Diversity and composition of cave methanotrophic communities

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

[a]Department of Earth and Atmospheric Sciences, Indiana University, Bloomington, Indiana, USA.

[b]Indiana Geological and Water Survey, Indiana University, Bloomington, Indiana, USA.

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

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

[e]Department of Biology, Indiana University, Bloomington, Indiana, USA.

*Present address: Kevin D. Webster, Planetary Science Institute, Tucson, AZ, USA.

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

ABSTRACT Methane oxidizing microorganisms (methanotrophs) are a major sink for the greenhouse gas methane (CH_4), and have been investigated in several environments. Recent studies show that CH_4 consumption in caves is pervasive and is a result of active methanotrophy. However, little is known about what controls the distribution and abundance of methanotrophs in caves. We sampled 42 sediments from 21 caves in North America to elucidate the factors shaping cave methanotroph communities. We hypothesized that cave methanotroph communities should be related to cave-air CH_4 concentrations and exhibit dispersal-limited biogeographical patterns due to the insular nature of caves. Using 16S rRNA sequencing, we recovered methanotrophs from 88 % of samples collected, including locations in caves where CH_4 concentrations were at or below detection limits (≤ 0.3 ppmv). Methanotrophs from the Methylocystaceae (Type II) were the dominant methanotrophs as has been seen in other environments with low CH_4 concentrations. Despite being insular ecosystems, we found that the composition of methanotrophs did not vary with distance, both within and among caves. Instead, we found evidence for a core microbiome, perhaps suggesting that high-affinity methanotrophs are not dispersal limited. Additionally, we observed that the relative abundance of methanotrophs was positively related the proportion of gravel in cave sediments and the relative abundance of methylotrophs. Last, we found that the relative abundance of methanotrophs was inversely correlated with cave-air CH_4 concentrations. Our results suggest that methanotrophs in caves have influences on cave biogeochemistry beyond CH_4 oxidation and that high-affinity methanotrophs may disperse easily into caves.

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 be acting as a substantial CH_4 sink. A detailed ecological understanding of the forces that shape methanotrophic communities in caves is lacking. We sampled cave sediments to better understand the community composition and structure of cave methanotrophs. Our results show that the relative abundance of methanotrophs was negatively correlated with CH_4 concentrations in cave air and that methanotrophic communities were similar to each other over distances of 10s of m to 100s of km.

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

INTRODUCTION

Methane (CH_4) is a greenhouse gas in Earth's atmosphere, an energy source for humanity, and has been observed in the atmosphere of Mars (1–3). Thus, enhancing the knowledge of the pathways that add and remove CH_4 to and from Earth's atmosphere has implications for several fields. Two major processes are responsible for the removal of CH_4 from Earth's atmosphere. Oxidation by hydroxyl radicals accounts for 90 % of atmospheric-CH_4 removal and methane-consuming microorganisms (i.e., methanotrophs) in surface soils account for an additional 5 % (4). Despite the influence that methanotrophs have on regulating Earth's climate, gaps in our understanding of these organisms, like how

63 they respond to changing CH_4 concentrations or how they disperse through the
64 environment, remain (2, 3).

65 Rocks that host caves are widespread on Earth, have been observed on other
66 planets, and are emerging as a global CH_4 sink (7–10) (Figure 1). The relative importance
67 of deterministic and stochastic factors responsible for shaping the methanotroph
68 community in caves is unknown (7, 9, 11–14). Observations of atmospheric CH_4
69 consumption in caves suggest that high-affinity or Type II methanotophs may be present in
70 cave environments. These organisms have broad phylogenetic affiliations (15–20), i.e. are
71 known from the Methylocystaceae, as well as several groups in clades affiliated with
72 upland soils (upland soil cluster methanotrophs). They tend to be the most prominent
73 members of the methanotrophic community in environments with low CH_4 concentrations
74 and high O_2 concentrations. They are less common in environments with high CH_4 and
75 low O_2 concentrations, like lakes or geological seeps, which are dominated by
76 methanotrophs from the Methylococcales (Type I methanotrophs) (21, 22).

77 The factors that control the distribution and abundance of methanotrophs in caves
78 are unknown. Previous work has shown that methanotrophs are frequently related to CH_4
79 concentrations, sediment texture, and other microorganisms, but it is unknown how these
80 factors influence methanotroph communities in caves. Type I and Type II methanotrophs
81 are known to exhibit niche partitioning between high and low CH_4 subenvironments
82 respectively (22, 23). However, it is unknown if Type II methanotrophs show similar
83 deterministic changes in community composition with atmospheric to subatmospheric CH_4
84 concentration gradients. Methanotrophs have also been shown to be more abundant in
85 clay-rich soils even though these soils may limit the diffusion of oxygen (24–26). Prior

86 work has documented associations between methanotroph and methylotroph abundances in
87 aquatic environments (18), and that methanotrophs interact with other microbial species
88 through volatile compounds in co-cultures (27).

89 Almost nothing is known about the importance of biogeography in structuring
90 high-affinity methanotroph communities. Previous observations have shown that
91 methanotrophs from Hawai'i exhibited close taxonomic affiliations with Arctic
92 methanotrophs (5). This suggests that these organisms may be able to readily disperse
93 through the environment. However, biota in caves are typically unique due to their
94 isolation from other environments (28, 29). Thus, it may be possible that dispersal
95 limitation is important in structuring high-affinity methanotrophic communities in caves.

96 We sampled 21 caves in North America, analyzed the 16S rRNA gene, and
97 developed the following hypotheses to examine the factors regulating the community
98 composition of cave methanotrophs (Figure 1). First, we hypothesized that that their
99 community structure should mirror the structure of methanotrophic communities in other
100 low CH_4 environments, and that methanotroph abundance should by correlated with cave-
101 air CH_4 concentrations. Additionally, we hypothesized that methanotrophs should show
102 relationships to sediment texture and other microorganisms in the environment. Finally, we
103 hypothesized that high-affinity methanotroph communities in caves should be dispersal-
104 limited due to the insular nature of caves.
105

106 **RESULTS**

Methanotroph community structure. Methanotrophs were recovered in 88 % of the cave-sediment samples including locations where the CH_4 concentrations were at or below our instrumental detection limits of ≤ 0.3 ppmv (Supplemental Table 1). In samples where methanotrophs were recovered, their relative abundances ranged from $1.1*10^{-5}$ to $2.2*10^{-2}$ ($Q_1 = 5.4*10^{-4}$, $Q_2 = 2.3*10^{-3}$, $Q_3 = 4.7*10^{-3}$). Cave methanotrophic communities were dominated by the Methylocystaceae, but members of the Methylococcales were also present (Figure 2). The median abundance of the Methylocystaceae at each site was 96 % of the total methanotrophic community. Members of the Methylocystaceae accounted for 51 unique operational taxonomic units (OTUs). Unknown OTUs at the genus level accounted for 94 % of the OTUs and virtually all (99.99 %) of the observed sequences from the Methylocystaceae. The Methylococcales were more diverse than the Methylocystaceae with 69 unique OTUs but only contributed 997 DNA sequences. 21 OTUs within the Methylococcales were identified from two undescribed families and accounted for 45 % of the total observed Methylococcales.

Methanotrophs from other clades constituted much less of the observed diversity. Seven OTUs were observed from the genus *Methylibium* (beta Proteobacteria: Comamonadaceae), and one DNA sequence from a member of the genus *Methylocella* (Beijerinckiaceae) was observed. Two OTUs were found from the family Acidimethylosilex (Verrucomicrobia) in Cave 2 in Indiana. We did not recover members from the candidate phylum NC10, or members of the USC methanotrophs.

Environmental influences on community structure. The CH_4 concentration in cave air was correlated with the relative abundance of methanotrophs. The relative abundances of both the Methylocystaceae ($log_{10}(RA) = -0.5 \pm 0.3*[CH_4] - 2.4 \pm 0.4$, r^2 =

0.24, $p = 1*10^{-3}$) and Methylococcales ($log_{10}(RA) = -0.5 \pm 0.2*[CH_4] - 3.6 \pm 0.3$), $r^2 =$ 0.41, $p = 1*10^{-4}$) increased with decreasing CH_4 concentration (Figure 2). An analysis of covariance (ANCOVA) test revealed that there was no significant difference between the slopes of the two clades (ANCOVA, $F = 1.6$, $p = 0.21$).

We investigated the influence of the amount of CH_4, gravel, sand, silt, and clay on methanotroph community composition using canonical correspondence analysis. As a whole, the composition of the methanotrophic community was not related to concentrations of CH_4, or the volumetric proportions of gravel, sand, silt, and clay at the sample locations. However, the community composition of the Methylococcales was related to CH_4 concentrations in cave air (Canonical Correspondence Analysis, $p = 0.001$) and the amount of clay at the sample location ($p = 0.008$) (Figure 4). OTUs within the total methanotrophic community and within Methylocystaceae were not related to any of the measured environmental parameters.

