Fluctuations in populations of subsurface methane oxidizers in coordination with changes in electron acceptor availability

C. Magnabosco[1*], P.H.A. Timmers[2], M.C.Y. Lau[3], G. Borgonie[4], B. Linage-Alvarez[5], O. Kuloyo[5,6], R. Alleva[7], T.L. Kieft[8], G.F. Slater[9], E. van Heerden[5,10], B. Sherwood Lollar[11], T.C. Onstott[3]

[1]Flatiron Institute Center for Computational Biology, Simons Foundation, New York USA
[2]Department of Microbial Physiology, Laboratory of Microbiology, Wageningen University, Wageningen Netherlands
[3]Department of Geosciences, Princeton University, Princeton USA
[4]Extreme Life Isyensya, Gentbrugge Belgium
[5] Department of Microbial, Biochemical and Food Biotechnology, University of the Free State Bloemfontein, South Africa
[6]Department of Geoscience, University of Calgary, Calgary Canada
[7]Department of Ecology and Evolutionary Biology, Princeton University, Princeton USA
[8]Department of Biology, New Mexico Institute of Mining and Technology, Socorro USA
[9]School of Geography and Earth Sciences, McMaster University, Hamilton Canada
[10]Biosaense Solutions, Bloemfontein, South Africa
[11]Department of Earth Sciences, University of Toronto, Toronto Canada
*To whom correspondence should be addressed: Cara Magnabosco; 162 Fifth Ave, New York, NY 10012
cmagnabosco@flatironinstitute.org; +1 646-603-3683

Running Title: Subsurface ANME in South Africa

Abstract:

The concentrations of electron donors and acceptors in the terrestrial subsurface biosphere fluctuate due to migration and mixing of subsurface fluids, but the mechanisms and rates at which microbial communities respond to these changes are largely unknown. Subsurface microbial communities exhibit long cellular turnover times and are often considered relatively static—generating just enough ATP for cellular maintenance. Here, we investigated how subsurface populations of CH_4 oxidizers respond to changes in electron acceptor availability by monitoring the biological and geochemical composition in a 1,339 meters-below-land-surface (mbls) fluid-filled fracture over the course of both longer (2.5 year) and shorter (2-week) time scales. Using a combination of metagenomic, metatranscriptomic, and metaproteomic analyses, we observe that the CH_4 oxidizers within the subsurface microbial community change in coordination with electron acceptor availability over time. We then validate these findings through a series of ^{13}C-CH_4 laboratory incubation experiments, highlighting a connection between composition of subsurface CH_4 oxidizing communities and electron acceptor availability.

Introduction:

The terrestrial subsurface is an energy-limited environment that is subject to changes in fluid chemistry over time (Onstott *et al.* 2006). Laboratory experiments have shown that when native fluids are supplemented with electron acceptors such as SO_4^{2-}, the activity of subsurface communities can be enhanced (Rajala *et al.* 2015). Large disturbances such as CO_2 (Morozova *et al.* 2010, 2011) and H_2 injection (Bagnoud *et al.* 2016), hydraulic fracking (Daly *et al.* 2016), and drilling (Purkamo *et al.* 2013) have also

been reported to alter natural subsurface communities. The response of microbial communities to natural fluctuations in their environment, however, is less understood.

In the South African subsurface, increases in the availability of electron acceptors such as NO_3^- (>10-fold higher) and SO_4^{2-} (>5-fold higher) over a 1.5 year period did not correspond to changes in the bacterial community (Magnabosco *et al.* 2014; Simkus *et al.* 2015). However, 16S SSU rRNA gene amplicon surveys of the archaeal communities from the same site and sampling points of the aforementioned studies showed a diverse collection of anaerobic methane oxidizing archaea (ANME) (Young *et al.* 2017) and methane oxidizing bacteria (Simkus *et al.* 2015).

ANME-1 and "*Candidatus* Methanoperedens nitroreducens", a member of ANME-2d, are among the best described ANME and couple the oxidation of CH_4 with SO_4^{2-} and NO_3^- reduction, respectively (Haroon *et al.* 2013; Wegener *et al.* 2015). Other ANME-2d have also been reported to couple CH_4 oxidation to the reduction of Mn^{4+} and/or Fe^{3+} (Ettwig *et al.* 2016). With the exception of "*Ca.* Methylomirabilis oxyfera" which has been suggested to generate intracellular O_2 for CH_4 oxidation from two molecules of NO (Ettwig *et al.* 2010), bacterial methanotrophs couple CH_4 oxidation with free O_2 in the environment. This potential relationship between CH_4 oxidizers (MOs) and electron acceptor availability provides a compelling avenue to explore the response of subsurface communities to natural changes in subsurface fluid chemistry.

Our study focuses on the subsurface microbial community of a 1,339 meters-below-land-surface (mbls) fluid-filled fracture (Be326 Bh2). Here, bulk bacterial phospholipid-derived fatty acid (PLFA) isotopic signatures have been shown to be consistent with the accumulation of ^{13}C-dissolved inorganic carbon (DIC) impacted by the microbial oxidation of CH_4 (Simkus *et al.* 2015) but the organisms responsible for CH_4 oxidation have not been well characterized (Magnabosco *et al.* 2014; Simkus *et al.* 2015; Young *et al.* 2017). To better describe the membership of native MO communities, their methane oxidizing genes, and their response to changes in fluid chemistry over time, we combined metagenomics, metatranscriptomics, and metaproteomics analyses with geochemical monitoring of Be326 Bh2's *in situ* fracture fluids over both 2.5 year and 2 week timescales. We also performed activity assays on the fracture fluids using ^{13}C-labeled CH_4 together with different electron acceptors to track anaerobic methane oxidizing activity in both short- and long-term incubations.

Materials and Methods:

Be326 Bh2 was accessed through a 57-m, horizontally drilled borehole that was drilled in 2007 and sealed with a high-pressure steel valve. The borehole is located on the 26 level of shaft 3 in the Beatrix Gold Mine (28.232288 S, 26.794365 E; Welkom, South Africa). Annual samples were collected during field trips in 2011, 2012, and 2013 and weekly samples were collected during the 2013 field trip with three time points designated as T_0, T_1, and T_2 with T_0 corresponding to the first day of the 2013 study and T_2 corresponding to the last day of the 2013 study.

Sampling

Sampling procedures have been outlined previously (Magnabosco *et al.* 2014). To sample, a sterile (combusted and autoclaved) stainless steel manifold with attached valves was attached to the Be326 borehole. A high-pressure steel valve was opened, allowing water to flow freely at ~4-6 L min^{-1} for at least 10 minutes. This manifold allows for the

attachment of airtight Teflon tubing and filters for sampling and chemical analysis. For the 2011, 2012, and 2013 samples, pre-autoclaved 0.1-µm Memtrex NY filters (MNY-91-1-AAS; General Electric Co., Minnetonka, MN USA) were left on the borehole for 6, 15, and 14 days, respectively. For the T_0, T_1, and T_2 samples, 0.2-µm Memtrex NY Capsule (CMNY) filters (General Electric Co., Minnetonka, MN USA) were left on the borehole for 2 hours with a flow rate of 500 mL per minute (equivalent to approximately 60 L filtered per time point) in 2013. The total volumes of water filtered for each time point in the 2.5-year study were 4,875 L, 12,604 L, and 6,635 L in 2011, 2012 and 2013, respectively.

Unfiltered water samples for direct cell counts were collected on the first day of the 2011, 2012, and 2013 studies and fixed with sterile formaldehyde (final concentration 4% v/v). Cell counts were obtained at the University of the Free State where samples were filtered through a sterile 0.22µm Millipore GTTP-type membrane filter, stained with DAPI, and visualized using fluorescence microscopy.

