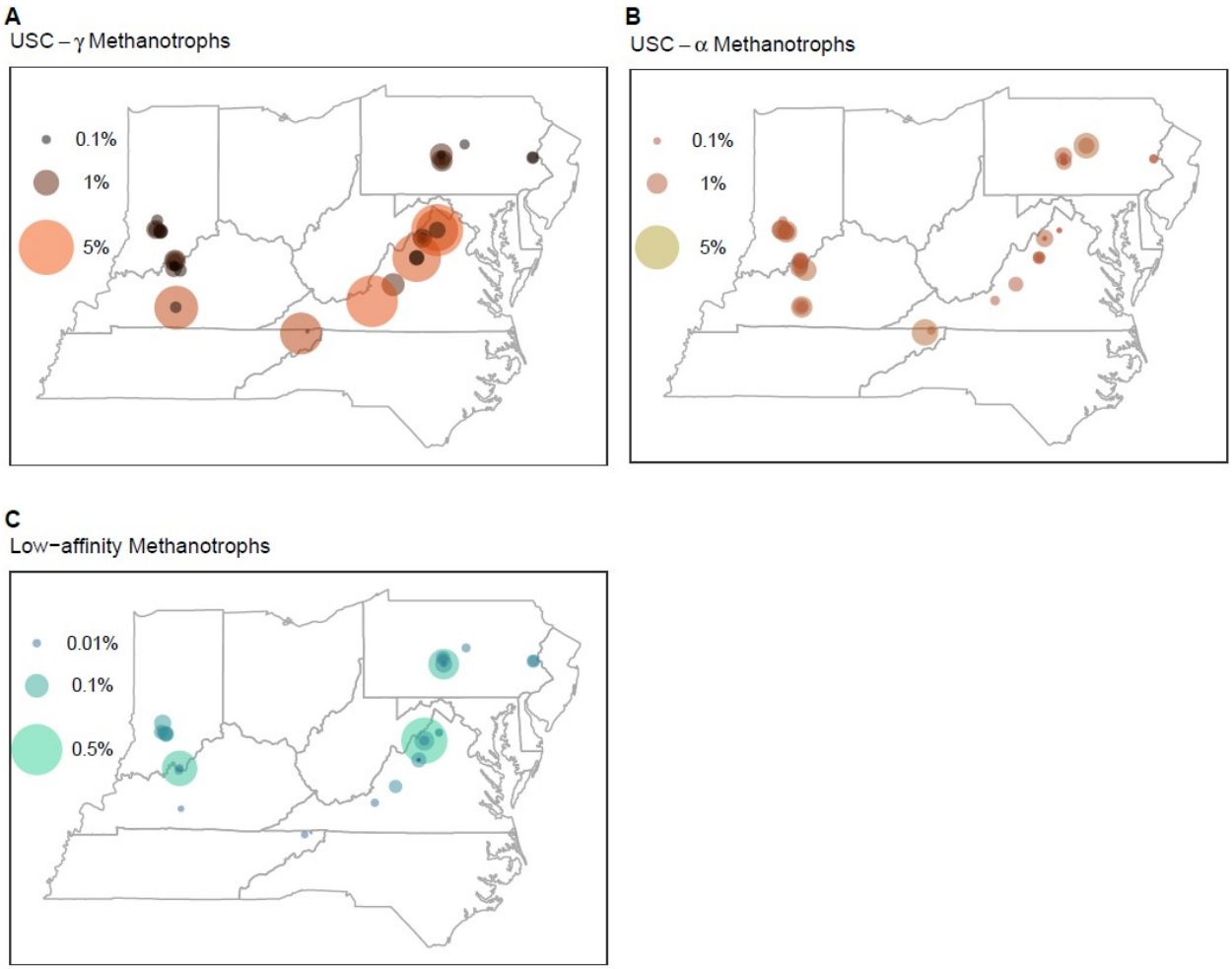

Figure 2: The relative abundance of A) USC-*γ* methanotrophs, B) USC-*α* methanotrophs, and C) Low-affinity methanotrophs in the sampled caves.