We tested if the proportions of gravel, sand, silt, and clay influenced the relative abundance of methanotrophs using Spearman's rank correlation tests. Methanotrophs from the Methylocystaceae, and thus of methanotrophic communities overall, were correlated with the volumetric abundance of gravel in a sample (Spearman's rank correlation, Methylocystaceae, $\rho = 0.39$, $S = 10231$, $p = 0.02$) (Figure 3). The relative abundance of methanotrophs from the Methylocystaceae was not related to any other sediment size fraction present in the samples. Furthermore, the relative abundance of the Methylococcales was not related to any sediment size fraction (Spearman's rank correlations, vs. Gravel $p = 0.7$; vs. Sand $p = 0.12$; vs. Silt $p = 0.76$; vs. Clay $p = 0.90$).

We examined potential interactions between methanotrophs and other microorganisms by testing if the relative abundance of methanotrophs was related to the abundance of methylotrophs in caves. The relative abundance of methanotrophs was positively correlated to the relative abundance of methylotrophs in the sampled cave sediments (Spearman's rank correlation, $\rho = 0.71$, $S = 3849$, $p = 1*10^{-7}$) (Figure 5). The relative abundance of methylotrophs was not, however, related to CH_4 concentrations in cave air (Spearman's rank correlation, $\rho = -0.25$, $S = 8914$, $p = 0.15$).

Biogeographical influences on community structure. We tested the ability of methanotrophs to disperse in the environment by measuring how the methanotroph community changed with measures of geographic distance. Cave methanotroph communities did not show significant relationships to measures of geographic distance. The relative abundance of methanotrophs was not correlated with the distance from a cave entrance (Spearman's rank correlation, $\rho = 0.33$, $S = 2432$, $p = 0.08$). Furthermore, distance-decay analysis—a measure of how the spatial distance between two samples affects their similarity—of the total methanotrophic community assemblage was not statistically significant (Mantel test, $r = 0.08$, $p = 0.14$).

We also tested methanotroph dispersal by determining if a core group of methanotrophs was present in the samples. We defined a core community as taxa present in at least 60 % of samples. If no group of methanotrophs could consistently be found among caves, this may suggest that methanotrophs are dispersal limited. We observed a core methanotroph biome of 3 OTUs. All 3 OTUs were from unidentified genera in the Methylocystaceae. Together these taxa were present in 88 % of samples, and one OTU

174 was present in 86 % of all the samples. The core methanotroph biome accounted for 59 %
175 of the obtained methanotroph sequences.

176

177 **DISCUSSION**

178 **Methanotroph community structure in caves.** Methanotrophs are widespread in
179 cave environments. The dominance of the Methylocystaceae in cave methanotroph
180 communities is consistent with observations that have been made in other low CH_4
181 environments. For example, high-affinity methanotrophs are observed to be dominant in
182 aerated riparian environments, Arctic soils, and soils associated with basaltic rocks (5, 23,
183 30, 31). Additionally, the relative abundance of methanotrophs observed in the sampled
184 cave sediments is similar to the relative abundance of methanotrophs in Arctic soils (0.026
185 to 0.589 %) (32), and from caves in Vietnam (9). However, we did not recover any USC
186 methanotrophs as have been observed in some soils that consume atmospheric CH_4 (16,
187 33). The lack of USC methanotrophs in our study contrasts with the methanotrophic
188 community structure present in a cave from Australia where USC methanotrophs
189 comprised 2 to 12 % of the community (10) and observations of USC-α methanotrophs in
190 several caves (19). Additionally, our results are somewhat inconsistent with observations
191 of methanotrophic communities from soils overlying caves in Australia where
192 *Methylocella* sp. and *Crenothrix* sp. comprised 12 to 15 % of the community (12).
193 The cave sediments in our study also contained relatively rare methanotrophs, but
194 the overall diversity of these taxa exceeded that of the dominant taxa. Minor contributors
195 to the methanotroph community consisted of the Methylococcales, as well as
196 methanotrophs from the Verrucomicrobia. The diversity of methanotrophs identified in our

197 study caves contrasts with the relatively depauperate assemblage found in a Spanish cave,
198 which reportedly consisted of only three species, *Methylocapsa aurea*, *Methylomicrobium*
199 *album*, and *Methylococcus capsulatus* (34). Our results suggest that CH_4 oxidation in
200 caves may be carried out by a relatively diverse yet uncharacterized methanotrophic
201 bacteria.

202 **Environmental influences on methanotroph communities.** The deterministic
203 relationships between cave methanotrophs and their environment may be best understood
204 from the perspective of metabolism. The process of life is fundamentally centered on
205 obtaining energy from the surrounding environment and methanotrophs are no exception.
206 Their distributions in caves appear to result from the fact that methanotrophs rely on both
207 CH_4 and oxygen to derive energy.

208 The relative abundance of methanotrophs was inversely related to CH_4
209 concentrations in cave air. This pattern was documented for two major groups of
210 methanotrophs, the Methylocystaceae and the Methylococcales. While the relative
211 abundance of methanotrophs is sometimes known to positively track CH_4 concentrations
212 (35), methanotrophs have been observed to be negatively correlated with CH_4
213 concentrations in some environments including glacial termini, geologic seeps, and a cave
214 (10, 22, 30, 36). One explanation for the inverse CH_4-methanotoph relationship is that
215 methanotrophs are actively driving down CH_4 concentrations in locations where they make
216 up a large fraction of the total microbial community. Another possibility is that
217 methanotroph abundance in caves may be positively related to CH_4 fluxes, but our
218 observed pattern is a result of seasonality in cave-air flow. Many caves are known to
219 exhibit faster air flow when the external atmospheric temperature is lower than the

220 temperature of cave air. This leads to higher CH_4 concentrations in caves during the winter
221 (13, 37). Our samples were acquired during summer, so it is possible that our sites
222 experience higher CH_4 concentrations during the winter and thus greater CH_4 fluxes on an
223 annual basis which may explain the observed negative relationship.

224 The permeability of cave sediments also appeared to influence the methanotroph
225 community. The relative abundance of methanotrophs in caves was positively related to
226 the amount of gravel in the samples. This may be due to the fact that cave sediments with
227 greater portions of larger clasts, like gravel, may decrease the water-holding capacity while
228 increasing the permeability of cave sediments and the diffusion of gasses (38). As the
229 proportion of larger clasts in a matrix increases, the permeability of the sediment tends to
230 increase (38). This gives methanotrophs greater access to CH_4 and O_2 which are crucial for
231 their metabolism. Previous research observed that methanotrophs tend to be more
232 abundant in silt and clay fractions of soils because these fractions made up the bulk of the
233 soils themselves (24). This previous study did not characterize larger sediment sizes, nor
234 did it examine clay floccules that may serve to enhance permeability.

235 The abundance of clay particles influenced the community composition of the
236 Methylococcales (a less dominant group of methanotrophs in this study) as did the CH_4
237 concentration of cave air. These relationships suggest that members of the
238 Methylococcales may be partitioning themselves across different niches with regard to
239 CH_4 and O_2 availability. The Methylococcales are typically associated with environments
240 that have low O_2 and high CH_4 concentrations. Members of Methylococcales may
241 therefore be responding to dysoxic to hypoxic environments in cave sediments that result
242 from tightly packed clay-sized particles and/or extra-atmospheric CH_4 sources (11, 13, 14).

The community gradient across CH_4 and clay particle size may represent the transition from potentially unknown members in the Methylococcales that consume CH_4 at low (i.e. atmospheric) concentrations to those more tolerant of lower O_2 concentrations.

Organisms alter the physical and chemical characteristics of the environment and thus may influence the distributions of other species in the environment. The positive correlation between methylotrophs and methanotrophs is consistent with observations of metabolic interactions between these functional groups of bacteria. For example, Methylotrophs are known to consume methanol that is produced by methanotrophs (39). Co-occurrences between members of Methylococcales and Methylophilaceae (methylotrophic) have been observed in environmental samples, and so has the transfer of methanol from methanotrophs to methylotrophs (40, 41). A similar process may be occurring in caves. Alternatively, methylotrophs may simply be responding to the presence of other organic compounds in cave air besides CH_4. In addition to consuming methanol, methylotrophs are known to consume methyl halides, like methyl chloride and methyl bromide which are present as trace gases in the atmosphere (42). These gases would be advected into cave air in the same fashion as atmospheric CH_4 (43).

Biogeography of methanotroph communities. Methanotrophic communities did not show statistically significant relationships with local or regional assessments of geographic distance. The composition of methanotrophic communities close to cave entrances was not different compared to those in more interior locations. This pattern has several possible explanations, but two seem to be the most probable. The methanotrophic community in caves is dominated by Type II methanotrophs and these methanotrophs do best under environmental conditions of low CH_4 concentrations and high O_2

266 concentrations. The change from about 2 to < 0.3 ppmv CH_4 in cave air may not impose
267 strong selective pressure on Type II methanotrophs. Additionally, the lack of a pattern in
268 methanotrophic community composition may be a result of the seasonality present in cave
269 air (44, 45). Thus, CH_4 concentrations in cave air may be near the atmospheric
270 concentration at one point in the year and shift to sub-atmospheric concentrations at other
271 times. This could cause location to be a poor predictor of the methanotrophic assemblage
272 due to microbial physiological states like dormancy (46).