Geochemical Measurements

Temperature, pH, and E_h were measured at the borehole using handheld probes (HANNA instruments, Woonsocket, RI). Gas samples (H_2, O_2, CH_4) were collected and analyzed by gas chromatography (Peak Performer 1 series, Peak Laboratories, USA) (Simkus *et al.* 2015). Oxygen was also measured at the borehole using a CHEMET kit (Chemetrics Inc.; Calverton, VA). NO_3^- and SO_4^{2-} were measured using an ion chromatograph coupled to an electrospray ionization-quadruple mass spectrometer (MS) (Dionex IC25 and Thermo Scientific MSQ, USA). $\delta^{18}O$ and $\delta^{2}H$ were measured as previously described (Simkus *et al.* 2015).

Preservation of Biomolecules and extraction

Filters were treated with an RNA preserve solution and stored in a -80°C freezer. The RNA preserve is a custom made solution of 20 mM EDTA, 0.3 M sodium citrate, and 4.3 M ammonium SO_4^{2-} (pH 5.2). The solution was autoclaved prior to sample application. Total protein, together with total DNA and RNA, was extracted using 2X CTAB lysis buffer and phenol/chloroform (pH=6.5-6.9), and re-suspended in 1 TE-buffer (Tris-EDTA, pH = 8) and stored in 1.5-mL Eppendorf tubes at −20°C. Extraction of biomolecules is further described elsewhere (Lau *et al.* 2014, 2016).

Amplicon Sequencing and Annotation

The V6 region of archaeal 16S rDNA molecules from the 2011 and 2012 time points was amplified using 958F (AATTGGANTCAACGCCGG) and 1048R (CGRCRGCCATGYACCWC) primers. The V4-V5 region of archaeal 16S rDNA molecules from all time points was amplified using the 517F (GCCTAAAGCATCCGTAGC; GCCTAAARCGTYCGTAGC; GTCTAAAGGGTCYGTAGC; GCTTAAAGNGTYCGTAGC; GTCTAAARCGYYCGTAGC) and 958R (CCGGCGTTGANTCCAATT) primers. For both amplicon datasets, forward primers included 5-nt multiplexing barcodes and a reverse 6-nt index. Triplicate PCR amplifications were performed in 33-µL reaction volumes with an amplification cocktail containing 10.0 U Platinum Taq Hi-Fidelity Polymerase (Invitrogen, Carlsbad, CA), 1X Hi-Fidelity buffer, 200 µM dNTP PurePeak DNA polymerase mix (ThermoFisher), 2 mM $MgSO_4$ and 0.3 µM of each primer. We added approximately 10-25 ng template DNA to each PCR and ran a control without template DNA for each primer pair. Amplification conditions were: initial 94°C, 3

minute denaturation step; 25 cycles of 94°C for 30 s, 60°C for 45 s, and 72°C for 60 s; and a final 2 minute extension at 72°C. The triplicate PCR reactions were pooled after amplification and purified using Qiagen MinElute plates followed by clean up, PicoGreen quantitation and Sage PippinPrep size selection. 101-nt paired-end sequencing was performed on an Illumina HiSeq 1000 at the Marine Biological Laboratory (Woods Hole, MA USA). Only reads that were identical in the overlapping regions of the forward and reverse reads were included for annotation. In the case of V6 data, this filter necessitates an exact match across both the forward and reverse read. Sequences that met the given quality criteria were annotated using GAST (Huse *et al.* 2008) and a GAST-formatted reference set.

Mapping of Metatranscriptomic Data to V6 Amplicons

Quality filtered RNA reads were mapped to the GAST-annotated V6 amplicons of Be326 Bh2 (Magnabosco *et al.* 2014; Young *et al.* 2017) using BLASTn (percent identity >97%; alignment length > 55 nucleotides (nt)). Reads that mapped positively to V6 sequences were assigned a consensus taxonomy based on the top three BLASTn hits.

Generation of Metagenomic and Metatranscriptomic Datasets

DNA samples from 2011 and 2012 were sequenced at the National Center for Genome Resources (Santa Fe, NM). The 2011 and 2012 metagenomic libraries were prepared from 500 ng of DNA using the KAPA High Throughput Library Preparation Kit (KAPA Biosystems), an insert size of approximately 280 bp, and followed by 8 PCR cycles. Paired-end sequencing (2 × 100 nt) was performed on an Illumina HiSeq 2000. DNA from 2013 was sequenced at Lewis Sigler Institute for Integrative Genomics (Princeton, NJ USA) using an Illumina HiSeq 2500. The metagenomic library was prepared using the TruSeq Rapid SBS Kit, size selected for 380-400 nt, and sequenced using paired-end sequencing (2 × 215 nt).

RNA samples from 2011, 2012, 2013, T_0, T_1, and T_2 were sequenced at Lewis Sigler Institute for Integrative Genomics (Princeton, NJ USA) on an Illumina HiSeq 2500 platform. Here, metatranscriptomic libraries were prepared using the Ovation RNA-Seq v2 System (NuGEN; San Carlos, CA USA), which involved 15-18 cycles of PCR. The 2011 and 2012 RNA samples were sequenced using 141-nt single-end sequencing while the 2013, T_0, T_1, and T_2 RNA samples were sequenced using 200-nt single-end reads.

Targeted Assembly and Tree Building

Reads were quality filtered and assembled using a targeted assembly pipeline (https://github.com/cmagnabosco/OmicPipelines). In summary, this pipeline involved quality filtering reads using default settings on the fastx toolkit (http://hannonlab.cshl.edu/fastx_toolkit/), targeted assembly of *mcrA* GENEs, protein prediction, alignment of reads to a set of well curated McrA peptides, and the construction of a phylogenic tree from the predicted proteins (PhyML (Guindon *et al.* 2010), Best of nearest-neighbor-interchange and subtree-pruning-regrafting, 8 rate categories, -b 100). This procedure was also repeated for the assembly of *mmo* genes using a MMO peptide database.

Assembled genes of 200 nt or longer were trimmed to include only the coding region of verified *mcrA* and *mmo* genes. This dataset was used as the reference database (Supplementary Data 2) to map quality-filtered DNA and RNA reads using Bowtie2 (--very-sensitive-local). The coverage was calculated as: (number of reads mapped × average read length)/length of the reference sequence. For correlation analyses, coverage

was normalized by dividing by the total number of reads in the dataset. Pearson correlation coefficients (*r*) were computed in Excel using the CORREL function.

Protein Identification from UPLC-MS/MS Data

Ultra performance liquid chromatography-tandem mass spectrometer (UPLC-MS/MS) data were analyzed in aggregate using the SEQUEST HT search engine in ProteomeDiscoverer v1.4 (ThermoFisher Scientific) using a custom database containing the translated genes from the targeted assembly and published McrA, pMoA, MmoX and AmoA (Supplementary Data 4). Search parameters included trypsin digestion with up to one missed cleavage, methionine oxidation and cysteine carbamidomethylation. A peptide-level false discovery rate (FDR) of 5% was achieved by using the Percolator node in ProteomeDiscoverer, which utilizes the frequency of matching against a reversed database as a rigorous model of the probability of error in the forward matches at given score thresholds. Proteins identified by matches to one unique peptide per run were tallied.

