Metagenomic identification of active methanogens and methanotrophs in serpentinite springs of the Voltri Massif, Italy

William J Brazelton Corresp., [1] , **Christopher N Thornton** [1] , **Alex Hyer** [1] , **Katrina I Twing** [2] , **August A Longino** [1] , **Susan Q Lang** [3, 4] , **Marvin D Lilley** [5] , **Gretchen L Früh-Green** [4] , **Matthew O Schrenk** [2]

[1] Department of Biology, University of Utah, Salt Lake City, Utah, United States

[2] Department of Earth and Environmental Sciences, Michigan State University, East Lansing, Michigan, United States

[3] Department of Earth and Ocean Sciences, University of South Carolina, Columbia, South Carolina, United States

[4] Department of Earth Sciences, ETH Zurich, Zurich, Switzerland

[5] School of Oceanography, University of Washington, Seattle, WA, United States

Corresponding Author: William J Brazelton
Email address: william.brazelton@utah.edu

The production of hydrogen and methane by geochemical reactions associated with the serpentinization of ultramafic rocks can potentially support subsurface microbial ecosystems independent of the photosynthetic biosphere. Methanogenic and methanotrophic microorganisms are abundant in marine hydrothermal systems heavily influenced by serpentinization, but evidence for methane-cycling archaea and bacteria in continental serpentinite springs has been limited. This report provides metagenomic and experimental evidence for active methanogenesis and methanotrophy by microbial communities in serpentinite springs of the Voltri Massif, Italy. Methanogens belonging to family Methanobacteriaceae and methanotrophic bacteria belonging to family Methylococcaceae were heavily enriched in three ultrabasic springs (pH 12). Metagenomic data also suggest the potential for hydrogen oxidation, hydrogen production, carbon fixation, fermentation, and organic acid metabolism in the ultrabasic springs. The predicted metabolic capabilities are consistent with an active subsurface ecosystem supported by energy and carbon liberated by geochemical reactions within the serpentinite rocks of the Voltri Massif.

Metagenomic identification of active methanogens and methanotrophs in serpentinite springs of the Voltri Massif, Italy

William J. Brazelton[1], Christopher N. Thornton[1], Alex Hyer[1], Katrina I. Twing[2], August A. Longino[1], Susan Q. Lang[3,5], Marvin D. Lilley[4], Gretchen L. Früh-Green[5], Matthew O. Schrenk[2]

[1]Department of Biology, University of Utah, Salt Lake City, UT, USA

[2]Department of Earth and Environmental Sciences, Michigan State University, East Lansing, MI, USA

[3]Department of Earth and Ocean Sciences, University of South Carolina, Columbia, SC, USA

[4]School of Oceanography, University of Washington, Seattle, WA, USA

[5]Department of Earth Sciences, ETH Zurich, Zurich, Switzerland

Corresponding author:
William J. Brazelton
Department of Biology
University of Utah
257 South 1400 East, Rm. 201
Salt Lake City, UT 84112-0840
(801) 587-9455
william.brazelton@utah.edu

24

25

Abstract

The production of hydrogen and methane by geochemical reactions associated with the serpentinization of ultramafic rocks can potentially support subsurface microbial ecosystems independent of the photosynthetic biosphere. Methanogenic and methanotrophic microorganisms are abundant in marine hydrothermal systems heavily influenced by serpentinization, but evidence for methane-cycling archaea and bacteria in continental serpentinite springs has been limited. This report provides metagenomic and experimental evidence for active methanogenesis and methanotrophy by microbial communities in serpentinite springs of the Voltri Massif, Italy. Methanogens belonging to family Methanobacteriaceae and methanotrophic bacteria belonging to family Methylococcaceae were heavily enriched in three ultrabasic springs (pH 12). Metagenomic data also suggest the potential for hydrogen oxidation, hydrogen production, carbon fixation, fermentation, and organic acid metabolism in the ultrabasic springs. The predicted metabolic capabilities are consistent with an active subsurface ecosystem supported by energy and carbon liberated by geochemical reactions within the serpentinite rocks of the Voltri Massif.

49

50

Introduction

Hydrothermal systems hosted in ultramafic rocks rich in serpentine minerals can support active and unique microbial communities (see Schrenk, Brazelton & Lang, 2011 for review). Microbial ecosystems associated with serpentinization (the process by which serpentine is formed) have been investigated by sampling seafloor hydrothermal chimneys (e.g. Kelley et al., 2005; Quéméneur et al., 2014), alkaline springs on land (Brazelton et al., 2013; Suzuki et al., 2013; Woycheese et al., 2015), and wells drilled into serpentinizing rocks (Tiago & Veríssimo, 2013; Cardace et al., 2013; Crespo-Medina et al., 2014). The microbial inhabitants of these environments are thought to feed on the hydrogen gas and organic compounds released as byproducts of serpentinization, but representative examples of specific organisms consuming specific products of serpentinization are lacking. For example, methane is typically abundant in serpentinite-hosted hydrothermal systems, but only a few studies have provided evidence for methanogenic or methanotrophic biological activity in such systems (Brazelton et al., 2011; Kohl et al., 2016).

The role of methane-cycling organisms in serpentinizing systems has been enigmatic since the discovery of thick biofilms of methanogens belonging to order Methanosarcinales in carbonate-brucite chimneys of the Lost City hydrothermal field (Schrenk et al., 2004). Whether these organisms are primarily involved in methane production or methane oxidation remains uncertain (Proskurowski et al., 2008; Bradley, Hayes & Summons, 2009; Brazelton et al., 2011; Lang et al., 2012; Méhay et al., 2013). Similar methanogens belonging to order Methanosarcinales have been identified in Prony Bay, New Caledonia, where freshwater serpentinizing fluids are venting into shallow marine waters and forming carbonate chimneys on the seafloor (Quéméneur et al., 2014).

16S rRNA genes associated with methanotrophic bacteria have been detected at Lost City (Brazelton et al., 2006) and Prony Bay (Quéméneur et al., 2014).

Recent studies of non-marine alkaline springs associated with serpentinization have identified 16S rRNA and/or methyl coenzyme-M reductase (*mcrA*) genes belonging to methanogens in California (The Cedars: Suzuki et al., 2013; Adobe Springs: (Blank et al., 2009), Costa Rica (Sánchez-Murillo et al., 2014), and the Philippines (Woycheese et al., 2015). At the Tablelands Ophiolite (Newfoundland, CA), however, experimental incubations have been unable to detect methanogenesis at high pH (Morrill et al., 2014), and a metagenomic study at this site did not detect any methanogenesis genes (Brazelton, Nelson & Schrenk, 2012; Brazelton et al., 2013). 16S rRNA and *mcrA* sequences belonging to the ANME-1a group of anaerobic methanotrophic archaea, but not typical methanogens, were detected in serpentinizing groundwater in the Cabeço de Vide Aquifer in Portugal (Tiago & Veríssimo, 2013). Biogeochemical studies at The Cedars have inferred biological methanogenesis from isotopic signatures of methane and experimental results, but these studies have not included any biological characterizations of the responsible organisms (Morrill et al., 2013; Wang et al., 2015; Kohl et al., 2016). Thus, methanogens have been detected in some continental serpentinite springs, but previous reports have not provided any quantitative measurements of their environmental distributions nor any genomic or metagenomic sequences from methanogens. Evidence for methanotrophic bacteria in these systems is even more scarce: only a single study of non-marine alkaline springs associated with serpentinization has previously reported methanotrophic bacteria (Sánchez-Murillo et al., 2014).

In the Voltri Massif (Italy), ultrabasic springs (pH 11-12) rich in calcium and methane are issuing from serpentinites and rare lherzolites and are associated with the precipitation of large amounts of carbonate (Bruni et al., 2002; Cipolli et al., 2004; Boulart et al., 2013; Schwarzenbach et al., 2013). Inorganic carbon geochemistry of the springs and associated travertine deposits suggest autotrophic microbial activity in the subsurface (Schwarzenbach et al., 2013). A previous study has characterized the archaeal and bacterial communities within the travertine deposits formed when the springs reach the surface (Quéméneur et al., 2015), but there are no previous microbiological studies of the spring water. In this report we show that pH 12 springs at the Voltri Massif are transporting distinct archaeal and bacterial communities, including methanogenic archaea and methanotrophic bacteria, from subsurface habitats where they are likely to be supported by hydrogen gas, methane, and possibly other products of serpentinization-associated reactions. We provide insights into the biology and potential subsurface habitats of these organisms with metagenomic and experimental studies.