273 Likewise, cave methanotrophic communities were not spatially autocorrelated at
274 the regional scale. The lack of a distance-decay relationship in methanotrophic
275 communities from the studied caves may indicate that the scale at which distance-decay
276 relationships occur in these communities is larger than the regional scale of 500 km in our
277 study. This may suggest that dispersal is high among methanotrophs.

278 We found a core group of three OTUs that were present in the studied caves. This
279 suggests that high-affinity methanotrophs are able to move about the environment
280 relatively easily. The fact that these methanotrophs accounted for 59 % of the observed
281 methanotrophs suggests these organisms are reproducing in these environments as well.
282 Previous observations are also indicative of high-dispersal capabilities for these organisms.
283 For example, methanotrophs from a newly formed soil in Hawai'i showed close taxonomic
284 affiliations with Arctic methanotophs (5). Additionally, microbes are known to travel
285 intercontinental distances on windblown dust (47), and perhaps windblown dust may carry
286 methanotrophs into caves. Our results suggest that in cases where microorganisms derive
287 part of their energetic and nutritional requirements from air, that barriers to dispersal into

other habitable subenvironments, like recurring slope lineae or caves on Mars (8, 48), may be decreased.

CONCLUSIONS

The structure of cave methanotroph communities appears to be best understood in terms of how cave subenvironments provide methanotrophs with access to CH_4 and oxygen. Cave methanotrophic communities showed relationships with cave-air CH_4 concentrations and the abundance of gravel-sized clasts in cave sediments. The CH_4 concentrations of cave air and the abundance of gravel both influence how much CH_4 and O_2 methanotrophs have access to. CH_4 is typically at low (i.e., atmospheric) concentrations in caves and as such, Methanotrophs from the Methylocystaceae were numerically dominant. This community structure mirrors that of other atmospheric CH_4 oxidizing communities (9, 32). Further, the community composition of the Methylococcales showed responses to CH_4 concentrations and the abundance of clay in cave sediments that could represent a transition from uncultivated members in the Methylococcales that consume CH_4 at low (i.e. atmospheric) concentrations to those more tolerant of lower O_2 concentrations.

Our results suggest that methanotrophs in caves have secondary influences on the cave-microbial community structure. The abundance of known methylotrophic organisms is correlated with the abundance of methanotrophs. This suggests that molecules produced along the methane oxidation pathway, like methanol, are leaked from methanotrophic cells

allowing for the growth of methylotrophs. Our findings suggest that CH_4 may be an overlooked energy source for cave communities.

Methantrophs did not show relationships with measures of geographic distance in this study. This study focused on environments that are thought to present atmospheric to subatmospheric CH_4 concentrations year-round and in sediments that are forming from the breakdown of similar rocks. The similarity of the environments may be one reason for the observed lack a relationship with measures of distance. This result also suggests that the dispersal barrier for high-affinity methanotrophs from cave to cave may be low.

MATERIALS AND METHODS

Microbial Sampling. We sampled microbial communities from caves along the western front of the Appalachian fold and thrust belt, in intracratonic settings of the USA, and in the Sierra Madre of Mexico. We obtained 42 sediment samples from caves along transects from cave entrances to interiors. Samples were scraped from ~0.1 m^2 large areas of cave walls and floors using a spatula that had been sterilized with 70 volume % ethanol in water (49). Samples were stored on ice until they could be transferred to a −80 °C freezer.

Environmental Parameters. We measured multiple environmental parameters to assess factors that potentially influence the composition of cave methanotrophic communities. We measured CH_4 concentrations *in-situ* with Fourier Transform Infrared Spectroscopy (FTIR) **(**Gasmet DX4030 – Milton Keynes, United Kingdom), or in the laboratory with FTIR gas-chromatography (Varian – Agilent Technologies, Palo Alto,

California) using discrete air samples collected in the field. Some of the CH_4 concentrations listed in this study have been previously published (14, 50), and new data were collected according to methods reported in the same publications. In cases where cave maps were available, the distance from the nearest entrance to the sampling location was calculated along the length of the cave passages. Cave sediment grain-size distributions were measured with a Malvern Mastersizer 3000 (Malvern Instruments Inc., Westborough Massachusetts). Raw data from the Mastersizer were converted to % gravel, % sand, % silt, and % clay sized particles by volume using the GRADISTAT software package (51). Metadata used for statistical analyses are reported in Supplemental Table 2.

Molecular Techniques. Genomic DNA was extracted from cave sediment samples with a MoBio PowerSoilTM extraction kit (MoBio, Carlsbad, California USA). About 10 ng of extracted DNA was used as a template for amplification by polymerase chain reaction (PCR). The V4 region of the 16S rRNA gene was amplified using 515F and 806R with barcoded primers designed to work with the Illumina MiSeq platform (52). DNA amplification was performed using 50 μL reactions from each sample. PCR reactions were carried out by a 3 min denaturing step at 94 °C, followed by 30 cycles of 94 °C for 45 s, 50 °C for 60 s, and 72 °C for 90 s. A final 10-minute extension was carried out at 72 °C. Quality of the PCR amplifications was checked by gel electrophoresis. Amplified DNA was cleaned using a commercial kit (Beckman Coulter Agencourt AMPure XP, Indianapolis, Indiana, USA). Cleaned DNA from the reactions was pooled to a final concentration of 20 ng per sample.

We sequenced the cleaned PCR amplicons using Illumina MiSeq technology (Illumina Reagent Kit v2, 500 reaction kit) at the Center for Bioinformatics and Genomics

354 at Indiana University. Data quality and unique sequences obtained from the PCR
355 amplifications were analyzed using mothur (53). DNA sequence data were aligned using
356 the Needleman algorithm. Sequences matching chimeras were removed using UCHIME
357 (54). Sequences that matched chloroplasts, Archaea, and other non-bacterial sequences
358 were also removed. OTUs were created by binning the data at 97 % sequence similarity.
359 OTUs were identified using a SILVA reference database (version 128).

360 **Statistical Analyses.** The R computing environment was used for all quantitative
361 analyses (55). We used 16S rRNA gene sequences to assess diversity and composition of
362 the methanotroph community. Sequences related to known methanotrophs and
363 methylotrophs were subset from the larger 16S rRNA dataset using the "subset_taxa"
364 function in the *phyloseq* software package to create OTU abundance tables and for only
365 the taxa of interest for a particular sample (56). Relationships between methylotrophs and
366 CH_4 concentrations were examined due to the similarities of their metabolic pathways with
367 methanotrophs. We defined the methanotrophic community as members of the
368 Methylocystaceae, Methylococcales, Methylocella, USC-α, and USC-γ (17, 18). We
369 defined the methylotrophic community as members of the Methylophilaceae and
370 Methylobacteriaceae.

371 We assessed the relationships between methanotrophs and environmental
372 conditions through a series of statistical analyses. Raw abundance data were transformed
373 into fractional abundance data to avoid variations in the number of reads between samples.
374 Community dissimilarity between samples was assessed with Bray-Curtis dissimilarity.
375 Dissimilarity measures were only calculated if the number of OTUs in a phylogenetic
376 grouping was greater than 100 OTUs to avoid statistical problems arising from small

377 sample size. The effect of CH_4 concentrations on the methanotrophic community was
378 assessed using ANCOVA and Mantel tests. We tested the effect of sediment grain size on
379 cave methanotrophic communities using canonical correspondence analysis that was
380 performed in the R package *vegan* (57).

381 The spatial autocorrelation of methanotrophic communities was tested by
382 performing distance-decay analyses. Distance-decay is a spatial-autocorrelation pattern
383 that describes the decreasing similarity of biological communities with increases in
384 geographic distance (58). We performed distance-decay analysis in R by first removing
385 data from Cave 38 (Cueva de Villa Luz) since it was roughly 2000 km away from the
386 nearest samples. We then tested for a distance-decay pattern using a Mantel test in the
387 package *vegan* (57).

388 We characterized the methanotroph core microbiome, which is defined as a suite of
389 members shared in an environment (59). We calculated the core microbiome based on a
390 cutoff of 60 %, a value that is in the middle of ranges that have used when calculating core
391 microbiomes in other ecosystems including corals and humans (60, 61). We also chose this
392 value because it is larger than the majority of samples, but accounts for some
393 unpredictability in an organism's distribution.

394 **Data and software availability.** All code and data used in this study can be found
395 in a public GitHub repository (https://github.com/websterkgd/CaveMethanotrophs).