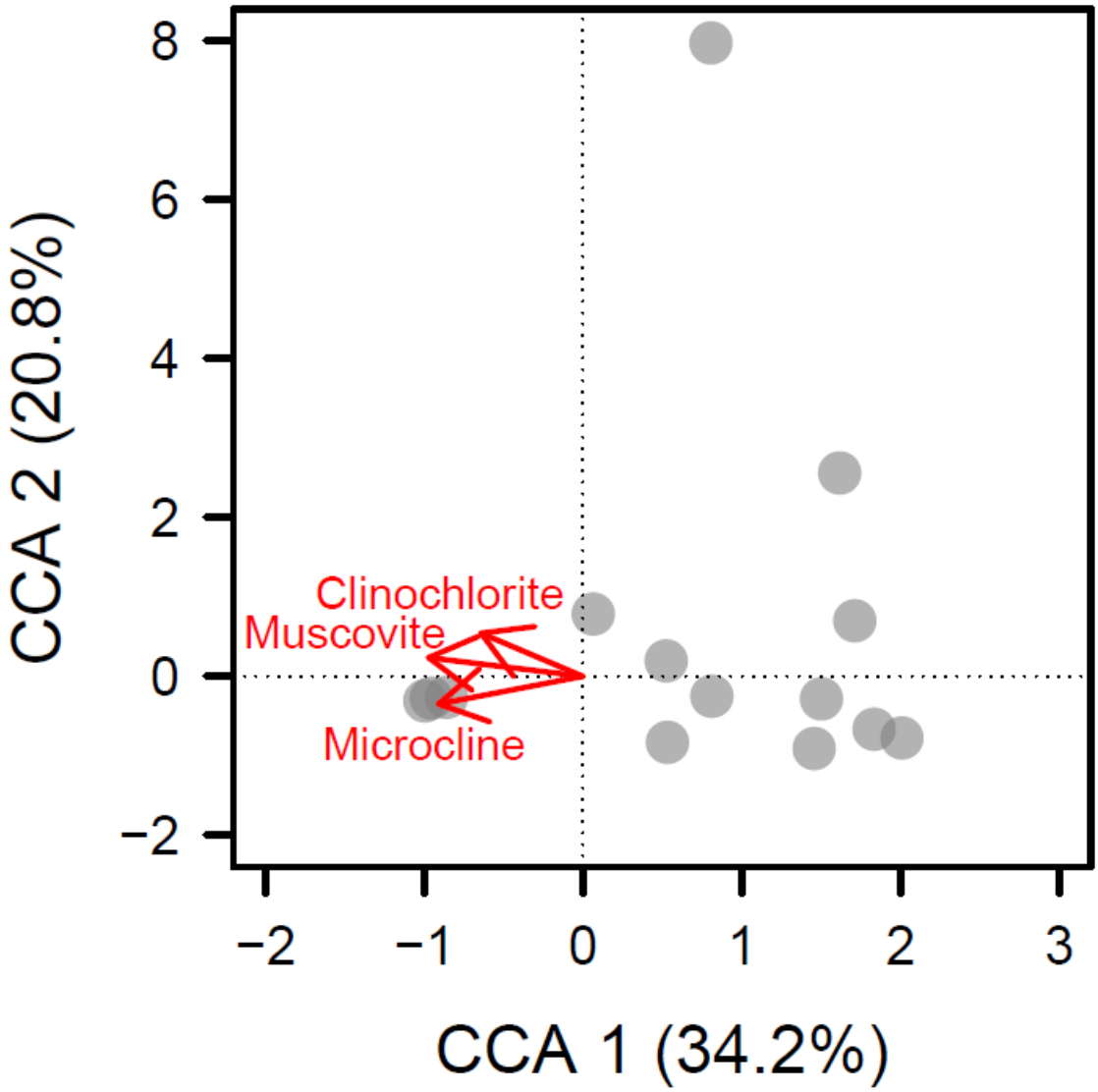

Figure 3: The total methanotrophic community showed relationships with the abundance of muscovite, clinochlorite, and microcline detected in the samples. The vectors show increasing abundances of minerals the samples.