Materials and Methods

Location and sample collection

The ultrabasic springs that were investigated in this study (BR2: 44.4512ºN, 8.7820ºE; GOR34: 44.5970ºN, 8.7833ºE) are located in the Voltri Massif near Genoa, Italy (Figure 1). Geochemical measurements of these springs were conducted in concert with the microbiological studies reported here and were reported by Schwarzenbach et al. (2013). At each of the BR2 and GOR34 locations, ultrabasic spring water and nearby river water were filtered through Millipore Sterivex 0.2 μm filter cartridges using a portable peristaltic pump. Measurements of pH and Eh were obtained with an Oakton PCS Testr 35 and an Oakton Testr 10, respectively, from pumped water

119 that had passed through the peristaltic tubing but before the Sterivex filter (Table 1). Eh readings
120 were corrected for the standard Ag/AgCl electrode (+200 mV).
121
122 GOR34-1-spring and GOR34-3-spring (Figure 1A) were sampled from two shallow pools within
123 a large deposit of travertine (Schwarzenbach et al., 2013). Each pool appeared to be fed by a
124 point source of spring water at the bottom, into which peristaltic tubing was inserted as far as
125 possible into the host rock (a few centimeters). The GOR34-1 pool was much smaller than that
126 of GOR34-3, and it was emptied before sampling by siphoning water out of the pool. During
127 pumping and sampling of water from the apparent subsurface source, the pump rate was adjusted
128 to approximately match the rate of inflow, determined by a constant water level in the pool. The
129 volume of water in the GOR34-3 pool did not visibly change during sampling, indicating that the
130 maximum pump rate was significantly slower than the rate of spring water inflow into the pool.
131 The pumped spring water from both GOR34-1 and GOR34-3 became more basic (higher pH)
132 and more reducing (lower Eh) during pumping of the first few liters, after which the readings
133 stabilized, and filtering commenced. GOR34-river was sampled from the surface of the adjacent
134 river at a site of exposed, rapid flow ~50 m upstream of the spring.
135
136 BR2-spring was sampled from a metal pipe (Figure 1B) through which spring water was flowing
137 at 492 mL s^{-1} into the adjacent river (Schwarzenbach et al., 2013). After filtering >500 L of
138 spring water from BR2-spring in 2012, less than 500 ng of DNA was recovered in initial
139 extractions, which was reduced to 50 ng after pooling and final purification. Even lower yields
140 were achieved in 2013 (Table 2). BR2-river was sampled from the surface of the river at a site of
141 exposed, rapid flow ~10 m upstream of the spring.

142

143 BR2-spring, BR2-river, and GOR34-river were sampled in June 2012 and August 2013.
144 GOR34-1 was sampled in October 2011 and August 2013, while GOR34-3 was only sampled in
145 June 2012. Field replicates were collected for all spring and river samples, but replicates were
146 pooled post-DNA-extraction when necessary to accumulate enough DNA for sequencing.
147 Furthermore, even after pooling of replicates, some samples contained only enough DNA for
148 either 16S rRNA amplicon sequencing or shotgun metagenomic sequencing, but not both. For
149 this reason, BR2-spring-2012, BR2-spring-2013, and GOR34-spring3-2012 are represented in
150 this study by shotgun metagenomic sequences but not 16S rRNA amplicon sequencing surveys
151 (Table 2). When available, field replicates are indicated with 'a' or 'b' at the end of the sample
152 name, such as GOR34-river-2013a and GOR34-river-2013b.

153

154 *Water chemistry*

155 Water samples were collected for measurements of hydrogen, methane, sulfate, and sulfide. Gas
156 samples were collected with 100 mL syringes by sampling 60 mL of water, minimizing any
157 incorporation of air. Subsequently, 40 mL of nitrogen or helium gas was introduced to the
158 syringe as headspace gas and the two phases were allowed to equilibrate. The gas phase was
159 injected into evacuated vacutainers for later analysis. The concentrations of gas were determined
160 by gas chromatography (GC), calibrated with commercially available gas standards of known
161 concentration. For hydrogen, samples were injected in split/splitless mode (5:1 split) onto a CP-
162 Molsieve 5Å column (50 m × 0.32 μm, 30 μm thickness) at 175°C with a helium flow rate of 1.8
163 mL/min and detected with a Pulsed Discharge Detector. Methane was determined similarly but
164 with a GS-Carbonplot column (30 m × 0.32 μm ID, 1.5 μm film thickness) held at 100°C for 3

minutes then ramped at 50 °C/min to 200 ºC and held for 0.6 minutes. Detection was with a flame ionization detector. The detection limit was ~0.6 nM for both hydrogen and methane.

Dissolved inorganic carbon (DIC) measurements were performed on samples that were collected by injecting 1-8 mL of 0.2 μm filtered water into Exetainers (Labco Limited, UK) that had been previously prepared with ~100 μL of phosphoric acid and flushed with helium. Concentrations were determined by injecting aliquots of the headspace with a calibrated gas-tight syringe into the same GC set-up used for hydrogen. DIC concentrations from 2011 were reported by Schwarzenbach et al. (2013).

Sulfide concentrations were determined by the methylene blue method and detected at 610 nm on a spectrophotometer (Cline, 1969). Concentrations were compared to standard curves of freshly prepared sodium sulfide. The detection limit was 3.1 μM. For sulfate, 100 mL of water was collected in high density polypropylene bottles and stored cool until analysis. Concentrations were determined by a DX-120 ion chromatograph equipped with an IonPac As14 column (4 × 250 mm). The detection limit was 1.0 μM. Organic acids were analyzed by high performance liquid chromatography (HPLC) by the method of Albert & Martens (1997) with minor modifications. After derivatization with 2-nitrophenyl hydrazine and 1-ethyl-3-(3-dimethylaminopropyl) carbodimide hydrochloride, acids were separated on a Prevail Organic Acid C18 column and detected at 400 nm. Adipic acid was used as an internal standard.

Enumeration of microbial cells

Fluids (50 mL per field replicate; three replicates per sample in Table 2) were preserved in the field for cell abundance enumeration at a final concentration of 3.7% formaldehyde and stored at 4°C. In the laboratory, preserved fluids (5 to 20 ml each) were filtered through a 0.2 μm black polycarbonate filter, and captured cells were stained with DAPI and counted with an epifluorescence microscope. At least 30 fields were counted for each field replicate and used to calculate the average cell concentration per mL of fluid for that replicate sample. The numbers reported in Table 2 reflect the mean cell concentration for three field replicates from each sample.

DNA extraction and sequencing

Filters were placed on ice immediately and stored within a few hours at liquid nitrogen temperature in a vapor shipper (MVE SC4/2V) for frozen transport to the home laboratory. Each Sterivex filter was extracted according to protocols modified from those described in (Huber, Butterfield & Baross, 2002; Sogin et al., 2006). Briefly, extractions were performed by lysis via freeze/thaw cycles and lysozyme/Proteinase K treatment and purified with phenol-chloroform extractions and precipitation in ethanol. Environmental DNA quantification was performed with the Qubit fluorometer (ThermoFisher). Extraction blanks never yielded quantifiable DNA and were not sequenced. Further purification of DNA was conducted with QiaAmp (Qiagen) columns according to the manufacturer's instructions for purification of genomic DNA.

Purified DNA was submitted to the Josephine Bay Paul Center, Marine Biological Laboratory (MBL) at Woods Hole for amplicon sequencing of archaeal and bacterial 16S rRNA genes with an Illumina MiSeq platform. The V4-V5 hypervariable region of the bacterial 16S rRNA gene

was targeted with one forward (518F; CCAGCAGCYGCGGTAAN) and three reverse primers (926R; CCGTCAATTCNTTTRAGT; CCGTCAATTTCTTTGAGT; CCGTCTATTCCTTTGANT). The V4-V5 hypervariable region of the archaeal 16S rRNA gene was targeted by a combination of five forward primer variants (517F; GCCTAAAGCATCCGTAGC; GCCTAAARCGTYCGTAGC; GTCTAAAGGGTCYGTAGC; GCTTAAAGNGTYCGTAGC; GTCTAAARCGYYCGTAGC) and a single reverse primer (958R; CCGGCGTTGANTCCAATT). Additional methodological details and description of primer development are published in (Nelson et al., 2014; Topçuoğlu et al., 2016). Amplicon sequences were screened for quality, including chimera-checking with UCHIME (Edgar et al., 2011) by the MBL as previously described (Huse et al., 2014a), and high-quality merged sequences were published on the Visualization and Analysis of Microbial Population Structures (VAMPS) website (Huse et al., 2014b).

The MBL also conducted shotgun metagenomic sequencing of a subset of these samples (see Table 2). Metagenomic libraries were constructed with the Nugen Ultralow Ovation kit according to the manufacturer's instructions. Paired-end sequencing with a 100 cycle Illumina HiSeq run generated partial ~30 bp overlaps, and six libraries were multiplexed per lane. 16S rRNA amplicon sequences are publicly available at the VAMPS database (http://vamps.mbl.edu) under the project code DCO_BRZ and sample code Serp_LIG. Shotgun metagenomic data is publicly available in MG-RAST under IDs 4545477.3, 4545478.3, 4545479.3, 4545480.3, 4537863.3, 4537864.3, 4537868.3, and 4537869.3. All sequence data related to this study are also available via the SRA identifier SRP049438 and BioProject PRJNA265986.