Activity assays

During the 2012 and 2013 sampling points, water samples were collected into 150-mL serum vials ("biovials") and stored at 4°C. Two days prior to sampling, 100 µL of MilliQ water was added to a 150-mL serum vial that had undergone combustion in an oven at 450°C overnight to deactivate spores. The serum vial was then sealed with a butyl rubber stopper that was cleaned by boiling in 0.1 M NaOH solution for 1 hour, rinsed and left to soak in MilliQ water until use. The serum vials were then crimped and left overnight. The next day, the serum vials were purged and pressurized with filtered N_2. The vials were then autoclaved and bubble-wrapped for transport underground. A needle was attached to the end of sterile teflon tubing that was attached to the manifold. Water was allowed to flow to rinse the attachment and needle. After 2 minutes of rinsing, the needle was inserted into the biovial and a second needle was added for pressure relief. Water was allowed to flow into the biovial until the relief needle overflowed. Extensive care was taken to ensure that no visible gas bubble remained in the biovial post-sampling.

Two sets of enrichment experiments were performed to monitor the response of the biovial communities to methane and a variety of electron acceptors. Experiment 1 contained samples from 2012 and 2013. Samples were incubated in triplicate with ^{13}C-CH_4 and SO_4^{2-} (20 mM) or ^{13}C-CH_4 and NO_3^- (20 mM) for 207 and 185 days, respectively. A control incubation with ^{13}C-CH_4 and no electron acceptors was monitored for 207 days to measure endogenous $^{13}CO_2$ production (Control A). Experiment 2 included samples from 2012 and 2013 that were incubated in triplicate with ^{13}C-CH_4 and SO_4^{2-} (20 mM), NO_3^- (20 mM) or O_2 (5% v/v). A killed (paraformaldehyde, 4% v/v) control without electron acceptors (Control 1) and a live control without electron acceptor or donor were included (Control 2). All samples of Experiment 2 were incubated and monitored over a 43-day period.

The samples for all experiments were prepared in 14-mL serum vials sealed with butyl rubber stoppers and aluminum caps. Prior to sample addition, vials were made anoxic by exchanging headspace gas with N_2 for 10 cycles and left pressurized (0.5 bar overpressure). 10 mL aliquots of fracture fluid were then added to the vials in an anoxic chamber and amended with the treatments described above. When ^{13}C-CH_4 was added, N_2 was added to a pressure of 130 kPa and 99.99% ^{13}C-CH_4 gas (Campro Scientific, Veenendaal, The Netherlands) was added to a final pressure of 180 kPa. Oxygen was

added afterwards, when applicable. All electron acceptor solutions were sterile and anoxic. The serum bottles were statically incubated at 37°C in the dark.

NO_3^- and SO_4^{2-} were analysed using an ion chromatography system equipped with an Ionpac AS9-SC column and an ED 40 electrochemical detector (Dionex, Sunnyvale, CA). The system was operated at a column temperature of 35°C, and a flow rate of 1.2 ml min^{-1}. Eluent consisted of a carbonate/bicarbonate solution (1.8 and 1.7 mM respectively) in deionized water. Headspace gas composition was measured on a gas chromatograph-mass spectrometer (GC-MS) composed of a Trace GC Ultra (Thermo Fisher Scientific, Waltham, MA) equipped with a Rt-QPLOT column (Restek, Bellefonte, PA), and a DSQ MS (Thermo Fisher Scientific). Helium was used as a carrier gas with a flow rate of 120 ml min^{-1} and a split ratio of 60. The inlet temperature was 80°C; the column temperature was set at 40°C and the ion source temperature was 200°C. CH_4 and CO_2 in the headspace were quantified from the peak areas in the gas chromatographs. The fractions of $^{13}CO_2$, $^{13}CH_4$ and $^{12}CH_4$ were derived from the mass spectrum (Shigematsu *et al.* 2004). Validation of the method was done using standards with known mixture of $^{13}CO_2$ and $^{12}CO_2$. The concentrations of total CO_2, total CH_4, and $^{13}CO_2$ were calculated following the method of Timmers *et al.* (2015). The pressure of the serum vials was determined using a portable membrane pressure unit (GMH 3150, Greisinger electronic GmbH, Regenstauf, Germany). The pH was checked by a standard pH electrode (QiS, Oosterhout, The Netherlands).

Results and Discussion:

A changing fluid chemistry over time

Over a period of 2.5 years, water isotope analysis revealed large changes in the fracture fluid's δ^2H and $\delta^{18}O$ isotopic signatures (Fig. 1; purple squares). These changes are inconsistent with contamination with service water in the mine and, instead, indicate mixing of different fracture waters within the system. In 2011, the $\delta^{18}O$ and δ^2H values were on the global meteoric water line (GMWL) and are indicative of paleometeoric water. This isotopic signature is consistent with other fluids located approximately 1,000 to 1,500 mbls in the Witwatersrand Formation (South Africa) (Fig. 1; green triangles) (Onstott *et al.* 2006; Sherwood Lollar *et al.* 2007). In 2012 and 2013, the $\delta^{18}O$ and δ^2H signatures of the fluids moved away from the GMWL, indicating mixing with more ancient fluids (Frape *et al.* 1984). A few meters away, Be326 Bh1 (a borehole located in the same mine and depth as the Be326 borehole of this study) exhibited a similar trend in $\delta^{18}O$ and δ^2H signatures over time—shifting away from the GMWL in 2012 and 2013 (Fig. 1; red diamonds). The displacements between the 2011/2012 and 2012/2013 $\delta^{18}O$ and δ^2H signatures of the two boreholes are in opposite directions—a pattern that would not be expected if the native fracture fluids were mixing with the mine's service water during this period. Measurements of the fracture fluid's δ^2H and $\delta^{18}O$ isotopic signatures over the 2-week study were not obtained.

Geophysico-chemical measurements were made for both the 2.5-year and 2-week time series (Table 1). Temperature, pH, and CH_4 concentrations were relatively unchanged in both of the time series but the degree to which E_h, SO_4^{2-}, NO_3^-, and H_2 concentrations changed was much greater over the 2.5 years (Table 1a). Within the 2.5-year time series, the fracture fluids shifted from a more reduced state (2011_{Eh} = -98 mV, $2011_{[H2]}$ = 130 nM) with limited electron acceptor availability ($2011_{[Sulph.]}$ = 137µM,

$2011_{[Nitr.]}$ = 0.4 μM) to a more oxidized state (2013_{Eh} = 21±28 mV, $2013_{[H2]}$ = 25 nM) with much greater electron acceptor availability ($2013_{[Sulph.]}$ = 479 μM, $2013_{[Nitr.]}$ = 4.5 μM). The 2-week time series did not exhibit as large of a shift in electron acceptor availability as the 2.5-year time series and maintained a positive E_h throughout (Table 1b).

Microbial community of Be326

The microbial communities of the 2011 and 2012 time points have been reported to be dominated by bacteria (98.5%) (Simkus *et al.* 2015) with the majority being related to Proteobacteria (Magnabosco *et al.* 2014). In order to investigate the community composition of the less numerous archaea, 16S rDNA amplicon sequencing of the archaeal V4-V5 hypervariable region was performed across all time points (Fig. 2a). For the long-term study (2011, 2012, 2013), the archaeal community shifted from an ANME-1- and Methanomicrobia-dominated community in 2011 to a Miscellaneous Crenarchaeotic Group-dominated community in 2012 and a Halobacteria-dominated community in 2013. For the short-term study (T_0, T_1, T_2), there were no noticeable changes between T_1 and T_2. A slight increase in Halobacteria and a decrease in ANME-1 with respect to T_1 and T_2 were observed in the T_0 time point. There is a noticeable difference between the T_0, T_1, T_2 samples community profiles and the 2013 community profile, despite being collected over the same two-week period. However, the filters used in the T_x and 2013 filtrations have varied pore sizes, geometries, and casings and should not be directly compared.