396

397 **ACKNOWLEDGEMENTS**

398 This material is based upon work supported by the U.S. Department of Energy,
399 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 Dimensions of Biodiversity, Grant/Award Number: 1442246 to JTL, and a US Army Research Office, Grant/Award Number: W911NF-14-1-0411 to JTL. We thank Brent Lehmkuhl for microbiological support. 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. LA Durden assisted in editing an early draft of this manuscript. 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. Webster CR, Mahaffy PR, Atreya SK, Moores JE, Flesch GJ, Malespin C, Mckay

CP, Martinez G, Smith CL, Martin-Torres J, Gomez-Elvira J, Zorzano M-P, Wong MH, Trainer MG, Steele A, Acher Jr. D, Sutter B, Coll PJ, Freissinet C, Meslin P-Y, Gough R V., House CH, Pavlov A, Eigenbrode JL, Glavin DP, Pearson JC, Keymeulen D, Christensen LE, Schwenzer SP, Navarro-Gonzalez R, Pla-Garía J, Rafkin SCR, Vicente-Retorillo Á, Kahanapää H, Viudez-Moreiras D, Smith MD, Harri A-M, Genzer M, Hassler DM, Lemmon M, Crisp J, Sander SP, Zurek RW, Vasavada AR. 2018. Background levels of methane in Mars' atmosphere show strong seasonal variations 360:1093–1096.

4. 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.

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

6. Degelmann DM, Borken W, Drake HL, Kolb S. 2010. Different atmospheric methane-oxidizing communities in european beech and norway spruce soils. Appl

444 Environ Microbiol 76:3228–3235.

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

447 8. Cushing GE, Titus TN, Wynne JJ, Christensen PR. 2007. THEMIS observes
448 possible cave skylights on Mars. Geophys Res Lett 34:4–8.

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

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

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

458 12. McDonough LK, Iverach CP, Beckmann S, Manefield M, Rau GC, Baker A, Kelly
459 BFJ. 2016. Spatial variability of cave-air carbon dioxide and methane
460 concentrations and isotopic compositions in a semi-arid karst environment. Environ
461 Earth Sci 75:700.

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

464 14. 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.

15. Ricke P, Kube M, Nakagawa S, Erkel C, Reinhardt R, Liesack W. 2005. First genome data from uncultured upland soil cluster alpha methanotrophs provide further evidence for a close phylogenetic relationship to *Methylocapsa acidiphila* B2 and for high-affinity methanotrophy involving particulate methane monooxygenase. Appl Environ Microbiol 71:7472–7482.

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

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

18. Lau E, Nolan IV EJ, Dillard ZW, Dague RD, Semple A, Wentzell WL. 2015. High throughput sequencing to detect differences in methanotrophic Methylococcaceae and Methylocystaceae in surface peat, forest soil, and *Sphagnum* moss in Cranesville Swamp Preserve, West Virginia, USA. Microorganisms 3:113–136.

19. 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.

20. 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.

21. 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.

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

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

24. 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.

25. Sessitsch A, Weilharter A, Gerzabek MH, Kirchmann H, Kandeler E. 2001. Microbial population structures in soil particle size fractions of a long-term fertilizer field experiment. Appl Environ Microbiol 67:4215–4224.

26. Hemkemeyer M, Christensen BT, Martens R, Tebbe CC. 2015. Soil particle size fractions harbour distinct microbial communities and differ in potential for microbial mineralisation of organic pollutants. Soil Biol Biochem 90:255–265.

507 27. Veraart AJ, Garbeva P, van Beersum F, Ho A, Hordijk CA, Meima-Franke M,
508 Zweers AJ, Bodelier PLE. 2018. Living apart together - Bacterial volatiles influence
509 methanotrophic growth and activity. ISME J 12:1163–1166.

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

512 29. Howarth FG. 1993. High-stress subterranean habitats and evolutionary change in
513 cave-inhabing arthropods. Am Nat 142:S65–S77.

514 30. Bárcena TG, Finster KW, Yde JC. 2011. Spatial patterns of soil development,
515 methane oxidation, and methanotrophic diversity along a receding glacier forefield,
516 Southeast Greenland. Arctic, Antarct Alp Res 43:178–188.

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

522 32. Martineau C, Pan Y, Bodrossy L, Yergeau E, Whyte LG, Greer CW. 2014.
523 Atmospheric methane oxidizers are present and active in Canadian high Arctic soils.
524 FEMS Microbiol Ecol 89:257–269.

525 33. Lau E, Ahmad A, Steudler PA, Cavanaugh CM. 2007. Molecular characterization of
526 methanotrophic communities in forest soils that consume atmospheric methane.
527 FEMS Microbiol Ecol 60:490–500.

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

531 35. Schütte UME, Cadieux SB, Hemmerich C, Pratt LM, White JR. 2016.
532 Unanticipated geochemical and microbial community structure under seasonal ice
533 cover in a dilute, dimictic Arctic lake. Front Microbiol 7:1–15.

534 36. Bárcena TG, Yde JC, Finster KW. 2010. Methane flux and high-affinity
535 methanotrophic diversity along the chronosequence of a receding glacier in
536 Greenland. Ann Glaciol 51:23–31.

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

539 38. Boadu FK. 2000. Hydraulic conductivity of soils from grain-size distribution: new
540 models. J Geotech Geoenvironmental Eng 126:739–746.

541 39. Kalyuhznaya MG, Martens-Habbena W, Wang T, Hackett M, Stolyar SM, Stahl
542 DA, Lidstrom ME, Chistoserdova L. 2009. *Methylophilaceae* link methanol
543 oxidation to denitrification in freshwater lake sediment as suggested by stable
544 isotope probing and pure culture analysis. Environ Microbiol Rep 1:385–392.

545 40. Beck DAC, Kalyuzhnaya MG, Malfatti S, Tringe SG, Glavina del Rio T, Ivanova
546 N, Lidstrom ME, Chistoserdova L. 2013. A metagenomic insight into freshwater
547 methane-utilizing communities and evidence for cooperation between the
548 *Methylococcaceae* and the *Methylophilaceae*. PeerJ 1:e23.

549 41. Xia F, Zou B, Shen C, Zhu T, Gao X-H, Quan Z-X. 2015. Complete genome
550 sequence of *Methylophilus* sp. TWE2 isolated from methane oxidation enrichment
551 culture of tap-water. J Biotechnol 211:121–122.

552 42. Michener JK, Vuilleumier S, Bringel F, Marx CJ. 2016. Transfer of a catabolic
553 pathway for chloromethane in *Methylobacterium* strains highlights different
554 limitations for growth with chloromethane or with dichloromethane. Front
555 Microbiol 7:1116.

556 43. Barton HA, Taylor MR, Pace NR. 2004. Molecular phylogenetic analysis of a
557 bacterial community in an oligotrophic cave environment. Geomicrobiol J 21:11–
558 20.

559 44. Kowalczk AJ, Froelich PN. 2010. Cave air ventilation and CO_2 outgassing by
560 radon-222 modeling: How fast do caves breathe? Earth Planet Sci Lett 289:209–
561 219.

562 45. Gregorič A, Vaupotič J, Gabrovšek F. 2013. Reasons for large fluctuation of radon
563 and CO_2 levels in a dead-end passage of a karst cave (Postojna Cave, Slovenia). Nat
564 Hazards Earth Syst Sci 13:287–297.

565 46. Jones SE, Lennon JT. 2010. Evidence for limited microbial transfer of methane in a
566 planktonic food web. Aquat Microb Ecol 58:45–53.

567 47. Barberán A, Ladau J, Leff JW, Pollard KS, Menninger HL, Dunn RR, Fierer N.
568 2015. Continental-scale distributions of dust-associated bacteria and fungi. Proc
569 Natl Acad Sci 112:5756–5761.

570 48. McEwen AS, Ojha L, Dundas CM, Mattson SS, Byrne S, Wray JJ, Cull SC,
571 Murchie SL, Thomas N, Gulick VC. 2011. Seasonal flows on warm Martian slopes.
572 Science (80-) 333:740–743.

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

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

579 51. Blott SJ, Pye K. 2001. Gradistat: A grain size Ddstribution and statistics package for
580 the analysis of unconcolidated sediments. Earth Surf Process Landforms 26:1237–
581 1248.

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

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

591 54. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. 2011. UCHIME improves
592 sensitivity and speed of chimera detection. Bioinformatics 27:2194–2200.

593 55. R Core Team. 2017. R: A language and environment for statistical computing. R
594 Foundation for Statistical Computing, Vienna, Austria.

595 56. McMurdie PJ, Holmes S. 2013. Phyloseq: An R package for reproducible
596 interactive analysis and graphics of microbiome census data. PLoS One 8:e61217.

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

600 58. Nekola JC, White PS. 1999. The distance decay of similarity in biogeography and
601 ecology. J Biogeogr 26:867–878.