Quantitative PCR of archaeal and bacterial 16S rRNA gene copies

The 16S rRNA gene copies of Bacteria and Archaea were quantified via quantitative PCR on a BioRad CFX Connect Optics Module with the BioRad SsoAdvanced SybrGreen assay and domain-specific primers targeting the V6 region of the 16S rRNA gene. Primers 958F and 1048R were used for Archaea, and primers 967F and 1064R were used for Bacteria (previously published in Sogin et al., 2006 and described on the VAMPS website https://vamps.mbl.edu/resources/primers.php). Gene copy numbers were calculated by plotting quantification values from environmental samples onto standard curves generated by amplification of DNA from *Methanocaldococcus jannaschii* (for Archaea) and *Escherichia coli* (for Bacteria) with the domain-specific primers. Bacteria:Archaea ratios were then calculated with gene copy numbers normalized to sample size, i.e. the volume of fluid filtered for that sample.

Analysis of 16S rRNA amplicon data

Additional quality screening of 16S rRNA amplicon sequences from VAMPS was conducted with the mothur (v.1.34.2) software platform (Schloss et al., 2009) to remove sequences with >9 homopolymers and >0 ambiguous bases. This screening step removed only 124 bacterial sequences and 5 archaeal sequences. The sequences were then pre-clustered with the mothur command pre.cluster (diffs=1), which reduced the number of unique sequences from 410,517 to 322,879 for bacteria and from 138,629 to 95,867 for archaea. This pre-clustering step removes rare sequences most likely created by sequencing errors (Schloss, Gevers & Westcott, 2011). The final unique sequence types were considered to be the operational taxonomic units (OTUs) for this study. We chose not to cluster sequences any more broadly (e.g. 97% sequence similarity)

256 because additional clustering inevitably results in a loss of biological information without
257 reliably improving sequencing error and because no arbitrary sequence similarity threshold can
258 be demonstrated to correspond to consistent species-like units across all taxa. Taxonomic
259 classification of all OTUs was performed with mothur using the SILVA reference alignment
260 (SSURefv119) and taxonomy outline (Pruesse, Peplies & Glöckner, 2012). Differences in the
261 relative abundances of OTUs between groups of samples were measured with the aid of the R
262 package edgeR v.3.6.8 (Robinson, McCarthy & Smyth, 2010) as recommended by (McMurdie &
263 Holmes, 2014). Results were visualized with the aid of the R package phyloseq v.1.9.13
264 (McMurdie & Holmes, 2013).
265
266 *Analysis of metagenomic data*
267 Detailed documentation of all computational analyses in this study, including actual software
268 commands, are provided on the Brazelton lab's website (https://baas-
269 becking.biology.utah.edu/data/category/18-protocols), and all custom software and scripts are
270 available at https://github.com/Brazelton-Lab. Quality-filtering of shotgun metagenomic
271 sequences included identification and removal of artifactual sequences with cutadapt v.1.9
272 (Martin, 2011) as follows. Reads found to have Illumina adapters starting at their 5'-end were
273 discarded, and reads containing Illumina adapters towards the 3'-end were trimmed where the
274 first adapter began. Identical and 5'-prefix replicates were also removed, as suggested by
275 (Gomez-Alvarez, Teal & Schmidt, 2009). Nucleotides (0-2) at the beginning and end of reads
276 were also cropped from all reads in that sample if those positions exhibited nucleotide
277 frequencies inconsistent with the nucleotide frequency distribution for the rest of the read. Low-
278 quality bases were removed from the ends of the reads, and the remaining sequence was scanned

20 base pairs at a time and trimmed where the mean quality score fell below a score of 8. Reads that did not pass a minimum length threshold of 62 bp after quality and adapter trimming were removed from the dataset. Phylogenetic affiliations of the unassembled, quality-checked paired reads were determined with PhyloSift v.1.0.1 (Darling et al., 2014).

Each metagenomic dataset was processed individually as described above. The four metagenomes from ultrabasic springs and the four metagenomes from adjacent rivers were then combined for two pooled assemblies with Ray Meta v.2.3.1 (Boisvert et al., 2012). A kmer of 41 was chosen after manual inspection of assemblies with kmer values of 31, 41, 51, and 61. High-quality reads from each sample were mapped onto the pooled assembly with Bowtie2 v.2.2.6 (Langmead & Salzberg, 2012) to obtain sample-specific coverages. The Prokka pipeline (Seemann, 2014) was used for gene prediction and functional annotation. The arguments --metagenome and --proteins were used with Prokka v.1.12 to indicate that genes should be predicted with the implementation of Prodigal v.2.6.2 (Hyatt et al., 2010) optimized for metagenomes and then searched preferentially against a custom protein database. The database provided was the Kyoto Encyclopedia of Genes and Genomes FTP release 2016-09-26 (Ogata et al., 1999). Abundances of metabolic pathways were obtained by mapping KEGG protein IDs and their normalized counts calculated with HTSeq v.0.6.1 (Anders, Pyl & Huber, 2015) onto the FOAM ontology (Prestat et al., 2014) with MinPath (Ye & Doak, 2009) as implemented in HUMAnN2 v.0.6.0 (Abubucker et al., 2012). Similar results were also achieved by annotating both pooled assemblies with the UniProtKB protein database (Consortium, 2015) and MetaCyc pathways (Caspi et al., 2014).

Contigs from the pooled spring assembly were binned according to their tetranucleotide frequencies with the aid of an emergent self-organized map (ESOM) as implemented by Databionic ESOM tools v.1.1 (Ultsch & Mörchen, 2005) and as described by (Dick et al., 2009). The resulting ESOM was annotated with phylogenetic markers classified by PhyloSift and selected predicted protein functions of interest identified by KEGG ID. Completeness and contamination of ESOM bins were evaluated with CheckM v.1.0.4 (Parks et al., 2015). Contamination levels as measured by CheckM do not always agree with the percentage of minority taxa reported by PhyloSift because CheckM only identifies contamination when single-copy markers are present as multiple, divergent sequences and because PhyloSift results are abundance-weighted.

Metabolic Activity Assays

During each sampling expedition, water samples collected from each of the BR2 and GOR34 springs were incubated with one of several ^{13}C-labeled carbon sources (obtained from Cambridge Isotope Laboratories, Tewksbury, MA): bicarbonate ($NaH^{13}CO_3$), formate ($Na^{13}CHOO$), acetate-^{13}C1 ($NaCH_3{}^{13}COO$), acetate-^{13}C2 ($Na^{13}CH_3COO$), propionate ($NaCH_3CH_2{}^{13}COO$), or methane ($^{13}CH_4$). Each 99% ^{13}C-labeled carbon compound was diluted 1:100 with the corresponding non-^{13}C-labeled compound to produce a final $\delta^{13}C$ of +691 ‰, and this compound was present in each incubation at a final concentration of 0.1 M after addition of 3 mL of water sample. No buffering nor additional amendments were applied to the incubations, and the experiments are assumed to have proceeded at the in situ pH of the spring water. ‘Dead’ control incubations were conducted by filter-sterilizing the spring water through a 0.2 µm syringe filter in the field and prior to incubation. Three live replicates and three dead replicates

from each spring were incubated for each treatment. Methanogenesis experiments received an addition of 5 cc H_2 gas to each replicate. Methanotrophy experiments received an addition of 5 cc $^{13}CH_4$ that had been diluted 1:10 with non-^{13}C-labeled methane. All incubations were conducted in gas-tight glass Exetainers (Labco Limited, UK) flushed with N_2 and 2-5% H_2 in a Coy anoxic chamber and flushed again outside the chamber and slightly over-pressured for transport with N_2 gas. Oxic methanotrophy experiments were conducted after exposing the contents to room air by opening the lids for a few minutes prior to adding $^{13}CH_4$. Some methanogenesis experiments were treated with a reducing agent in the form of 0.3 mL of 1% dithiothreitol (DTT) or 2.5% sodium sulfide, but no differences were observed with or without the additional reducing agent. After incubation at ambient temperature (20-35 ºC) in the dark for 4-8 days, Exetainers were injected with 0.2 mL of 50% phosphoric acid to convert all dissolved inorganic carbon into CO_2, and the $^{13}C/^{12}C$ ratios of CO_2 and CH_4 were measured by gas chromatography isotope ratio mass spectrometry using a Thermo Fisher Delta V Plus Isotope Ratio Mass Spectrometer, interfaced with a Trace gas chromatograph, a GC IsoLink interface and a ConFlo IV at the Stable Isotope Laboratory of the ETH in Zurich.