Around 1% of V4-V5 amplicons were related to "*Ca.* M. nitroreducens" in all time points except for the 2011 dataset that contained only 2 archaeal amplicons relating to "*Ca.* M. nitroreducens" (Supplementary Data 1). These relative abundances of "*Ca.* M. nitroreducens" are lower than what was reported using archaeal V6 primers on 2011 and 2012 samples. With archaeal V6 primers, "*Ca.* M. nitroreducens" accounts for 1.5% of the archaeal community in 2011 and 10.6% of the archaeal community in 2012, while ANME-1 accounts for ~10% of the archaeal community at each time point (Young *et al.*, 2017; Supplementary Fig. S1). Despite differences in the relative abundances of taxa based on the primers used, community membership does not appear to be significantly different between the 2 primer sets.

In order to estimate the relative activity of each taxonomic group, metatranscriptomic data from each sample were mapped to a database of Be326 Bh2 16S rDNA V6 bacterial (Magnabosco *et al.* 2014) and archaeal (Young *et al.* 2017) sequences. The V6 sequences were selected over the V4-V5 sequences due to their shorter length and stringent quality filtering procedure (see Materials & Methods). Proteobacteria within the family Rhodocyclaceae dominated all of the RNA datasets except for the 2011 dataset that was dominated by Hydrogenophilaceae. The number of bacterial and archaeal species (unique hits within the 16S rDNA V6 database) observed in the metatransciptomic data ranged from 204 in 2013's T_0 time point to 414 in the 2012 sample (Table 2). MO archaea and bacteria account for only a small percentage of the V6 rRNA sequences identified in the metatranscriptomic data (0.3-3%). Notably, the number of species observed in each metatranscriptome's V6 rRNA pool was not correlated to the number of reads generated for each metatranscriptome ($R^2 < 0.01$); however, the number of species observed in each metatranscriptome is almost 10 times less than the number of

$OTU_{0.97}$ obtained from the bacterial 16S rDNA V6 dataset (2,478 in 2011 and 3,987 in 2012) (Magnabosco *et al.* 2014).

A detectable shift in the MO community over time

As the organisms responsible for CH_4 oxidation in Be326 were present at relatively low abundances, a targeted assembly pipeline (https://github.com/cmagnabosco/OmicPipelines) that employs the PRICE assembler (Ruby, Bellare and DeRisi 2013) was implemented to assemble methyl-coenzyme M reductase (*mcrA*)—the gene for the first step in the anaerobic oxidation of methane (AOM) (Thauer 2011) or the last step in methanogenesis—and a suite of CH_4 monooxygenases (*mmo*) that are known to play a role in aerobic oxidation of methane (McDonald *et al.* 2008) from the metagenomic and metratranscriptomic datasets. Notably, *mcrA* was selected as an indicator for ANME presence because its phylogeny is congruent with MO phylogeny.

Following targeted assembly and annotation, two complete *mcrA* genes related to ANME-1 and "*Ca.* M. nitroreducens", an ANME-2d, were assembled. Only one complete *mmo* gene closely related to *Methylococcus capsulatus* was recovered from the metagenomic and metatranscriptomic data (Fig. 2b, Supplementary Data 2). Partial *mcrA* related to Methanomicrobia and Methanobacteria were also identified in the high-throughput data (Fig. 2b, Supplementary Data 2), but partially assembled *mmo*-related genes were omitted in downstream analyses due to the difficulty in distinguishing *mmo* from homologous ammonia monooxygenases genes (Holmes *et al.* 1995).

Assembled *mcrA* and *mmo* were translated into peptide sequences (Supplementary Data 3). Notably, the Be326 "*Ca.* M. nitroreducens"-type McrA was 99% identical to the McrA of the reference "*Ca.* M. nitroreducens" genome (Supplementary Data 3, 4). Metaproteomic data were searched against the collection of assembled McrA and MMO and a database of known McrA and MMO peptide sequences (Supplementary Data 4) to confirm that the transcribed genes were translated into proteins. The predicted amino acid sequences (Supplementary Data 3) from the assembled ANME-1, "*Ca.* M. nitroreducens", and *Methylococcus* genes were all identified within the metaproteomic data (Fig. 3, Supplementary Data 5) which further confirmed the presence and activity of these groups of organisms.

The abundances (based on coverage) of each MO gene (*mcr,A mmo*) and MO 16S SSU rRNA gene were calculated using Bowtie2 (Langmead and Salzberg 2012) (-very-sensitive-local) and BLASTn, respectively. These analyses provided evidence that the dominant members of the MO community in the metagenomes and metatranscriptomes shifted from ANME-1 to "*Ca.* M. nitroreducens" during the 2.5-year sampling period (Fig. 2) but remained constant during the 2-week sampling period (Supplementary Table 1). These observations were consistent with the relative changes in geochemistry over both time scales (Table 1). Notably, "*Ca.* M. nitroreducens" accounted for ≤1% of the archaeal community, as revealed using archaeal V4-V5 16S rDNA primers (Fig. 1), which is in contrast to the higher estimates of "*Ca.* M. nitroreducens" found when using archaeal V6 16S rDNA (Young *et al.*, 2017; Supplementary Fig. S1), metagenomic and metatranscriptomic data. This discrepancy suggests that "*Ca.* M. nitroreducens" sequences are recovered at a lower efficiency in archaeal V4-V5 16S SSU rRNA gene surveys relative to archaeal V6 16S SSU rRNA gene surveys, metagenomic and metatranscriptomic studies.

When metagenomic MO *mcrA* and *mmo* abundances of the long-term study (Supplementary Table 2) were correlated to geophysico-chemical measurements, "*Ca.* M. nitroreducens" was positively correlated to NO_3^- (R^2=0.99) and SO_4^{2-} (R^2=0.98) concentrations but negatively correlated to CH_4 (R^2=0.99) and H_2 (R^2=0.99) concentrations. ANME-1 *mcrA* abundances showed an opposite trend and were positively correlated to CH_4 (R^2=0.96) and H_2 (R^2=0.97) concentrations but negatively correlated NO_3^- (R^2=0.88) and SO_4^{2-} (R^2=0.85) concentrations (Table 3). Correlation of metatranscriptomic 16S rRNA and *mcrA* and *mmo* abundances to geophysico-chemical measurements exhibited similar trends (Supplementary Table 3) and are consistent with a transition from an ANME-1-dominated MO community to a "*Ca.* M. nitroreducens"-dominated community. As NO_3^--coupled CH_4 oxidation is more energetic than SO_4^{2-}-coupled CH_4 oxidation (Caldwell *et al.* 2008), an energetic advantage, presumably, provides "*Ca.* M. nitroreducens" a competitive advantage against ANME-1 when NO_3^- concentrations are sufficient.

O_2 concentrations in 2011 and 2013 were below detection limit (Table 1) but were detectable in 2012 (0.47 μM). Likely related to the increased availability of O_2, aerobic *Methylococcus*-related *mmo* genes exhibited their highest relative abundances within metagenomic and metatranscriptomic MO gene profiles during 2012 and a minimal presence throughout the remainder of the time points (Fig. 2, Supplementary Table 1). "*Ca.* M. nitroreducens" was the dominant member (73.2±2.8%) of the MOs community during the 2-week time series (Supplementary Table 1) when fracture fluids contained high concentrations of SO_4^{2-} (496±36 μM) and NO_3^- (4.8±0.9 μM) along with a positive E_h (21±28 mV).