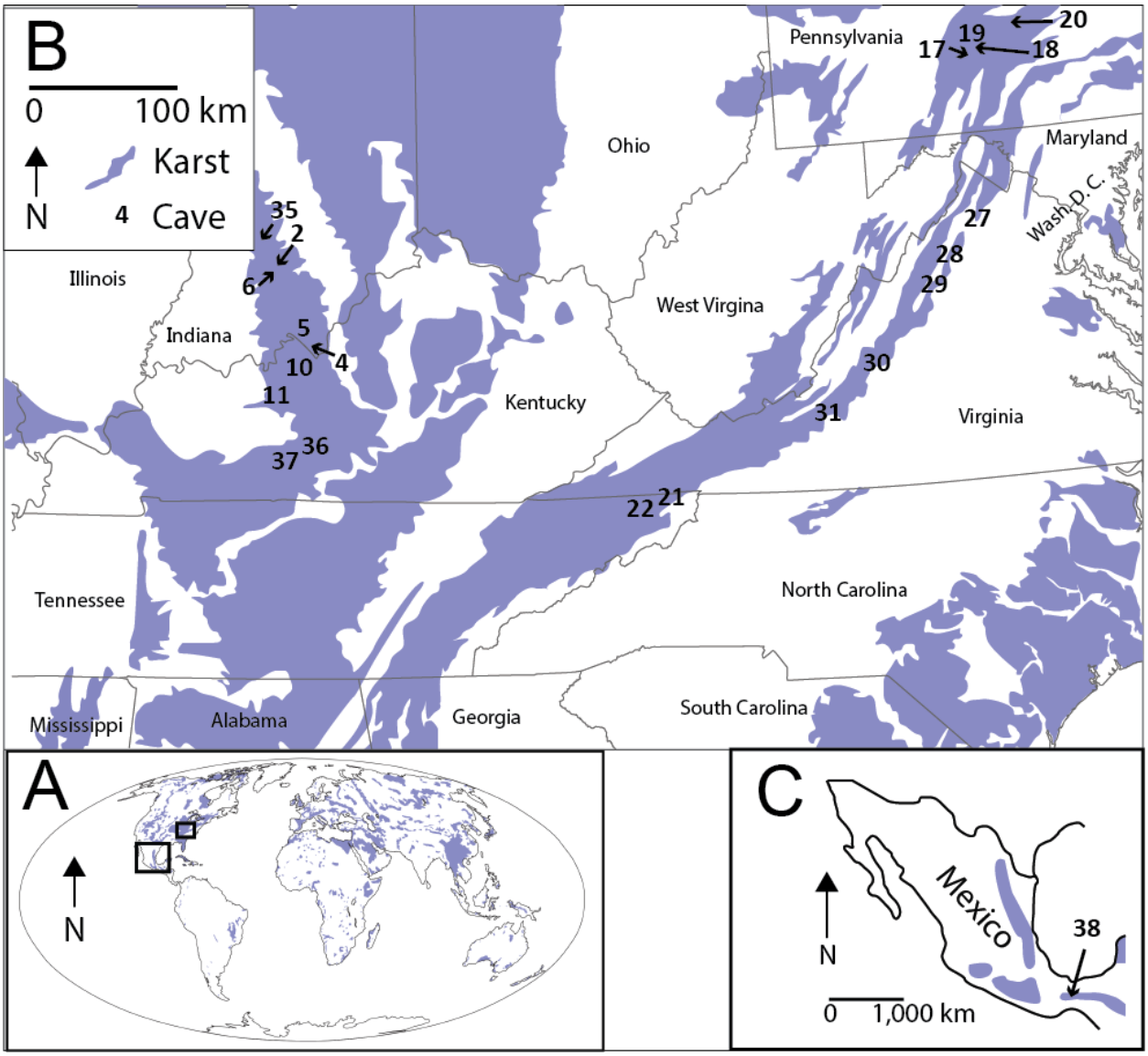
602 59. Shade A, Handelsman J. 2011. Beyond the Venn diagram: the hunt for a core
603 microbiome. Environ Microbiol 14:4–12.

604 60. Björk JR, O'Hara RB, Ribes M, Coma R, Montoya JM. 2017. The dynamic core
605 microbiome: structure, dynamics and stability. bioRxiv 137885.

606 61. Hernandez-Agreda A, Gates RD, Ainsworth TD. 2017. Defining the core
607 microbiome in corals'microbial soup. Trends Microbiol 25:125–140.

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

610



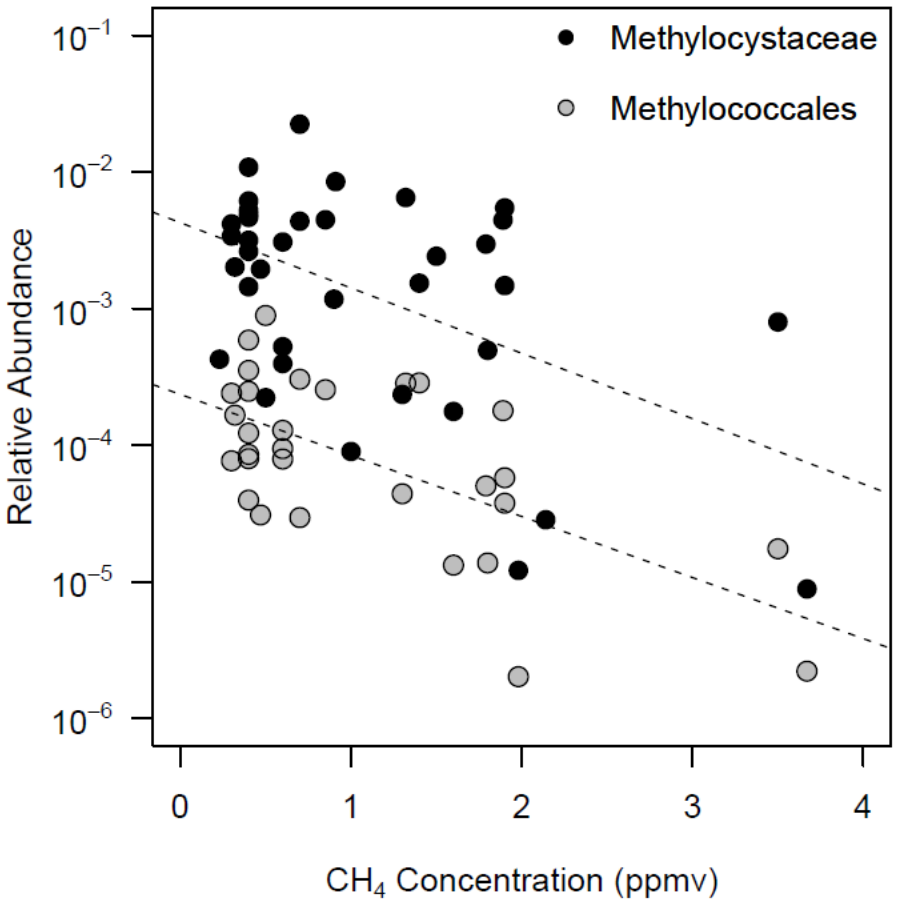
611

612 **Figure 1:** (A) Locations of karst at the global scale. Inserts show the occurrence of karst

613 (B) in the eastern United States (C) and in Mexico. Numbers represent the locations of

614 sampled caves in this study. Karst land cover data were obtained from (7, 62).

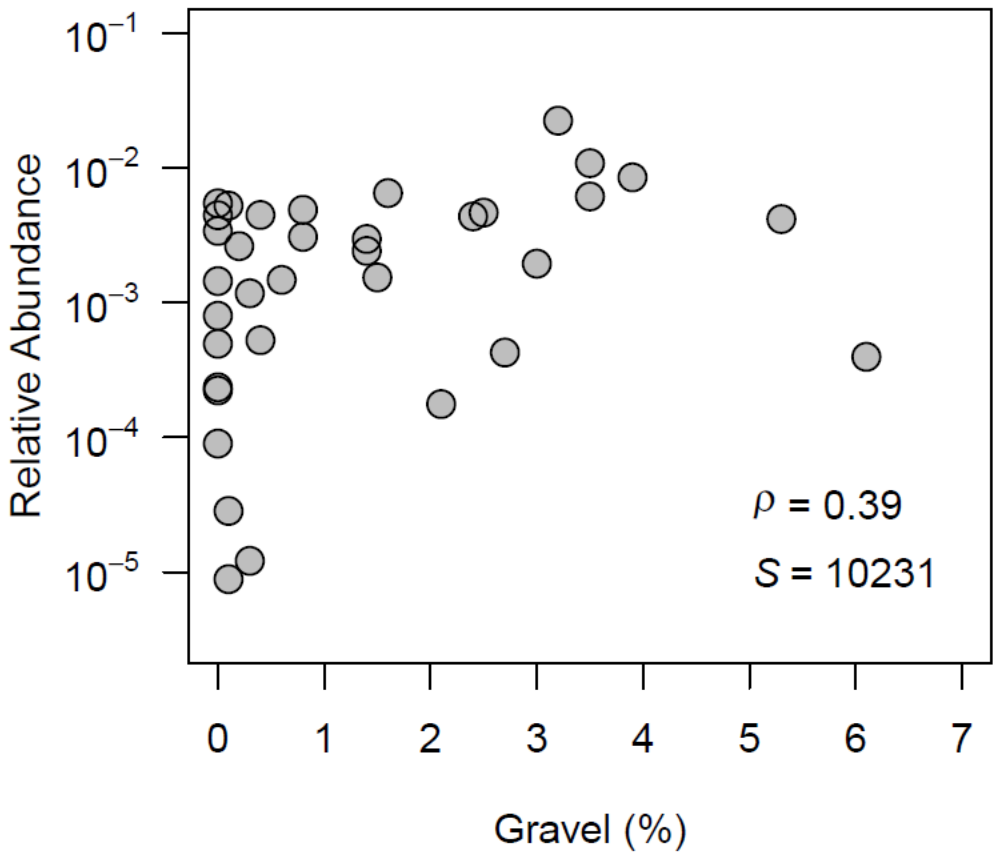
615



616

617 **Figure 2:** Relative abundances of methanotrophic community members plotted against the
618 CH_4 concentration at each sample location. Members of the Methylocystaceae, typically
619 associated with high-affinity methanotrophs, were more abundant than members from the
620 Methylococcales, typically associated with low-affinity methanotrophs.