Results

Biogeochemical contrasts between ultrabasic springs and surface rivers

At both the BR2 and GOR34 sites, the spring water had much higher pH, lower Eh (i.e. more reducing), fewer particulates, and lower biomass than the adjacent river (Tables 1-2). The high pH (up to 12.3) and low Eh are consistent with the spring water being sourced from ultramafic rocks where serpentinization-like reactions are occurring now or have occurred in the past (Alt et al., 2013). The difference in particulates was noticeable due to the much greater volumes of

spring water, compared to surface river water, that could be filtered through a single Sterivex cartridge before clogging (more than 20-50 L for springs, only 2-5 L for surface rivers). The difference in biomass was also evident from the DNA yield: the GOR34 springs yielded 3-5-fold less DNA per liter of water compared to the adjacent river, and the BR2 spring yielded >1,000-fold less DNA per liter of water compared to the adjacent river. The cell densities (as measured by enumeration of DAPI-stained cells under a microscope) of the BR2 spring (10^2 – 10^3 cells per mL of water) were 2-40-fold lower than the cell densities of the adjacent river (Table 2). The cell densities of the GOR34 springs were 2-20-fold lower than the adjacent river.

BR2 spring waters contained higher methane and lower hydrogen concentrations compared to the GOR34 springs (Table 1). At both sites, spring water contained much lower dissolved inorganic carbon (DIC) than in the adjacent rivers, likely due to the precipitation of calcium carbonate at high pH (Schwarzenbach et al., 2013; Alt et al., 2013). Sulfate was also very low in all springs, though detectable at BR2. Sulfide was elevated in all springs compared to the adjacent rivers.

Taxonomic survey of bacterial 16S rRNA genes

The bacterial taxonomic composition of the pH 12 spring GOR34-spring1 was distinct from that of surface rivers, as measured by high-coverage Illumina MiSeq sequencing of bacterial 16S rRNA gene amplicons. Comamonadaceae (a member of order Burkholderiales, class Betaproteobacteria) was the most abundant family in GOR34-spring1 in 2013 (35-40% of all bacterial 16S rRNA amplicon sequences, Supplementary File S1). The most abundant OTU within this family is 100% identical to the most abundant OTU recovered from ultrabasic

371 serpentinite springs at the Tablelands Ophiolite in Newfoundland, Canada (Brazelton et al., 2013)
372 and to the most abundant OTU in borehole fluids from the Coast Range Ophiolite Microbial
373 Observatory in California (Crespo-Medina et al., 2014). This same bacterial taxon has been
374 isolated from serpentinite springs at The Cedars (northern California) by Suzuki et al. (2014),
375 who have proposed the novel genus *Serpentinomonas* for these organisms. Curiously,
376 Comamonadaceae sequences only comprised 14% of all bacterial sequences in GOR34-spring1
377 in 2011. Instead, the most abundant sequences were those classified as Candidate Division OD1
378 (21% of sequences in GOR34-spring1 in 2011 but only 4% in 2013).
379
380 Comamonadaceae sequences were also well represented in surface river samples (Supplementary
381 File S1), but the Comamonadaceae OTUs from river samples were distinct from the
382 Comamonadaceae OTUs from ultrabasic spring samples (only 94% sequence identity over 373
383 bases between the most abundant Comamonadaceae OTU in GOR34-spring1-2013b compared to
384 the most abundant Comamonadaceae OTU in GOR34-river-2013). Neither of these sequences
385 could be classified at a lower taxonomic rank than family. This result highlighted the need to
386 conduct comparisons of bacterial community compositions at the level of individual OTUs
387 (Supplementary File S2), rather than taxonomic classifications.
388
389 *Differential abundance of bacteria*
390 Although the bacterial community compositions of ultrabasic springs and the adjacent rivers
391 were generally distinct, the most abundant OTUs in rivers were also present, at low relative
392 abundances, in ultrabasic springs, and the converse was also true. In other words, the most
393 abundant 16S rRNA OTUs at each site were not completely exclusive to that site. For example,

394 the most abundant OTU in GOR34-river-2013b (classified as Sphingomonadaceae) occurred
395 8515 times in that sample and 0 and 4 times in GOR34-spring1-2013a and GOR34-spring1-
396 2013b, respectively. Conversely, the most abundant OTU in GOR34-spring1-2013b (the
397 Comamonadaceae *Serpentinomonas* OTU described above) occurred 32,866 times in that sample
398 and 71 and 0 times in GOR34-river-2013a and GOR34-river-2013b, respectively
399 (Supplementary File S2). As with any environmental study, these sequences could have appeared
400 in multiple locations because of natural mixing in the environment or due to accidental
401 contamination during sampling or sequencing. Regardless of the cause, the source of shared
402 sequences can be inferred by making the parsimonious assumption that the site of higher
403 abundance is closer to that organism's true habitat. Therefore, we looked for potential
404 subsurface-specific organisms by testing for significant differences in the abundances of
405 individual OTUs between ultrabasic spring water and the adjacent river.
406
407 For this analysis, we contrasted the counts of OTUs in GOR34-spring1 to those in GOR34-river.
408 Each site was represented by two field replicates collected in 2013 (Table 2). In Figure 2, each
409 data point of the plot represents the difference in abundance of a single OTU between GOR34-
410 spring1 and GOR34-river. The plot's x-axis displays the mean abundance (in units of $\log_2$ counts
411 per million) of each OTU across all four samples (two GOR34-spring1 replicates and two
412 GOR34-river replicates). Red data points indicate OTUs whose differential abundance between
413 GOR34-spring1 and GOR34-river passed a significance test (false discovery rate < 0.05)
414 implemented by the edgeR package (Robinson, McCarthy & Smyth, 2010). Spring-enriched
415 OTUs include the Comamonadaceae described above as well as a group of uncultured
416 Bacteroidetes known as aquatic group ML635J-40, Candidate Division OD1, an uncultured

group within the Thermoanaerobacterales, and Methylococcaceae (Figure 2A). Sequences that were not consistently enriched in one sample type compared to the other are represented by black points in Figure 2. These sequences could represent potential contaminants or truly cosmopolitan species, or their abundances may be correlated to factors not captured by the study design. The data visualization approach in Figure 2 highlights the key differences between the spring and river microbial communities by acknowledging the reality that all environmental samples are mixtures formed by multiple sources of organisms and without making the naive and extreme assumption that all shared sequences must be some kind of contamination.

Taxonomic survey of archaeal 16S rRNA genes

As with the bacterial communities, the archaeal communities in the ultrabasic spring waters were distinct from those in the surface rivers, as measured by high-coverage Illumina MiSeq sequencing of archaeal 16S rRNA gene amplicons. More than 88% of total archaeal sequences in GOR34-spring1-2013, for example, belong to the Methanobacteriaceae family of methanogens. The closest match (~97% sequence similarity) in the GenBank, VAMPS, and SILVA databases to these Methanobacteriaceae sequences is a clone from subsurface water collected in a deep South Africa mine (GenBank DQ230925). The taxonomic composition of GOR34-river was nearly completely different; ~72% of sequences in both 2012 and 2013 were assigned to Marine Group I, within the Thaumarchaeota (Supplementary File S3). As with the bacterial sequences, the most abundant archaeal OTUs in GOR34-river were also present, at low relative abundances, in GOR34-spring1, and the converse was also true. For example, the most abundant OTU in GOR34-spring1-2013b occurs 89,407 times in that sample, 127 times in GOR34-river-2013, and 24 times in GOR34-river-2012 (Supplementary File S4).

We identified potential subsurface-specific archaeal 16S rRNA OTUs by contrasting the archaeal sequences from GOR34-spring1 to those of GOR34-river. Unfortunately, GOR34-spring1-2013b was the only spring sample to yield sufficient DNA and sufficient read coverage for this analysis of archaeal 16S rRNA sequences. GOR34-river is represented in this analysis by two samples: GOR34-river-2012a and GOR34-river-2013a. Not surprisingly, considering the differential abundances reported above, the Methanobacteriaceae sequences were identified as the most abundant OTUs that were significantly enriched in GOR34-spring1 (Figure 2B). Other families of methanogens in addition to Methanobacteriaceae were identified in the 16S rRNA sequences, but none were significantly more abundant in the ultrabasic springs compared to the adjacent river. One sequence belonging to family Methanosaetaceae (order Methanosarcinales) was enriched in the GOR34 spring compared to the adjacent river in 2013, but Methanosaetaceae sequences were equally abundant in the river in 2012 (Supplementary Files S3-S4). The only OTUs that could be considered significantly enriched in GOR34-river belong to the Thaumarchaeota Marine Group I family. The variability of less abundant OTUs between the two samples from GOR34-river prevented them from passing our significance test, even though many of them were extremely rare or absent in GOR34-spring1.