The correlations between MO abundances and fluid chemistry suggest that a relationship between electron acceptor availability and populations of MOs exists. We therefore wanted to experimentally validate the response of the MOs to changes in electron acceptor availability. As ANME-1, "*Ca.* M. nitroreducens", and *Methylococcus* are best described as a SO_4^{2-}-dependent ANME (Wegener *et al.* 2015), NO_3^--dependent ANME (Haroon *et al.* 2013), and aerobic methanotrophs (Kleiveland *et al.* 2012), respectively, we designed experiments to test whether or not each MO lifestyle would respond to an increase in the aforementioned electron acceptor.

Validation of the MO community through $^{13}CH_4$ enrichments

To better understand the response of the MO community to changes in electron donor/acceptor balance, two sets of ^{13}C-CH_4 laboratory enrichment experiments were performed on fracture water collected in 2012 and 2013. The first experiment (Experiment 1) was a long-term experiment analyzed over 207* days and contained fracture fluid samples from 2012 and 2013 enriched with either ^{13}C-CH_4 and no additional electron acceptor (endogenous activity control), ^{13}C-CH_4+SO_4^{2-} (to stimulate SO_4^{2-}-dependent AOM), or ^{13}C-CH_4+NO_3^- (to stimulate NO_3^--dependent AOM). A second set of 2012 and 2013 fracture fluid enrichments (Experiment 2) was analyzed for 43 days. Experiment 2 contained ^{13}C-CH_4 treatments of ^{13}C-CH_4+formaldehyde (4%, v/v) (killed control) to rule out non-biological sources of ^{13}C-CH_4 production, ^{13}C-CH_4+SO_4^{2-}, ^{13}C-CH_4+O_2 (to simulate aerobic methane oxidation), and ^{13}C-CH_4+NO_3^- as well as an electron acceptor- and donor-free control (methanogenesis control). The

* The long term ^{13}C-CH_4+NO_3^- enrichments were run for 183 days.

methanogenesis control was intended to detect whether or not methanogenesis, and therefore also trace CH_4 oxidation (TMO),was occurring within the samples (Zehnder and Brock 1979). Due to the limited amount of sample, we were unable to test the potential occurrence of Mn^{4+}-or Fe^{3+}-driven methane oxidation.

Unlike the correlations observed between expression data and geochemical parameters, an increase in the proportion of ^{13}C-CO_2 relative to total CO_2 ($\%^{13}C$-CO_2) in the $^{13}CH_4$ enrichments over time provides definitive evidence of $^{13}CH_4$ oxidation under different conditions. Notably, we chose to express our results as $\%^{13}C$-CO_2 rather than the absolute concentration (molar) of ^{13}C-CO_2 to account for CO_2 production from other substrates. In Experiment 1, the ^{13}C-CH_4+NO_3^- enrichments exhibited the greatest rate of $\%^{13}C$-CO_2 production and, in 2012, the rate of $\%^{13}C$-CO_2 production (0.017±0.005 $\%^{13}C$-CO_2 day^{-1}) was found to be significantly greater (paired one-tailed Student t-test; p=0.02) than in Control A (0.004±0.001 $\%^{13}C$-CO_2 day^{-1}) (Fig. 4). No samples from Experiment 2 exhibited an increase in $\%^{13}C$-CO_2 production (Supplementary Data 5).

Although ANME-1 were present in the metatranscriptomic and metaproteomic data during the 2012 and 2013 sampling points, there was not a significant difference in $\%^{13}C$-CO_2 production between Experiment 1's ^{13}C-CH_4+SO_4^{2-} incubations and the endogenous activity controls (Fig. 4). TMO was probably not responsible for the ^{13}C-CO_2 production in the endogenous activity controls of Experiment 1; only the methanogenesis control of the 2013 sample showed minor methanogenesis (and thus TMO) activity (Supplementary Data 5). It is conceivable, however, that SO_4^{2-}-dependent AOM occurred in both the endogenous controls and ^{13}C-CH_4+SO_4^{2-} incubations, as the concentrations of SO_4^{2-} in the controls ($[SO_4^{2-}{}_{2012}]$ = 623 μM; $[SO_4^{2-}{}_{2013}]$ = 479 μM) are well within the lower range of SO_4^{2-} concentrations (100-1200 μM) that have been reported for SO_4^{2-}-coupled AOM (Beal, Claire and House 2011; Segarra *et al.* 2015). Combined, these findings suggest that SO_4^{2-}-coupled AOM likely occurred in the controls of Experiment 1 and 2.

Conclusions:

Metagenomic, metatranscriptomic, and metaproteomic data suggest that community composition, activity, and function are changing in response to natural fluctuations in fluid chemistry. The observed CH_4 oxidation in the controls and dominance of ANME-1 in the 2011 samples (when SO_4^{2-} concentrations were lowest) indicate that the *in situ* fluids contain enough SO_4^{2-} to power SO_4^{2-}-coupled MO. The increase in $\%^{13}C$-CO_2 production in the ^{13}C-CH_4+NO_3^- enrichments and correlation of "*Ca.* Methanoperedens" abundances to electron acceptor concentrations *in situ* suggest that electron acceptor availability plays an important role in MO population dynamics. Together, these results provide the most conclusive biological evidence to date that CH_4 oxidation occurs and is an integral component of the deep terrestrial subsurface carbon cycle.

Data Availability:

Metagenomic and metatranscriptomic data are available at NCBI BioProject PRJNA308990. 16S amplicon data are available under NCBI BioProject PRJNA263371.

Tables and Figures:

TABLE 1a. Geophysico-chemical measurements for the long-term study

	2011	2012	2013
Dates Sampled	Jan 21-27, 2011	July 12-27, 2012	Aug 1-15, 2013
Temperature (°C)	36.9	37.3±1.2	35.1±0.8
pH	8.83	8.55	8.17±0.7
E_h (mV)	-98	-27.5±6.4	21±28
SO_4^{2-} (µM)	137	623	479
NO_3^- (µM)	0.4	6.0	4.5
O_2 (Chemet) (µM)	b.d.	0.47	b.d.*
H_2 (nM)	130	9	25
CH_4 (mM)	2.0	0.9	1.1

b.d. = below detection (<0.31 µM)
b.d.* = below detection (<1.6 µM)

Table 1b. Geophysico-chemical measurements for the short-term study

	T_0	T_1	T_2
Dates Sampled	Aug 1, 2013	Aug 8, 2013	Aug 15, 2013
Temperature (°C)	35.8	35.3	28.6
pH	8.2	8.1	7.9
E_h (mV)	9	53	1
SO_4^{2-} (µM)	531	499	459
NO_3^- (µM)	5.3	5.3	3.7
O_2 (Chemet) (µM)	b.d.*	b.d.*	b.d.*
H_2 (nM)	5.8×10^{-5}	10.4	39.8
CH_4 (mM)	0.8	1.2	1.0

b.d.* = below detection (<1.6 µM)

Table 2: Metatranscriptomic species richness observed and estimated by Chao 1 and ACE metrics and cell concentrations measured in the field.

	2011	2012	2013	T_0	T_1	T_2
Species (S) Observed	302	414	247	222	204	248
S.Chao1	344±22	440±10	256±7	233±9	216±8	276±16
S.ACE	317±8	450±10	253±6	233±5	212±7	260±8
Cell concentration ($\times 10^3$ cells mL^{-1})	1.3^-18.3*	2.2*-31^	3.2^	3.2^	n.m.	n.m.