621



622

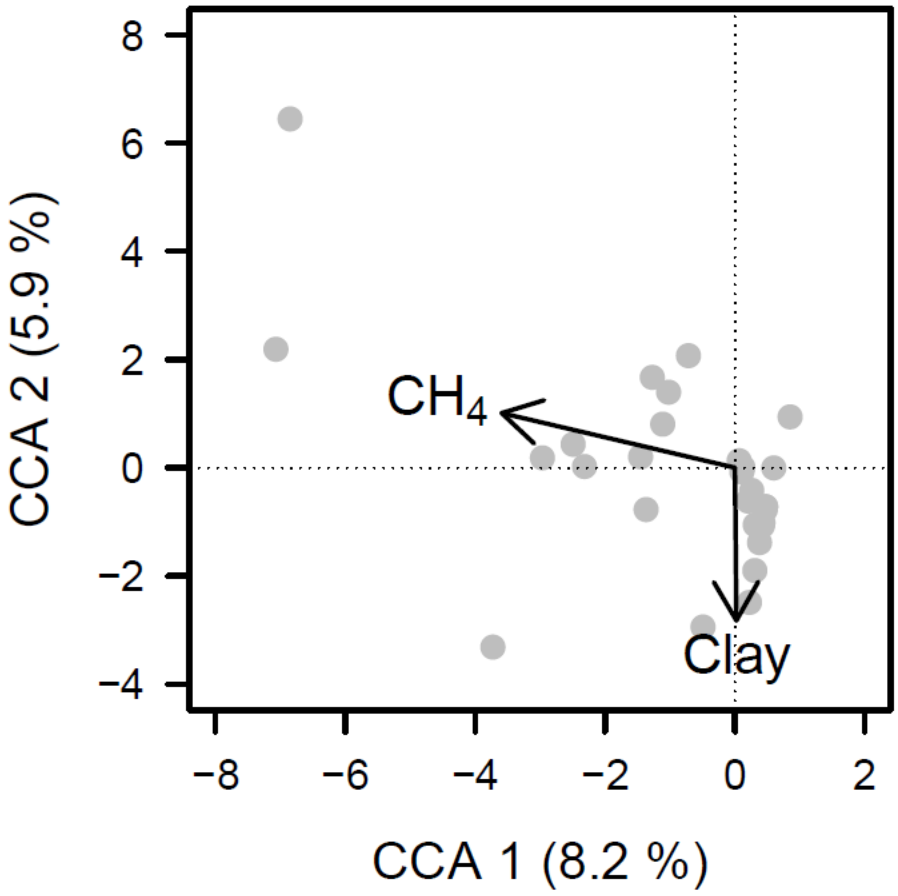
623 **Figure 3.** The relative abundance of methanotrophs from the Methylocystaceae plotted

624 against the volumetric proportion of gravel in a sample (Spearman's rank correlation, p =

625 0.02).

626

627



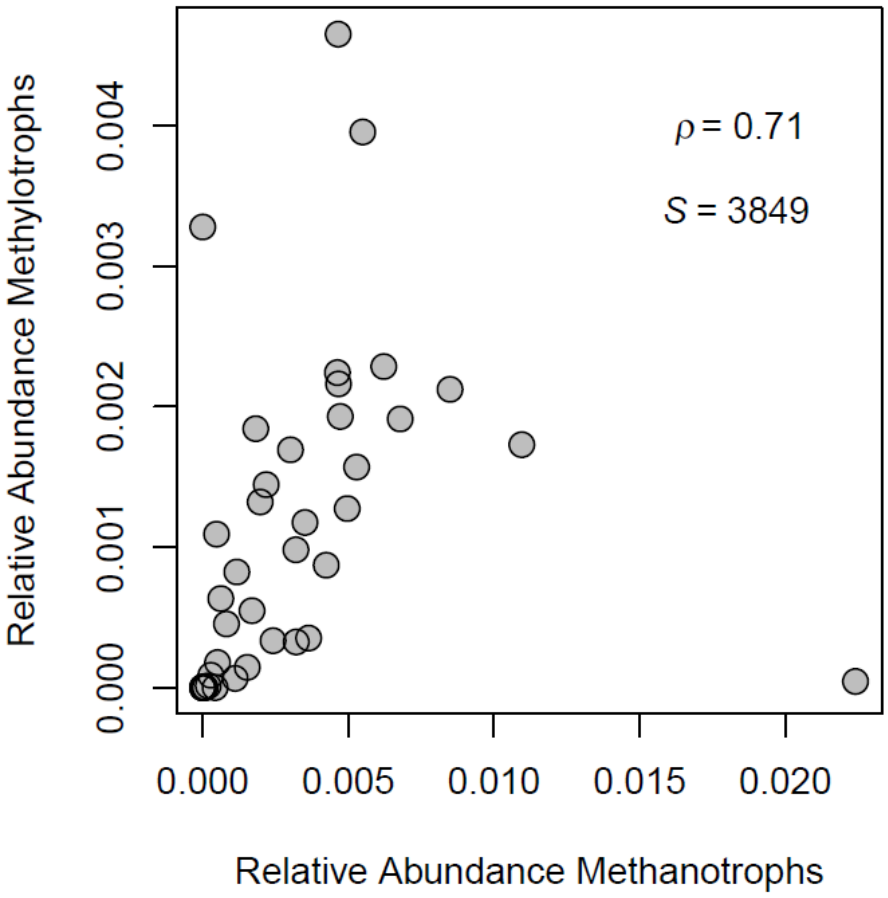
628

629 **Figure 4:** Canonical correspondence analysis of the community structure of the

630 Methylococcales in caves. Community structure of this group was significantly related to

631 clay content and CH_4 concentrations.