Taxonomic survey of shotgun metagenomic sequences

Phylogenetic classifications (from PhyloSift) of unassembled metagenomic sequences were largely consistent with the 16S rRNA amplicon results described above. For example, even though Comamonadaceae was the most common family identified in metagenomic sequences from the ultrabasic springs as well as the river at GOR34 (Supplementary File S5), the sample

types differed at the genus level: *Hydrogenophaga* was the most common member of Comamonadaceae in the GOR34-spring1 and GOR34-spring3 metagenomes, whereas *Acidovorax* was the predominant representative of Comamonadaceae in GOR34-river (Supplementary File S6). Comamonadaceae sequences were also abundant in metagenomic data from the ultrabasic spring at BR2, although this site appears to be dominated by Desulfovibrionales, a member of Deltaproteobacteria (Supplementary File S5).

Metagenomic sequences predicted to represent methanogens (family Methanobacteriaceae) comprised ~1% of shotgun sequences in BR2-spring-2012 and GOR34-spring3-2012 (Supplementary File S5). Curiously, no methanogen sequences were found in GOR34-spring1-2013b (Supplementary File S6), despite their dominance in the archaeal 16S rRNA amplicon data from this same sample (Figure 2). None of the 28 million sequence pairs in this metagenome were classified as domain Archaea, so it is possible that the ratio of bacterial to archaeal DNA in this sample was very high. A quantitative PCR assay of bacterial and archaeal 16S rRNA gene copies indicated a 10:1 ratio of Bacteria:Archaea (Table 2) in GOR34-spring1-2013b, which would predict a minority, but perhaps not a complete absence, of archaeal sequences in its metagenome. The Bacteria:Archaea ratio for BR2-spring-2012 was less than 2:1, and only 3% of shotgun metagenomic sequences from this sample were classified as Archaea, indicating a lack of direct correspondence between quantitative PCR results and numbers of shotgun sequences. When archaea were detected in metagenomic sequences from the rivers adjacent to the BR2 and GOR34 springs, they were dominated by Thaumarchaeota (Supplementary File S5), consistent with the archaeal 16S rRNA amplicon results (Figure 2). No Methanobacteriaceae sequences were detected in the metagenomic data from either river.

Metagenomic assembly and metabolic pathway prediction

Shotgun metagenomic sequences were assembled into contiguous genomic fragments (contigs), which were used for predictions of genes, protein functions, and metabolic pathways. Because the four metagenomes from ultrabasic springs (BR2-spring-2012, BR2-spring-2013, GOR34-spring1-2013b, and GOR34-spring3-2012) were similar to each other in taxonomic composition (Supplementary Files S5-S6), they were combined in a pooled assembly called 'combined-spring'. The N50 (length of median contig) of this pooled assembly was 2.4 kb (Table 3), and the longest contig was 99 kb. Of the 104,651 predicted proteins in the combined-spring assembly, 49,794 could be annotated with the KEGG database. A large proportion (45%) of predicted proteins had no matches in any of the protein databases and could not be assigned a function (Table 3). The four metagenomes from adjacent rivers (BR2-river-2012a, BR2-river-2013, GOR34-river-2012a, and GOR34-river-2013b) were also combined for a pooled assembly called 'combined-river', which has similar assembly statistics to the combined-spring assembly (Table 3).

The abundance of metabolic pathways was determined by mapping predicted protein functions (as annotated with KEGG) and their sequencing read coverages onto the FOAM ontology. Many of the most abundant predicted pathways in the ultrabasic springs involved transporters and stress responses, possibly indicating cellular responses to the extreme environmental conditions. The most abundant energy-conserving metabolic pathways included those involved in fermentation, methanogenesis, sulfur oxidation, denitrification, methylotrophy, and hydrogen oxidation and production (Supplementary File S7). The presence of genes associated with the

509 aerobic oxidation of hydrogen and methanogenesis are are expected from the dominance of the
510 *Hydrogenophaga* and Methanobacteriaceae sequences described above. The abundance of
511 several pathways involved in fermentation and hydrogen production is consistent with the
512 presence of Clostridiaceae and Bacteroidetes, and the detection of methylotrophy pathways is
513 consistent with the presence of Methylococcaceae in the ultrabasic springs (Supplementary File
514 S7).
515
516 Methanogenesis pathways using carbon dioxide and formate were among the most abundant
517 pathways in the spring metagenomes, but other FOAM-defined methanogenesis pathways
518 predicted to utilize acetate and methylated compounds were also predicted to be present in these
519 springs (Supplementary File S7). To better understand the evidence for different kinds of
520 methanogenesis pathways in these springs, we examined the presence of each key
521 methanogenesis-associated gene. All of the steps in the 'core' methanogenesis pathway were
522 represented by at least one predicted protein (as defined by KEGG) in the combined-spring
523 assembly (green highlighting in Figure 3). The only exception is the lack of sequences encoding
524 the enzyme Hmd (H_2-forming methylenetetrahydromethanopterin dehydrogenase), but this gene
525 is not present in many methanogens (Fricke et al., 2006), including the five *Methanobacterium*
526 genomes currently available in the Integrated Microbial Genomes database (img.jgi.doe.gov).
527
528 All of the genes encoding formate dehydrogenase (Fdh), acetate kinase (Ack),
529 phosphotransacetylase (Pta), AMP-forming acetyl-CoA synthetase (Acs), and carbon monoxide
530 dehydrogenase / acetyl-CoA decarbonylase/synthase (Cdh/ACDS) are present in the combined-
531 spring assembly. The presence of these genes indicates the genetic potential for both formate and

acetate to be used as substrates for methanogenesis, although these proteins can be utilized by non-methanogens as well, which can explain why some of these proteins are also found in the combined-river assembly (yellow highlighting in Figure 3). Two of the three genes encoding Mta (methanol-specific methylcobalamin:coenzyme M methyltransferase) are also present, suggesting methanol as another potential substrate for methanogenesis.

Binning of metagenomic assembly

In order to examine the distribution of predicted protein functions and metabolic pathways among putative organisms within the combined-spring assembly, fragments of metagenomic contigs were binned according to their tetranucleotide compositions with an emergent self-organizing map (ESOM). Each point in the ESOM in Figure 4 represents a contig or a ~5 kb fragment of a contig if the original contig was larger than 5 kb. In Figure 4A, contigs containing PhyloSift taxonomic markers that correspond to the enriched taxa in Figure 2 are colored accordingly. Figure 4B is the same map as Figure 4A, but in this case, fragments are color-coded according to the presence of predicted proteins associated with metabolic functions of interest (Supplementary File S8). Five bins of metagenomic fragments with relatively homogenous taxonomic assignments were identified in this ESOM (Table 4). Unfortunately, none of the five bins contained any 16S rRNA genes, so direct links to the 16S rRNA gene amplicon data could not be made, and comparisons with PhyloSift taxonomic classifications of protein-encoding taxonomic markers are reported below instead.

ESOM Bin 1 (44% complete, 1% contamination) corresponds to the Methanobacteriaceae 16S rRNA sequences that were enriched in the ultrabasic springs (Figure 2) because 88% of its

555 sequences encoding taxonomic markers were classified as Methanobacteriaceae by PhyloSift
556 (Supplementary File S6). Furthermore, genomes belonging to genus *Methanobacterium* were the
557 most common source of best hits when sequences belonging to ESOM Bin 1 were competitively
558 recruited to a set of all methanogen genomes available from GenBank. ESOM Bin 1 sequences
559 were recruited to many other types of methanogens as well, indicating that this bin does not have
560 a single nearly identical whole-genome match in GenBank. ESOM Bin 1 includes many of the
561 methanogenesis-associated sequences in the combined-spring assembly (black borders in Figure
562 3), including genes involved in the utilization of formate, acetate, and methanol as substrates for
563 methanogenesis. Two of the genes encoding nitrogenase (*nifDK*) were also identified in this bin,
564 suggesting the potential for nitrogen fixation (Supplementary File S8), which is consistent with
565 the low levels of fixed nitrogen expected to be in the spring water (Cipolli et al., 2004).
566 Curiously, ESOM Bin 1 does not include any of the genes that encode formylmethanofuran
567 dehydrogenase, which catalyzes the first step of methanogenesis when carbon dioxide is the
568 carbon source (Figure 3). Although the absence of genes in an incomplete metagenomic bin
569 cannot be definitive, it is nevertheless interesting to observe that the only methanogens known to
570 lack all of the formylmethanofuran dehydrogenase genes (according to a comparison of KEGG
571 annotations on JGI's Integrated Microbial Genomes website) are members of genus
572 *Methanomasssiliicoccus*, which use methanol as their carbon source and cannot grow on carbon
573 dioxide.
574
575 ESOM Bin 2 (42% complete, 12% contamination) includes sequences that were classified as
576 Comamonadaceae (43% of taxonomic markers) and Xanthomonadaceae (27% of taxonomic
577 markers) (Supplementary File S6). Both of these bacterial families were also among the most