^DAPI-derived cell concentration

*PLFA-derived cell concentration using a conversion factor of 6×10^4 cells/pmol PLFA as previously reported in Simkus et al. (2015)

n.m. = not measured

Table 3: Correlation (R^2) between normalized metagenomic *mcrA* or *mmo* PEG coverage and geochemical parameters. Positive correlations are underlined.

		NO_3^-	SO_4^{2-}	CH_4	E_h	pH	H_2
2.5-Year Time Series	"*Ca.* Methanoperedens"	<u>0.99</u>	<u>0.98</u>	0.99	<u>0.70</u>	0.52	0.99
	ANME-1	0.88	0.85	<u>0.93</u>	0.90	<u>0.76</u>	<u>0.95</u>
	Methylococcales	<u>0.12</u>	<u>0.14</u>	0.07	0.10	<u>0.25</u>	0.05

Figure 1. δ^2H_{H2O} and $\delta^{18}O_{H2O}$ for Beatrix (Be) Gold Mine fracture fluids.

This figure plots the $\delta^{18}O$ (x-axis) against the δ^2H (y-axis) for each of the three time points (purple squares) in the long-term study of Be326 (2011, 2012, 2013). The black diagonal line represents the global meteoric water line (Craig 1961) (GMWL). The 2011 sample sits on the GMWL, consistent with historic Be datapoints (green triangles) and paleometeoric water (Lippmann *et al.* 2003; Ward *et al.* 2004). The 2012 and 2013 time points are displaced away from the GMWL along trends that have been observed to be consistent with mixing with older, more saline fracture fluids (Ward *et al.* 2004; Sherwood Lollar *et al.* 2007). The dotted blue lines indicate possible mixing patterns. Be326_Bh1 (red diamonds) is a borehole located in the same mine and at the same depth as the borehole sampled in this study. EV=Evander Mine; KL=Kloof. Data for EV, KL, and Adams are adapted from Ward *et al.* (2004).

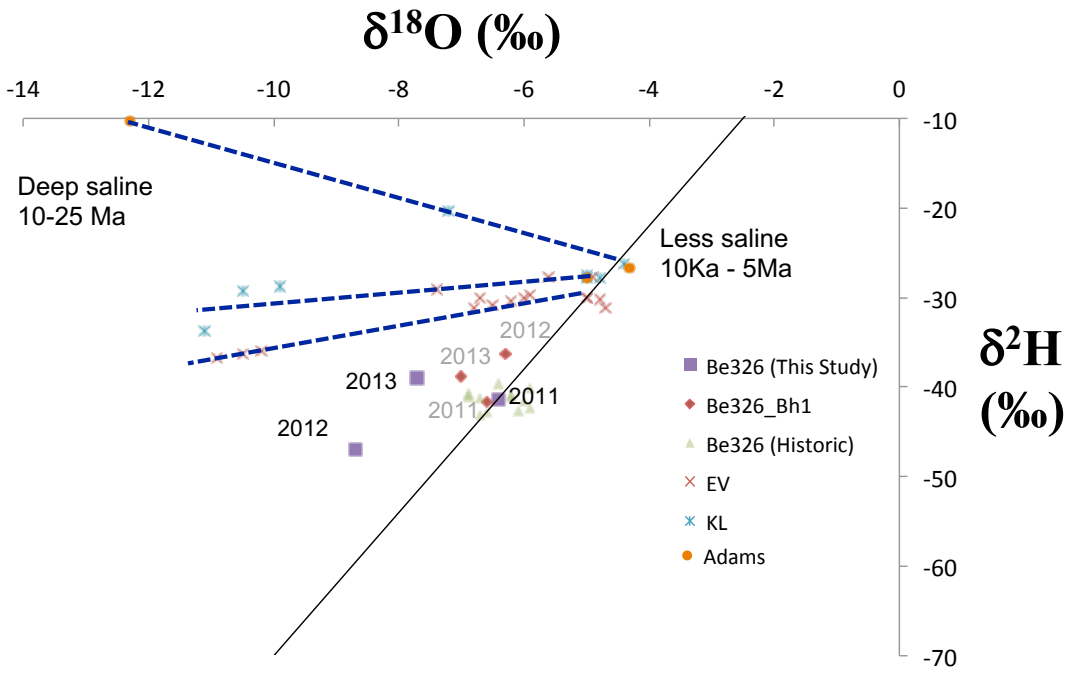


Figure 2. Change in the CH_4 oxidizing community and activity over time.
a) The relative abundance (bar height) of major archaeal groups (colors of bars) identified in the archaeal V4-V5 16S rDNA surveys are shown. b) A phylogenetic tree was constructed using PhyML from a McrA peptide database (Supplementary Data 6) and the predicted proteins from the *mcrA* targeted assembly (designated by a *). The predicted MMO from the targeted assembly is also displayed. The 2×3 blocks represent the relative abundance of each taxon (DNA, left column; RNA, right column) with respect to their CH_4-related protein of interest (McrA or MMO) over time (row 1, 2011; row 2, 2012; row 3; 2013). Notably, Methanobacteria-related McrA PEGs were selected as the outgroup in this presentation of the McrA phylogenetic tree due to Methanobacteria's placement in species tree of Euryarchaeota.