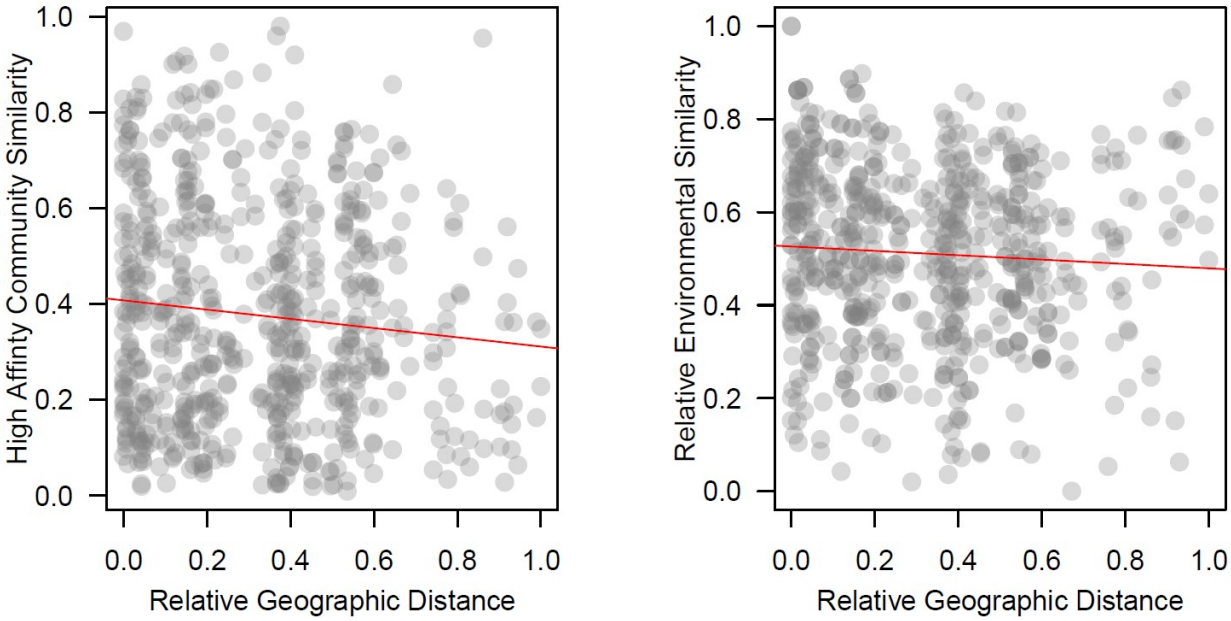

Figure 4: Distance decay analyses of the high affinity methanotrophs. The change in community similarity with geographic distance was statistically indistinguishable from the change in environmental similarity with distance suggesting the organisms are adept at dispersing through the environment. Note that the samples from Cave 38 have been removed from this analysis due to the large geographic separation of these samples from the rest of the samples.