578 abundant taxa identified in the bacterial 16S rRNA data (Figure 2). Repeated attempts to
579 partition this bin to separate the Comamonadaceae and Xanthomonadaceae sequences were
580 unsuccessful, suggesting shared sequence compositions and/or mis-assembly. ESOM Bin 2
581 encodes RuBisCo (ribulose-1,5-bisphosphate carboxylase/oxygenase), the key enzyme in the
582 Calvin-Benson-Bassham pathway of carbon fixation. The predicted protein sequence of the large
583 subunit has 83% identities and 91% similarities with the Comamonadaceae RuBisCo sequence
584 identified in metagenomic sequences from a serpentinite-hosted ultrabasic spring at the
585 Tablelands Ophiolite in Newfoundland, Canada (Brazelton, Nelson & Schrenk, 2012). Nearly all
586 predicted proteins in the combined-spring assembly involved in the degradation of chlorinated
587 aromatic compounds are also found in ESOM Bin 2 (Supplementary File S7), and these
588 sequences have close matches in other Comamonadaceae genomes. This bin does not include
589 any sequences encoding a NiFe-hydrogenase, but there is at least one large contig in the
590 combined-spring assembly that encodes a NiFe-hydrogenase with 84% amino acid identities
591 with the group 1 NiFe-hydrogenase from *Serpentinomonas* strain H1, a Comamonadaceae isolate
592 from a serpentinite spring at The Cedars, California (Suzuki et al., 2014).
593
594 ESOM Bin 3 (28% complete, 2% contamination) is a collection of sequences with PhyloSift
595 classifications consistent with that of the Bacteroidetes uncultured aquatic group 'ML635J-40',
596 which was identified as the second-most abundant bacterial taxon in the ultrabasic springs
597 (Figure 2). 92% of all taxonomic markers in ESOM Bin 3 were classified as Bacteroidetes
598 (Supplementary File S6), although these sequences were somewhat evenly distributed among
599 several different families within the Bacteroidetes, consistent with the undetermined
600 phylogenetic placement of the ML635J-40 aquatic group (Nolla-Ardèvol, Strous & Tegetmeyer,

2015). This bin includes 50% of all combined-spring metagenomic sequences assigned to the FOAM pathway 'Pyruvate fermentation to acetate III' and is also rich in sugar transporters (Supplementary File S7). ESOM Bin 3 encodes multiple [FeFe]-hydrogenases with high similarity to those encoded by other Bacteroidetes including *Lentimicrobium saccharophilum* (GenBank GAP44922.1) and *Alistipes* sp. ZOR0009 (GenBank WP_047449271).

ESOM Bin 4 (48% complete, 6% contamination) corresponds to the Methylococcaceae 16S rRNA sequences identified in Figure 2. 60% of taxonomic markers in this bin were classified as family Methylococcaceae, and 82% were classified as order Methylococcales (Supplementary File S6). Bacteroidetes sequences comprised 12% of taxonomic markers in this bin. Predicted protein sequences for particulate and soluble methane monooxygenase (*pmoCAB*, K10944-46; *mmoXYBZDC*, K16157-16162) were prominent in ESOM Bin 4 (Supplementary File S8), and they shared 83-90% amino acid identities with sequences from *Methylobacter* and *Methylomicrobium* species. The most abundant Methylococcaceae 16S rRNA gene amplicon sequences in the ultrabasic springs are most similar to a clone recovered from a deep mine in South Africa (Blanco et al., 2014), which is most closely related to *Methylosoma difficile* (Rahalkar, Bussmann & Schink, 2007).

ESOM Bin 5 (60% complete, 6% contamination) appears to capture much of the genomic content associated with the Desulfovibrionales 16S rRNA sequences that dominate the BR2-spring-2012 metagenome. 83% of taxonomic markers in this bin were classified as Desulfovibrionales (Supplementary File S6). Furthermore, [NiFe]-hydrogenase sequences in ESOM Bin 5 have high similarity to other Desulfobivrionales genomes; for example, one

predicted protein sequence has 90% identity with the hydrogenase 2 large subunit of *Desulfonatronum lacustre* (Pikuta et al., 2003). The *Desulfonatronum* genus includes several alkaliphilic, hydrogen-oxidizing species, although there are no reports of *Desulfonatronum* growth above pH 10. In addition to [NiFe]-hydrogenase (*hydAB*; K18008, K00437), most of the combined-spring metagenomic sequence coverage of carbon monoxide dehydrogenase (*cooS*; K00198), acetyl-coA synthase (*acsB*; K14128), three subunits of acetyl-CoA decarbonylase/synthase complex (Cdh/ACDS; K00194, K00197, K00198), dissimilatory sulfite reductase (*dsrAB*; K11180-81), heterodisulfide reductase (*hdrABC*; K03388-K03390), and phosphate transacetylase (*pta*; K13788) were included in this bin (Supplementary File S8).

Metabolic activity assays

The activity of methanogenesis, methanotrophy, and oxidation of organic acids in ultrabasic springs was tested by measuring the conversion of ^{13}C-labeled carbon compounds during incubations of spring water samples at ambient temperatures (~20-25 ºC) in the dark. A single plus symbol in Table 5 indicates that the ^{13}C ratio of the product for that reaction was higher in all three live replicates compared to all three dead replicates, which had been filter-sterilized prior to incubation. Double plus symbols indicate that the live replicates were 10-60 ‰ heavier than the corresponding dead replicates, and triple plus symbols indicate that the live replicates were 100-700 ‰ heavier than the dead replicates.

Methanogenesis from acetate was detected in GOR34-spring3-2011. Unfortunately, insufficient material was collected from this spring in 2011 for high-quality sequencing studies, and subsequent metabolic activity experiments failed to detect methane production, most likely due

to technical difficulties with performing these experiments. Curiously, methane production was only detected from the carboxyl group (C1 atom) of acetate. Acetoclastic methanogenesis generates methane by reduction of the methyl group (C2 atom), so this result suggests that the carboxyl group of acetate was first oxidized to inorganic carbon (carbon dioxide, bicarbonate, or carbonate) and then utilized by methanogens. Indeed, oxidation of acetate to carbon dioxide was observed in GOR34-spring3-2011. Methane oxidation to carbon dioxide was also observed at least once in all three springs. Methane oxidation signals were greatly increased when incubations were exposed to oxygen (indicated by the third plus symbol in GOR34-spring1-2012 in Table 5) and were completely eliminated by addition of a reducing agent. This result strongly indicates that the measured methane oxidation was aerobic, although some activity was also detected even without oxygen exposure.

Discussion

Methanogenesis

This study of ultrabasic springs at the Voltri Massif provides the first metagenomic evidence for methanogens and methanogenesis pathways in a continental serpentinizing system. All genes required for methanogenesis are present in the combined-spring assembly (Figure 3), and many of these were collected in a single metagenomic bin corresponding to the Methanobacteriaceae taxa that are enriched in the ultrabasic springs (Figure 2). The combined-spring assembly includes the genetic potential for the utilization of carbon dioxide, formate, acetate, and methanol as methanogenic substrates. Some of the genes involved in carbon dioxide, formate, and acetate utilization were also found in genomic bins for non-methanogens and in the adjacent rivers, but the two proteins required for the utilization of methanol as a methanogenic substrate were unique

to the Methanobacteriaceae bin. Furthermore, the Methanobacteriaceae bin lacked all known genes for the utilization of carbon dioxide for methanogenesis. There are no previous examples of organisms within family Methanobacteriaceae that can use methanol as a substrate, so this result should be confirmed with additional genomic and experimental data. Nevertheless, methanol and formate are among the single carbon compounds predicted to be abiogenic products of serpentinization (Shock, 1992; Shock & Schulte, 1998; Seewald, Zolotov & McCollom, 2006), so these metagenomic data are consistent with the possibility that Methanobacteriaceae methanogens at this site are directly supported by subsurface serpentinization reactions.