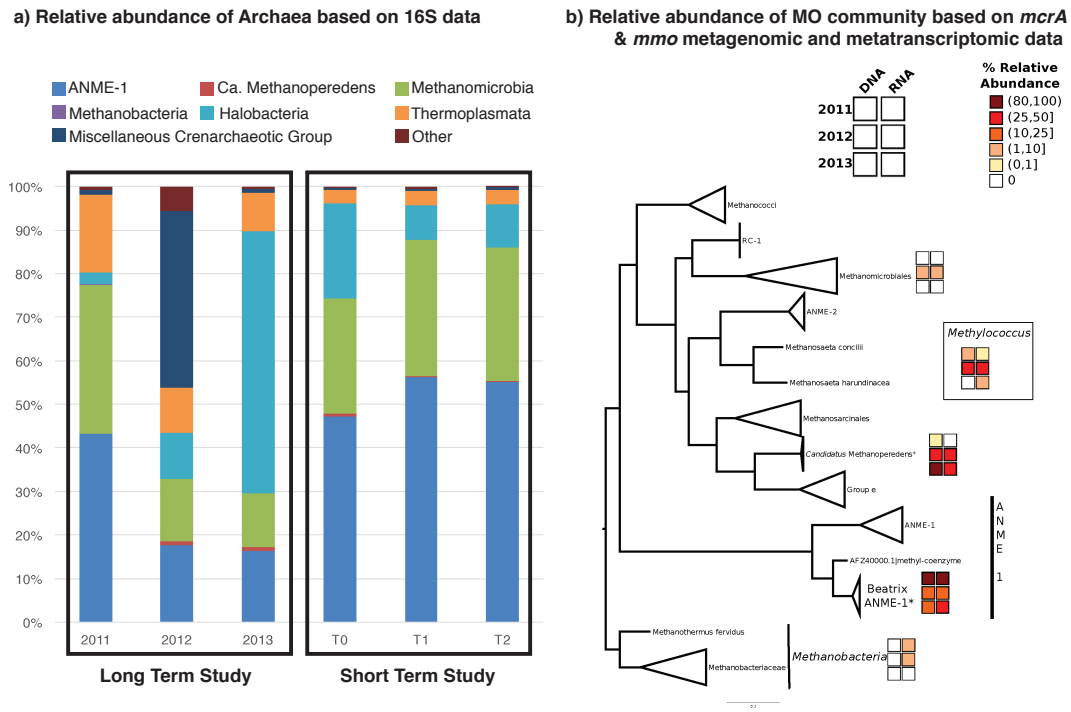
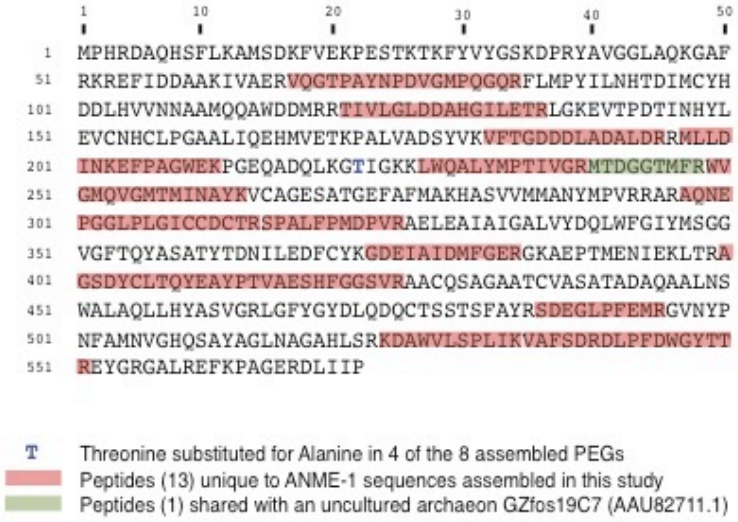
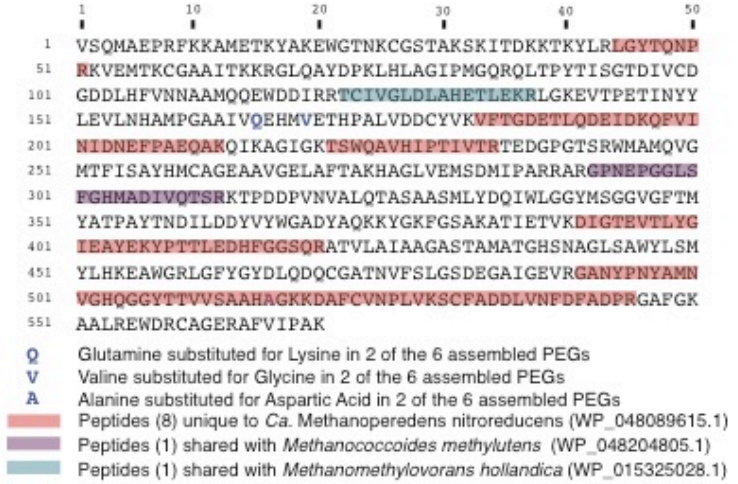


Figure 3: Peptide mapping to identified McrA and MMO proteins

A)



B)



C)

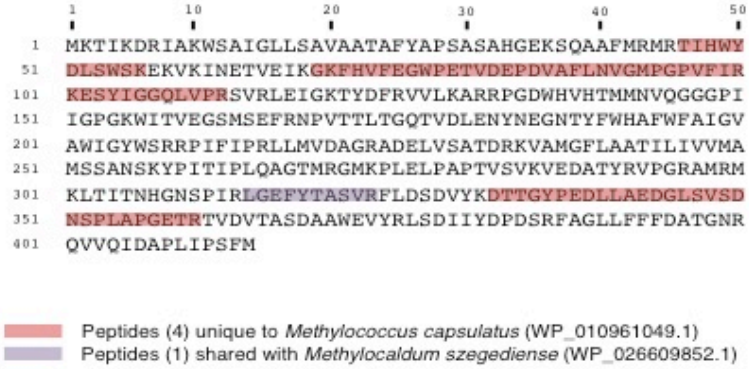
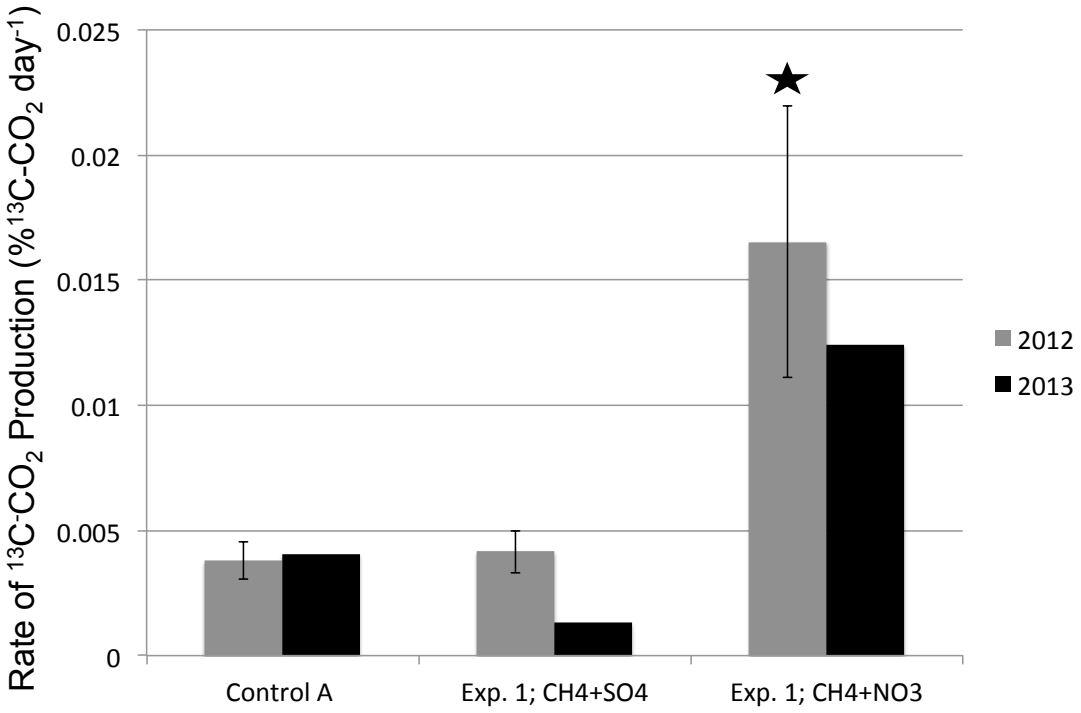


Figure 4: ^{13}C-CO_2 production after ^{13}C-CH_4 enrichment.
The rates of ^{13}C-CO_2 production (%^{13}C-CO_2 day^{-1}) from the 2012 (gray) and 2013 (black) fluids of Experiment 1 are displayed. The 2012 samples were run in triplicate and the standard deviations are shown. The star above the 2012 CH4+NO3 bar indicates that the rate of ^{13}C-CO_2 production in the ^{13}C-CH4+NO_3^- enrichment was significantly greater (p=0.02) than the electron acceptor-free control (Control A).