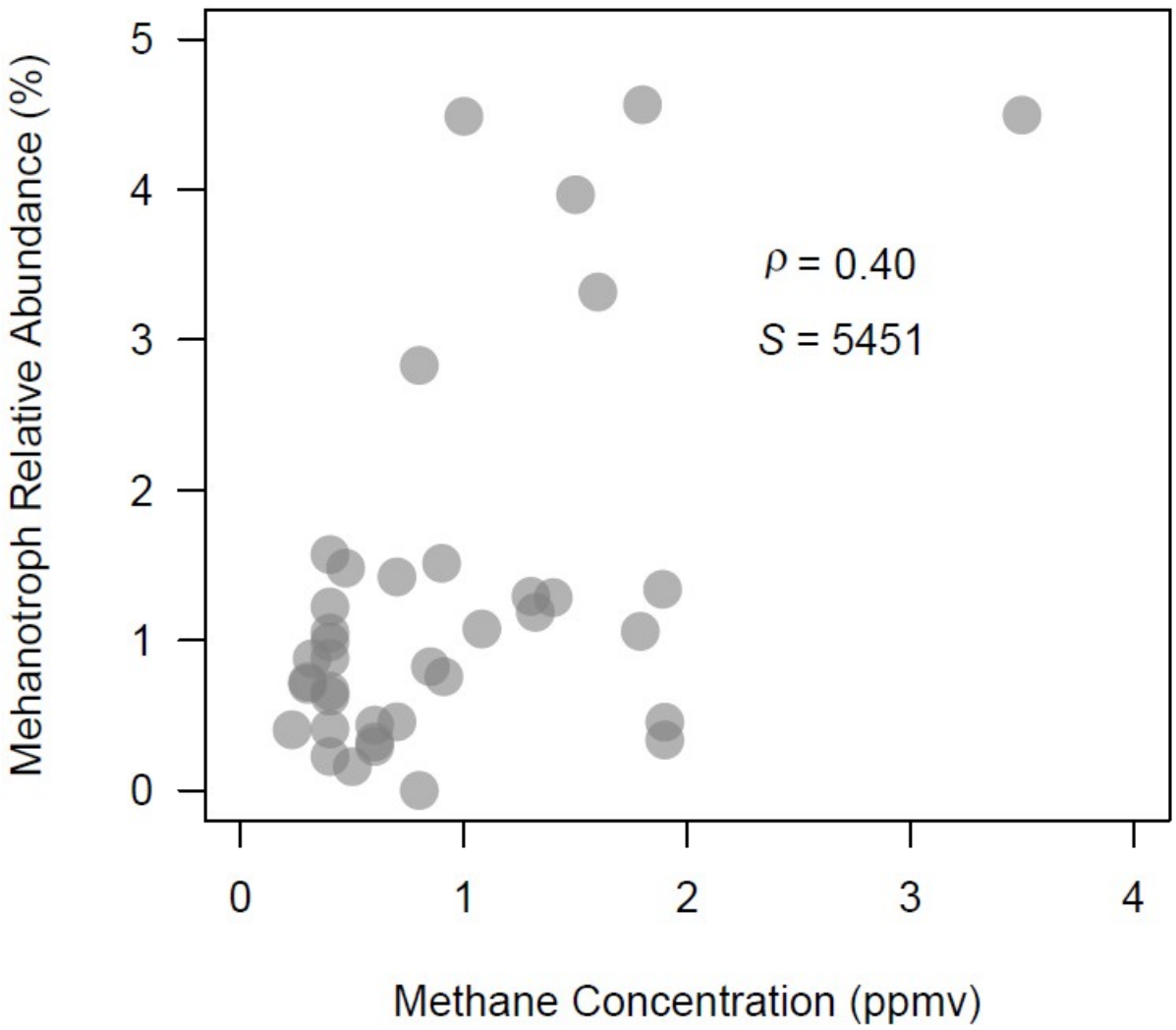

Figure 5: The relative abundance of members of the methanotrophic community plotted against the CH_4 concentration at each sample location ($\rho = 0.40$, S-statistic: S = 5451, $p < 0.01$).

Table 1: The proportional abundance of minerals from the sampled locations

Sample	Cave	Quartz	Muscovite	Clinochlorite	Albite	Orthoclase	Calcite	Microcline	Dolomite	Anorthite	Gypsum
21-10d	Cave 21	0.72	0.14	0.01	0.06	0.04	0.00	0.00	0.00	0.03	0.00
22-11d	Cave 22	0.65	0.16	0.01	0.07	0.00	0.01	0.11	0.00	0.00	0.00
17-1b	Cave 17	0.48	0.28	0.01	0.00	0.00	0.23	0.00	0.00	0.00	0.00
17-1c	Cave 17	0.89	0.09	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00
27-5b	Cave 27	0.33	0.19	0.01	0.06	0.00	0.07	0.28	0.05	0.00	0.00
27-5d	Cave 27	0.50	0.15	0.01	0.00	0.00	0.16	0.14	0.04	0.00	0.00
30-8d	Cave 30	0.40	0.13	0.00	0.02	0.00	0.19	0.09	0.18	0.00	0.00
31-9b	Cave 31	0.40	0.20	0.01	0.00	0.00	0.05	0.16	0.19	0.00	0.00
31-9c	Cave 31	0.14	0.20	0.00	0.00	0.00	0.06	0.08	0.52	0.00	0.00
35-1	Cave 35	0.70	0.09	0.00	0.08	0.04	0.00	0.08	0.00	0.00	0.00
38-9	Cave 38	0.38	0.29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.33
10-1	Cave 10	0.76	0.11	0.00	0.06	0.00	0.00	0.06	0.01	0.00	0.00
4-1	Cave 4	0.21	0.24	0.00	0.00	0.00	0.55	0.00	0.00	0.00	0.00
2-3	Cave 2	0.82	0.11	0.00	0.00	0.00	0.00	0.07	0.00	0.00	0.00
2-2	Cave 2	0.75	0.12	0.00	0.04	0.00	0.02	0.07	0.00	0.00	0.00
2-1	Cave 2	0.74	0.10	0.00	0.06	0.00	0.00	0.10	0.00	0.00	0.00