We detected active methanogenesis in GOR34-spring3 (Table 5), which is assumed to have occurred at close to the in situ pH of the spring (pH 11.8). To our knowledge, this would represent the highest pH known to support methanogenesis, although no pH measurements were made during or after the incubation experiments. Methane produced during this experiment was derived from the carboxyl group (C1 atom) of acetate, and no methane was produced from the methyl group (C2 atom). This is contrary to the expectation for acetate-based methanogenesis, which produces methane by reduction of the methyl group. Furthermore, Methanobacteriaceae methanogens are not known to be able to use acetate as a substrate for methanogenesis. Therefore, it is more likely that the carboxyl group was oxidized to inorganic carbon (activity for which was also detected in this spring; see Table 5) and subsequently utilized for methanogenesis directly or indirectly. Unfortunately, methanogenesis from inorganic carbon (carbon dioxide, bicarbonate, or carbonate) was not measured during that experiment, and subsequent experiments failed to detect methanogenesis from any substrates. These results are

693 also not easily reconciled with the absence of genes required for the utilization of carbon dioxide
694 for methanogenesis in the Methanobacteriaceae ESOM bin. Thus, the genomic evidence for the
695 potential of formate, acetate, and methanol to support methanogenesis by Methanobacteriaceae
696 organisms remains only incompletely confirmed by metabolic activity experiments. Future
697 experimental studies designed to test the utilization of specific carbon compounds during active
698 methanogenesis at high pH would provide stronger evidence for the link between subsurface
699 carbon sources and methanogens in serpentinite springs.
700
701 Methanobacteriaceae (as well as orders Methanosarcinales and Methanocellales) have been
702 previously detected in travertine deposits formed by ultrabasic springs of the Voltri Massif
703 nearby, but not the same as, the springs described here (Quéméneur et al., 2015). The presence of
704 Methanobacteriaceae in the carbonate deposits raises the question of whether they primarily
705 inhabit surface or subsurface environments. Methanobacteriaceae and predicted proteins in the
706 methanogenesis pathway are abundant in the BR2 spring (Supplementary Files S5-S8), which is
707 not submerged by an overlying pool, does not appear to come into contact with surface travertine
708 deposits, and can be sampled directly from the spring's source (Figure 1). Therefore, the high
709 abundance of Methanobacteriaceae in this spring suggests a subsurface habitat for these
710 methanogens, although it is possible that these organisms could be active in both subsurface
711 springs and surface deposits where biofilms may create anoxic micro-environments.
712
713 *Hydrogen metabolism*
714 In addition to hydrogen gas consumption by methanogens, the Voltri Massif springs encode
715 abundant genes for hydrogen oxidation (Supplementary File S8). [NiFe]-hydrogenases

716 associated with Comamonadaceae and Desulfovibrionales, which dominate the bacterial
717 communities of the ultrabasic springs, were abundant in the metagenomic data from these
718 springs. The Comamonadaceae 16S rRNA, [NiFe]-hydrogenase, and Rubisco sequences are
719 highly similar to those reported for Comamonadaceae from other continental sites of
720 serpentinization including the Cabeço de Vide Aquifer (Tiago & Veríssimo, 2013), the
721 Tablelands Ophiolite (Brazelton, Nelson & Schrenk, 2012), The Cedars (Suzuki et al., 2014),
722 and the Coast Range Ophiolite Microbial Observatory (Crespo-Medina et al., 2014). The
723 presence of these organisms in alkaline springs appears to be a consistent indicator of
724 serpentinization (Schrenk, Brazelton & Lang, 2013), and the genus *Serpentinomonas* has been
725 proposed for them (Suzuki et al., 2014). *Serpentinomonas* 16S rRNA sequences have also been
726 recovered from surface carbonate deposits formed by a spring at the Voltri Massif near to those
727 reported in this study (Quéméneur et al., 2015), which is consistent with the interpretation that
728 these organisms live at the surface or the shallow subsurface where serpentinizing fluids mix
729 with the oxygenated atmosphere (Schrenk, Brazelton & Lang, 2013; Brazelton et al., 2013).
730
731 By contrast, the Desulfovibrionales represented by ESOM Bin 5 appear to be anaerobic,
732 hydrogen-utilizing, sulfate-reducing bacteria and may live in anoxic, deep subsurface habitats
733 underlying the springs. These bacteria, like the methanogens described above, were most
734 abundant in the BR2 spring, which also contains higher sulfide levels. Therefore, this spring
735 appears to provide more direct access to organisms that may have been recently active in anoxic,
736 methanogenic, and sulfate-reducing subsurface habitats.
737

738 [FeFe]-hydrogenase sequences were also abundant, and many were associated with the
739 unclassified Bacteroidetes group enriched in the ultrabasic springs (Figure 2). These genes are
740 typically involved in fermentative hydrogen production and appear to have a broad taxonomic
741 distribution in our study, as illustrated by their wide distribution in the ESOM (Figure 4B).
742 Clostridiales, which were also enriched in the springs, are expected to encode some of the
743 [FeFe]-hydrogenase sequences (e.g. Mei et al., 2014), but no ESOM bins with taxonomic
744 markers consistently classified as Clostridiales could be identified.
745
746 These results are consistent with previous observations of abundant hydrogenases in
747 metagenomic data from marine and continental serpentinite springs (Brazelton, Nelson &
748 Schrenk, 2012). Phylogenetic analyses of those sequences indicated that [NiFe]-hydrogenases
749 associated with Comamonadaceae and [FeFe]-hydrogenases associated with Firmicutes are
750 likely to be involved in hydrogen consumption and production, respectively. The co-occurrence
751 of Betaproteobacteria (often order Burkholderiales, including family Comamonadaceae) and
752 Firmicutes (in particular order Clostridiales) is a common feature of serpentinite springs and
753 perhaps the mixing of deep subsurface and surface fluids in general (reviewed by Schrenk,
754 Brazelton & Lang, 2013; see also Purkamo et al., 2016). This report contributes to that emerging
755 trend by providing another example of Comamonadaceae organisms with [NiFe]-hydrogenases
756 and also reports [NiFe]-hydrogenases in Desulfovibrionales and [FeFe]-hydrogenases in
757 Bacteroidetes in serpentinite springs for the first time.
758
759 *Methanotrophy*

760 Aerobic methane oxidation was detected in all three of the ultrabasic springs in this study (Table
761 5), and the Methylococcaceae family of aerobic methanotrophs was among the most enriched
762 taxa in the ultrabasic springs (Figure 2). The genomic bin corresponding to this family (ESOM
763 Bin 4) includes sequences encoding both the particulate and soluble forms of methane
764 monooxygenase (*pmoABC* and *mmoXYBZDC*). (Ward et al., 2004). Being aerobic, these bacteria
765 are unlikely to be active in anoxic subsurface environments. However, a completely surface-
766 exposed habitat also seems unlikely because they were not detected in a travertine deposit
767 formed by one of these springs (Quéméneur et al., 2015). Therefore, they may inhabit a shallow
768 subsurface mixing zone where low levels of oxygen reach methane-rich subsurface water,
769 similar to the inferred habitat for the aerobic, hydrogen-oxidizing Comamonadaceae described
770 above.
771
772 *Fermentation*
773 Multiple fermentation pathways were enriched in the ultrabasic springs, and fermentation genes
774 were widely distributed among metagenomic fragments in the ESOM (Figure 4). Candidate
775 Division OD1 (now called Parcubacteria) was enriched in the ultrabasic springs here and has also
776 been found in serpentinite springs at The Cedars that are expected to represent deeper,
777 subsurface habitats (Suzuki et al., 2013). These bacteria are expected to be involved with
778 fermentation (Hu et al., 2016), but we were unable to verify this due to the lack of any ESOM
779 bins with a strong representation of Candidate Division OD1. Additional work will be required
780 to better characterize the genomes of these organisms, which have been reported to be small and
781 lacking genes typically thought to be essential for basic cellular processes (Nelson & Stegen,
782 2015; Brown et al., 2015).

783

784 An uncultured group of Bacteroidetes known as 'ML635J-40 aquatic group' was also enriched in
785 the ultrabasic springs and was well-represented in the ESOM. The bin of metagenomic
786 sequences associated with this group (ESOM Bin 3) encoded proteins predicted to be involved in
787 pyruvate fermentation, sugar uptake, and hydrogen production. A member of ML635J-40 aquatic
788 group was recently identified in a pH 10 anaerobic reactor in which it was predicted to be
789 primarily responsible for hydrolysis of organic matter and supply of hydrogen gas to
790 methanogens (Nolla-Ardèvol, Strous & Tegetmeyer, 2015). The Bacteroidetes in the BR2 and
791 GOR34 ultrabasic springs could have a similar syntrophic relationship with methanogens,
792 although the source of hydrolyzable organic matter in this case is unclear. The ML635J-40
793 aquatic group was not detected in a previous study of surface travertine deposits (Quéméneur et
794 al., 2015), indicating that they do not inhabit a completely surface-exposed environment.
795 Although great care was taken to avoid surface contamination during field sampling, including
796 measures to sample the subsurface source of the spring rather than the overlying pool (see
797 Methods), it remains possible that the Bacteroidetes sequences in the GOR34 springs represent
798 pool-dwelling organisms that are dependent on terrestrial organic matter that falls into the
799 surface-exposed pools. This interpretation is consistent with the lower abundance of
800 Bacteroidetes in the BR2 spring (Supplementary File S6), which is not submerged by an
801 overlying pool (Figure 1). It is also possible that they inhabit a shallow subsurface transition
802 zone, similar to that inferred above for the Comamonadaceae and Methylococcaceae. Future
803 studies should investigate the carbon source for these organisms in order to determine whether
804 they are supported by compounds synthesized by subsurface serpentinization-associated
805 reactions.