Supplement:

Supplementary Figure S1 – Relative abundance of archaeal 16S SSU rRNA gene v6 amplicons for the 2011 and 2012 time points

Supplementary Table 1 – CH_4 oxidizing community with respect to their CH_4-related gene

Supplementary Table 2 – Normalized coverage CH_4-related genes abundances

Supplementary Table 3 – Correlation of CH_4 oxidizing community activity to fluid chemistry

Supplementary Data 1 – Archaeal V4-V5 16S rDNA annotations

Supplementary Data 2 – Assembled *mcrA* and *mmo* genes

Supplementary Data 3 – Predicted proteins from assembled *mcrA* and *mmo*

Supplementary Data 4 – Datasets used for protein mapping

Supplementary Data 5 – Protein mapping results

Supplementary Data 6 – Enrichment experiment results

Supplementary Data 7 – McrA protein library used to generate phylogenetic tree

Acknowledgments

This research was supported by funding from the National Science Foundation to T.C.O. (EAR-0948659) and to TLK (EAR-0948335) and the Deep Carbon Observatory (Alfred P. Sloan Foundation) to M.C.Y.L. (Sloan 2013-10-03, subaward 48045). Research of P.H.A.T. is supported by the Soehngen Institute of Anaerobic Microbiology (SIAM) Gravitation grant (024.002.002) of the Netherlands Ministry of Education, Culture and Science and the Netherlands Organisation for Scientific Research (NWO). Partial support for isotopic analyses was provided by the Natural Sciences and Engineering Research Council of Canada. We are indebted to the logistical support of Sibanye Gold Limited, the management and staff of Beatrix Gold Mine and specifically to S. Maphanga of Beatrix gold mine. We thank Matthew Cahn (Department of Molecular Biology, Princeton University) and the staff of Research Computing (Office of Information Technology, Princeton University), especially Robert Knight, for their technical support with the computational analyses.

References

Bagnoud A, Chourey K, Hettich RL *et al*. Reconstructing a hydrogen-driven microbial metabolic network in Opalinus Clay rock. *Nat Commun* 2016;**7**.

Beal EJ, Claire MW, House CH. High rates of anaerobic methanotrophy at low sulfate concentrations with implications for past and present methane levels. *Geobiol* 2011;**9**:131–9.

Caldwell SL, Laidler JR, Brewer EA *et al*. Anaerobic oxidation of methane: mechanisms, bioenergetics, and the ecology of associated microorganisms. *Environ Sci Technol* 2008;**42**:6791–9.

Daly RA, Borton MA, Wilkins MJ *et al*. Microbial metabolisms in a 2.5-km-deep ecosystem created by hydraulic fracturing in shales. *Nat Microbiol* 2016;**1**:16146.

Ettwig KF, Butler MK, Le Paslier D *et al*. Nitrite-driven anaerobic methane oxidation by oxygenic bacteria. *Nature* 2010;**464**:543–8.

Ettwig KF, Zhu B, Speth D *et al*. Archaea catalyze iron-dependent anaerobic oxidation of

methane. *Proc Natl Acad Sci* 2016;**113**:12792–6.

Frape SK, Fritz P, McNutt RH t. Water-rock interaction and chemistry of groundwaters from the Canadian Shield. *Geochim Cosmochim Acta* 1984;**48**:1617–27.

Guindon S, Dufayard J-F, Lefort V *et al*. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 2010;**59**:307–21.

Haroon MF, Hu S, Shi Y *et al*. Anaerobic oxidation of methane coupled to nitrate reduction in a novel archaeal lineage. *Nature* 2013;**500**:567–70.

Holmes AJ, Costello A, Lidstrom ME *et al*. Evidence that participate methane monooxygenase and ammonia monooxygenase may be evolutionarily related. *FEMS Microbiol Lett* 1995;**132**:203–8.

Huse SM, Dethlefsen L, Huber JA *et al*. Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLoS Genet* 2008;**4**, DOI: 10.1371/journal.pgen.1000255.

Kleiveland CR, Hult LTO, Kuczkowska K *et al*. Draft genome sequence of the methane-oxidizing bacterium Methylococcus capsulatus (Texas). *J Bacteriol* 2012;**194**:6626.

Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;**9**:357–9.

Lau MCY, Cameron C, Magnabosco C *et al*. Phylogeny and phylogeography of functional genes shared among seven terrestrial subsurface metagenomes reveal N-cycling and microbial evolutionary relationships. *Front Microbiol* 2014;**5**.

Lau MCY, Kieft TL, Kuloyo O *et al*. An oligotrophic deep-subsurface community dependent on syntrophy is dominated by sulfur-driven autotrophic denitrifiers. *Proc Natl Acad Sci* 2016:201612244.

Magnabosco C, Tekere M, Lau MCY *et al*. Comparisons of the composition and biogeographic distribution of the bacterial communities occupying South African thermal springs with those inhabiting deep subsurface fracture water. *Front Microbiol* 2014;**5**:679.

McDonald IR, Bodrossy L, Chen Y *et al*. Molecular ecology techniques for the study of aerobic methanotrophs. *Appl Environ Microbiol* 2008;**74**:1305–15.

Morozova D, Wandrey M, Alawi M *et al*. Monitoring of the microbial community composition in saline aquifers during CO 2 storage by fluorescence in situ hybridisation. *Int J Greenh Gas Control* 2010;**4**:981–9.

Morozova D, Zettlitzer M, Let D *et al*. Monitoring of the microbial community composition in deep subsurface saline aquifers during CO 2 storage in Ketzin, Germany. *Energy Procedia* 2011;**4**:4362–70.

Onstott TC, Lin L-H, Davidson M *et al*. The origin and age of biogeochemical trends in deep fracture water of the Witwatersrand Basin, South Africa. *Geomicrobiol J* 2006;**23**:369–414.

Purkamo L, Bomberg M, Nyyssönen M *et al*. Dissecting the deep biosphere: retrieving authentic microbial communities from packer-isolated deep crystalline bedrock

fracture zones. *FEMS Microbiol Ecol* 2013;**85**:324–37.

Rajala P, Bomberg M, Kietäväinen R *et al*. Rapid reactivation of deep subsurface microbes in the presence of C-1 compounds. *Microorganisms* 2015;**3**:17–33.

Ruby JG, Bellare P, DeRisi JL. PRICE: software for the targeted assembly of components of (meta) genomic sequence data. *G3 Genes| Genomes| Genet* 2013;**3**:865–80.

Segarra KEA, Schubotz F, Samarkin V *et al*. High rates of anaerobic methane oxidation in freshwater wetlands reduce potential atmospheric methane emissions. *Nat Commun* 2015;**6**.

Sherwood Lollar B, Voglesonger K, Lin L-H *et al*. Hydrogeologic controls on episodic H2 release from Precambrian fractured rocks-Energy for deep subsurface life on Earth and Mars. *Astrobiology* 2007;**7**:971–86.

Shigematsu T, Tang Y, Kobayashi T *et al*. Effect of dilution rate on metabolic pathway shift between aceticlastic and nonaceticlastic methanogenesis in chemostat cultivation. *Appl Environ Microbiol* 2004;**70**:4048–52.

Simkus DN, Slater GF, Lollar BS *et al*. Variations in microbial carbon sources and cycling in the deep continental subsurface. *Geochim Cosmochim Acta* 2015.

Thauer RK. Anaerobic oxidation of methane with sulfate: on the reversibility of the reactions that are catalyzed by enzymes also involved in methanogenesis from CO 2. *Curr Opin Microbiol* 2011;**14**:292–9.

Timmers PHA, Gieteling J, Widjaja-Greefkes HCA *et al*. Growth of anaerobic methane-oxidizing archaea and sulfate-reducing bacteria in a high-pressure membrane capsule bioreactor. *Appl Environ Microbiol* 2015;**81**:1286–96.

Wegener G, Krukenberg V, Riedel D *et al*. Intercellular wiring enables electron transfer between methanotrophic archaea and bacteria. *Nature* 2015;**526**:587–90.

Young ED, Kohl IE, Lollar BS *et al*. The relative abundances of resolved l2 CH 2 D 2 and 13 CH 3 D and mechanisms controlling isotopic bond ordering in abiotic and biotic methane gases. *Geochim Cosmochim Acta* 2017;**203**:235–64.

Zehnder AJ, Brock TD. Methane formation and methane oxidation by methanogenic bacteria. *J Bacteriol* 1979;**137**:420–32.