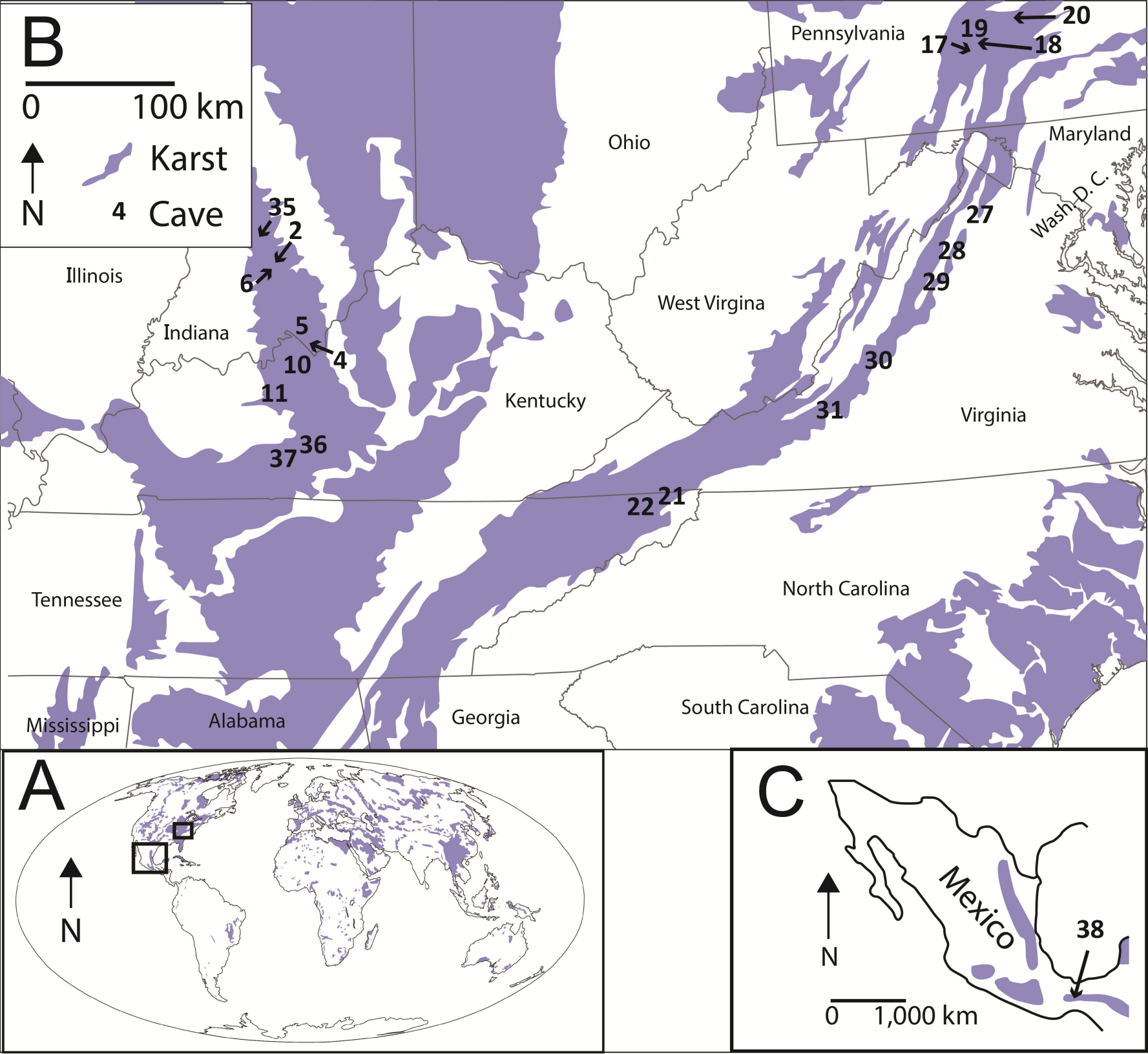
632

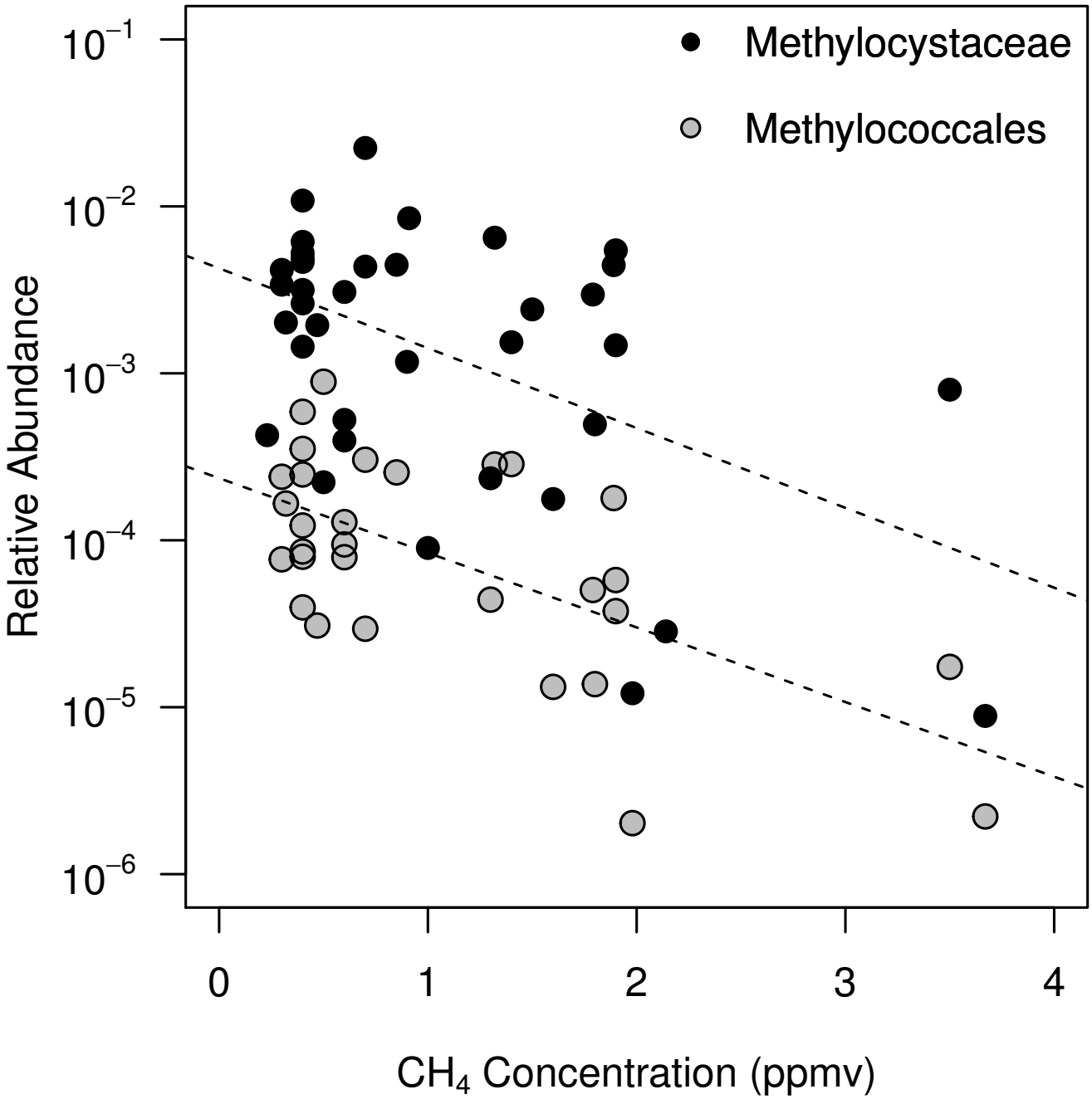


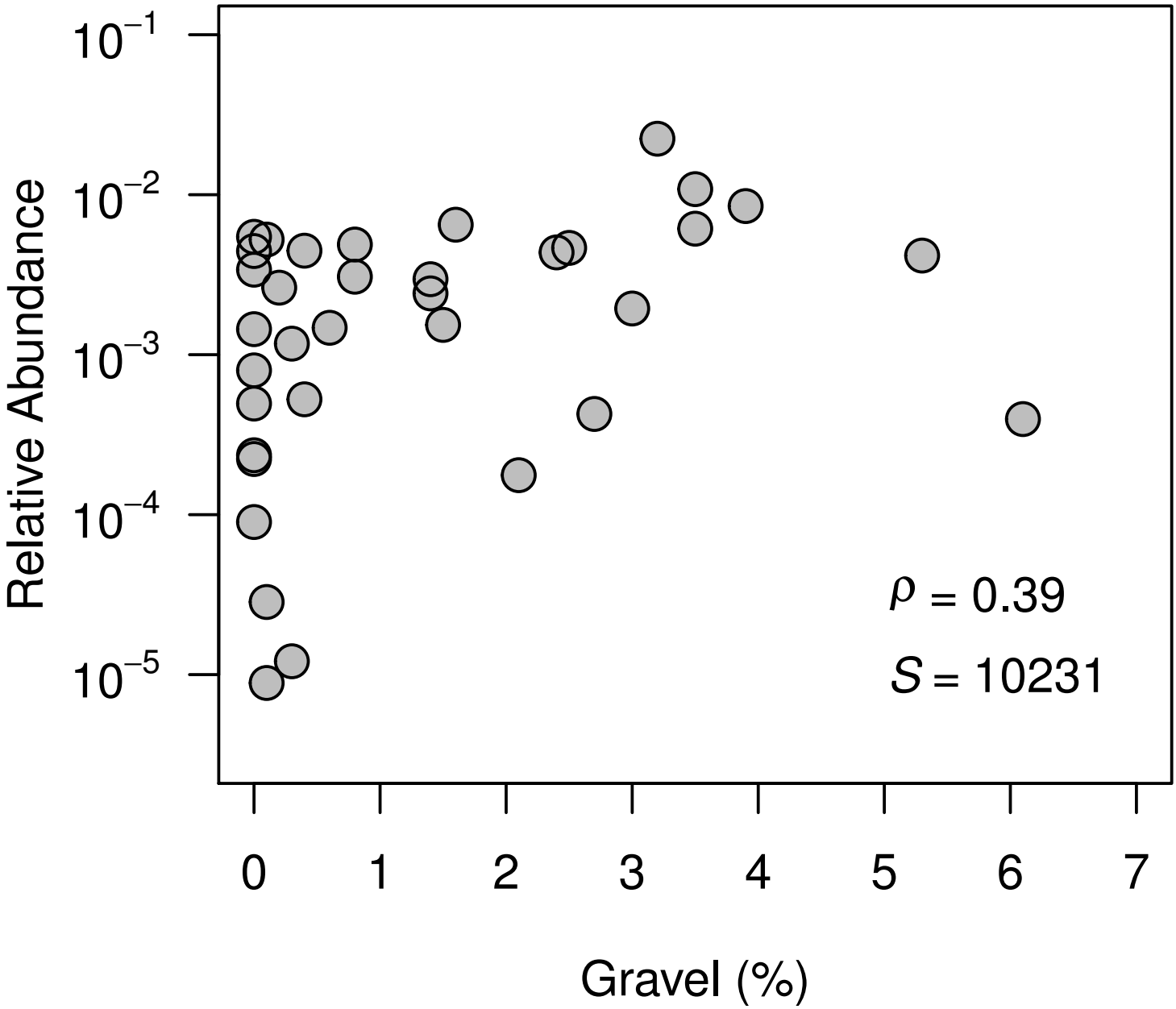
633

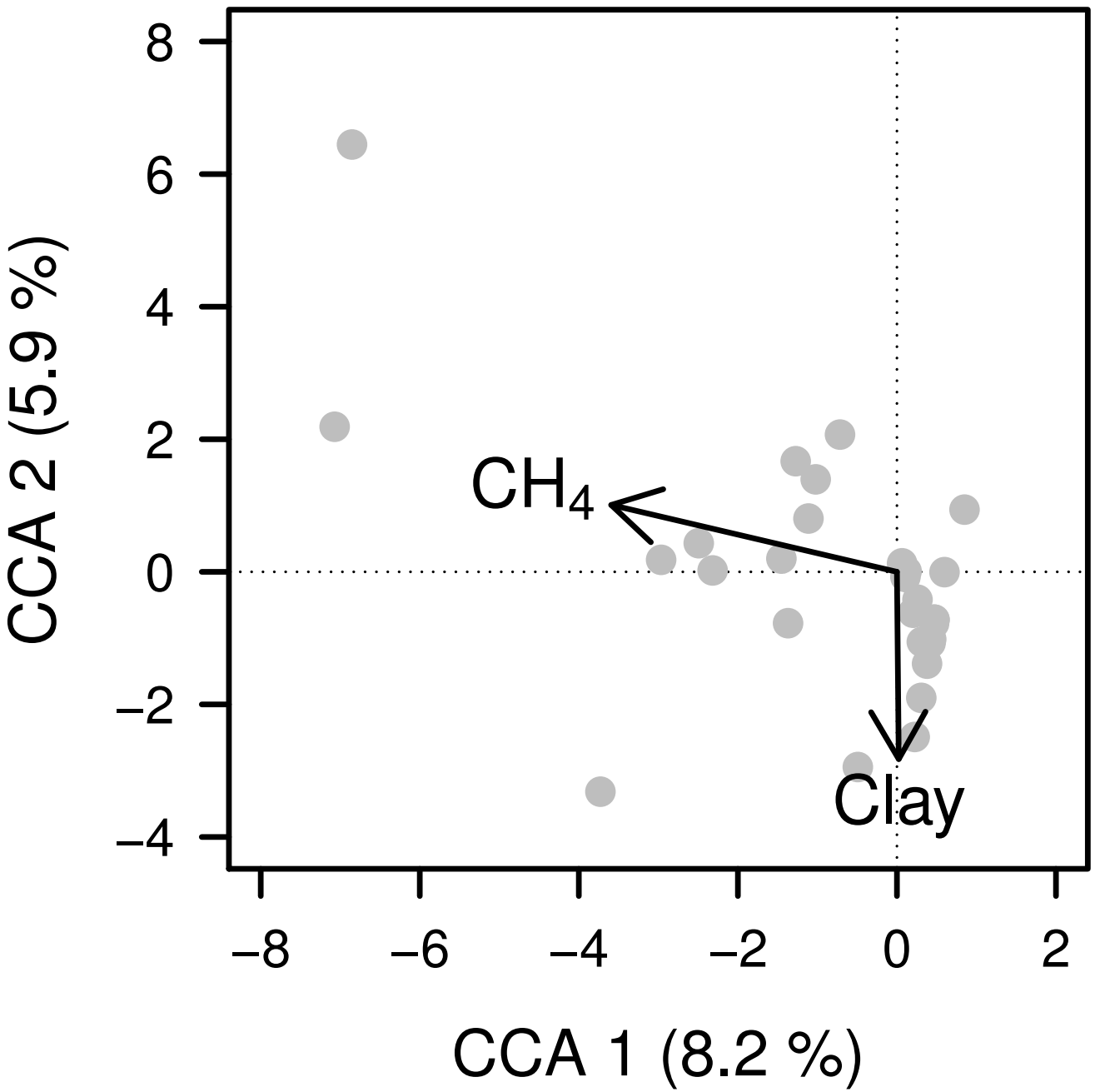
634 **Figure 5:** Relative abundance of methylotrophs plotted against the relative abundance of

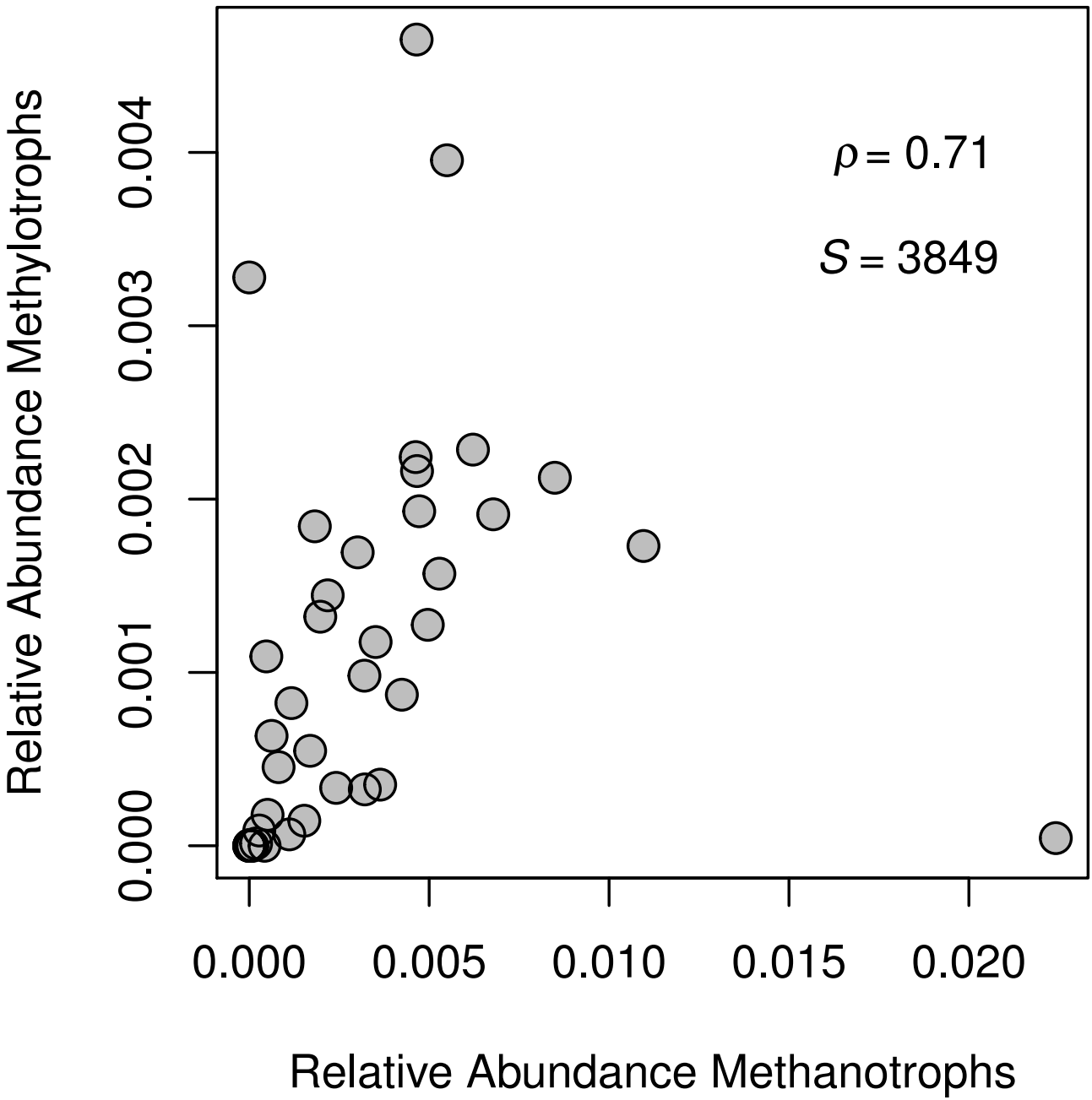
635 methanotrophs from sampled caves (Spearman's rank correlation, $p = 1*10^{-7}$).


B
0
100 km
N
Karst
4 Cave
Pennsylvania
Ohio
Maryland
Wash. D.C.
Illinois
Indiana
West Virgina
Kentucky
Virginia
Tennessee
North Carolina
South Carolina
Georgia
Mississippi
Alabama
35
2
6
5
4
10
11
37
36
22
21
17
19
20
18
27
28
29
30
31
A
N
C
N
Mexico
38
0
1,000 km


Methylocystaceae
Methylococcales
Relative Abundance
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
0
1
2
3
4
CH_4 Concentration (ppmv)


Relative Abundance
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
$\rho = 0.39$
$S = 10231$
0
1
2
3
4
5
6
7
Gravel (%)


8
6
4
2
0
−2
−4
−8
−6
−4
−2
0
2
CCA 2 (5.9 %)
CCA 1 (8.2 %)
CH4
Clay


Relative Abundance Methylotrophs
0.000
0.001
0.002
0.003
0.004
$\rho = 0.71$
$S = 3849$
0.000
0.005
0.010
0.015
0.020
Relative Abundance Methanotrophs