806

807 *Formate and acetate metabolism*

808 Both formate and acetate are present at elevated concentrations in fluids venting from the
809 chimneys of the Lost City hydrothermal field (Lang et al., 2010), and formate is expected to be
810 an abiogenic product of serpentinization-associated reactions (Shock, 1992; Shock & Schulte,
811 1998; McCollom & Seewald, 2001, 2003; Seewald, Zolotov & McCollom, 2006). Metagenomic
812 sequences predicted to encode proteins associated with formate and acetate metabolism were
813 abundant (Supplementary File S8) and widespread among taxa (green and blue points,
814 respectively, in Figure 4B) in the combined-spring assembly, including the ESOM bins
815 representing Methylococcaceae and Desulfovibrionales (Table 4). The wide taxonomic
816 distribution of these genes (formate c-acetyltransferase, formate dehydrogenase, acetate kinase,
817 phosphotransacetylase, AMP-forming acetyl-CoA synthetase, and carbon monoxide
818 dehydrogenase / acetyl-CoA decarbonylase/synthase (Cdh/ACDS)) clearly indicates that they are
819 not restricted to methanogens in this system and that this ecosystem features a broad metabolic
820 potential for utilizing formate and acetate. We measured organic acid concentrations in Voltri
821 Massif springs in 2010, before the microbiological studies described here. Formate and acetate
822 were present at low but detectable concentrations (0.4 and 2.7 μM, respectively) at GOR34-
823 spring3 and were below detection in all other springs. It is unclear whether the low levels
824 indicate a lack of production or rapid consumption, but our initial metabolic activity assays have
825 demonstrated that acetate can be oxidized to carbon dioxide at in situ conditions characterized by
826 extremely high pH. Additional work is required to test whether the availability of formate and
827 acetate are correlated with the activity of specific genes, including those identified by this study.
828 Furthermore, future studies should attempt to demonstrate in the case of each group of organisms

reported here whether formate and acetate are consumed or produced, whether they are cycled intracellularly, and whether they are obtained from or released into the environment.

Conclusions

The potential significance of serpentinization-supported microbial ecosystems has been widely recognized since the discovery of the Lost City hydrothermal field (Kelley et al., 2001), but we are still in the early stages of characterizing the organisms and pathways that may benefit from the subsurface geochemical reactions associated with serpentinization. This study of ultrabasic springs at the Voltri Massif, Italy provides the first evidence for active methanogenesis and aerobic methanotrophy in continental serpentinite springs as well as genomic information about the specific organisms likely to be responsible for these processes. Both methanogenesis and methanotrophy were active at very high pH, potentially raising the upper pH limit known to support both processes. These potential activity experiments were not intended to quantify in situ rates, however, and do not preclude the possibility that the vast majority of methane in these springs is produced by non-biological processes. The data reported here also contribute additional metagenomic evidence for the importance of hydrogenases involved in both the consumption and production of hydrogen gas in serpentinite systems.

The unavailability of inorganic carbon at the extremely high pH of these springs is likely to be a limiting factor for autotrophic activity, and small organic compounds such as formate, acetate, and methanol may serve as the primary carbon sources in such ecosystems. The absence of genes in the Methanobacteriaceae bin required for the reduction of carbon dioxide may be a reflection of the lack of inorganic carbon in this system. Furthermore, we found evidence for inorganic

carbon as a substrate for methanogenesis only after it had been liberated by the oxidation of acetate. The genetic potential for metabolism of formate and acetate is widespread among taxa in this environment, which is consistent with an ecosystem being at least partially supported by organic carbon synthesized by subsurface serpentinization-associated reactions.

These genomic and metabolic data were obtained despite the extremely low biomass of these springs, suggesting that our results are the first clues into the few organisms capable of survival, and potentially growth, in the harsh conditions of the pH 12 springs. Because many of these organisms were only found in the spring waters and not in surface travertine deposits (Quéméneur et al., 2015), they are likely to represent inhabitants of subsurface environments in the serpentinite rocks underlying the springs. Many of these putative subsurface organisms may have been plucked from dense biofilm communities attached to serpentinites and then flushed out to the surface by the spring water. Residence times of ~700 years have been estimated for spring water in this system (Cipolli et al., 2004), suggesting that the subsurface communities inferred by this study may persist for long time periods in isolation from the surface. We anticipate that the metagenomic inventories and data analysis tools described here will provide a foundation for future studies to investigate how these organisms make a living in such unusual conditions and to test whether such ecosystems can be supported solely by subsurface serpentinization-associated reactions.

Acknowledgements

We are grateful to have had access to the excellent facilities of the ETH Stable Isotope Lab directed by Stefano Bernasconi. Esther Schwarzenbach, Melitza Crespo-Medina, Bridget Nelson,

and Elena Amador provided valuable assistance and fun conversations in the field. Hilary Morrison, Sharon Grim, Mitch Sogin, and Rick Colwell facilitated the early stages of this project through the Census of Deep Life.

References

Abubucker S., Segata N., Goll J., Schubert AM., Izard J., Cantarel BL., Rodriguez-Mueller B., Zucker J., Thiagarajan M., Henrissat B., White O., Kelley ST., Methé B., Schloss PD., Gevers D., Mitreva M., Huttenhower C. 2012. Metabolic reconstruction for metagenomic data and its application to the human microbiome. *PLoS Computational Biology* 8:e1002358. DOI: 10.1371/journal.pcbi.1002358.

Albert DB., Martens CS. 1997. Determination of low-molecular-weight organic acid concentrations in seawater and pore-water samples via HPLC. *Marine Chemistry* 56:27–37. DOI: 10.1016/S0304-4203(96)00083-7.

Alt JC., Schwarzenbach EM., Früh-Green GL., Shanks WC., Bernasconi SM., Garrido CJ., Crispini L., Gaggero L., Padrón-Navarta JA., Marchesi C. 2013. The role of serpentinites in cycling of carbon and sulfur: Seafloor serpentinization and subduction metamorphism. *Lithos* 178:40–54. DOI: 10.1016/j.lithos.2012.12.006.

Anders S., Pyl PT., Huber W. 2015. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31:166–9. DOI: 10.1093/bioinformatics/btu638.

Blanco Y., Rivas LA., García-Moyano A., Aguirre J., Cruz-Gil P., Palacín A., van Heerden E., Parro V. 2014. Deciphering the Prokaryotic Community and Metabolisms in South African Deep-Mine Biofilms through Antibody Microarrays and Graph Theory. *PLoS ONE* 9:e114180. DOI: 10.1371/journal.pone.0114180.

Blank JG., Green SJ., Blake D., Valley JW., Kita NT., Treiman A., Dobson PF. 2009. An alkaline spring system within the Del Puerto Ophiolite (California, USA): A Mars analog site. *Planetary and Space Science* 57:533–540. DOI: 10.1016/j.pss.2008.11.018.

Boisvert S., Raymond F., Godzaridis E., Laviolette F., Corbeil J. 2012. Ray Meta: scalable de novo metagenome assembly and profiling. *Genome Biology* 13:R122. DOI: 10.1186/gb-2012-13-12-r122.

Boulart C., Chavagnac V., Monnin C., Delacour A., Ceuleneer G., Hoareau G. 2013. Differences in gas venting from ultramafic-hosted warm springs: The example of Oman and Voltri ophiolites. *Ofioliti* 38:143–156. DOI: 10.4454/ofioliti.v38i2.423.

907 Bradley A., Hayes J., Summons R. 2009. Extraordinary ^{13}C enrichment of diether lipids at the
908 Lost City hydrothermal field indicates a carbon-limited ecosystem. *Geochimica et*
909 *Cosmochimica Acta* 73:102–118. DOI: 10.1016/j.gca.2008.10.005.

910 Brazelton WJ., Schrenk MO., Kelley DS., Baross JA. 2006. Methane- and sulfur-metabolizing
911 microbial communities dominate the Lost City hydrothermal field ecosystem. *Applied and*
912 *Environmental Microbiology* 72:6257–6270. DOI: 10.1128/AEM.00574-06.

913 Brazelton WJ., Mehta MP., Kelley DS., Baross JA. 2011. Physiological Differentiation within a
914 Single-Species Biofilm Fueled by Serpentinization. *mBio* 2:e00127-11. DOI:
915 10.1128/mBio.00127-11.

916 Brazelton WJ., Morrill PL., Szponar N., Schrenk MO. 2013. Bacterial communities associated
917 with subsurface geochemical processes in continental serpentinite springs. *Applied and*
918 *Environmental Microbiology* 79:3906–16. DOI: 10.1128/AEM.00330-13.

919 Brazelton WJ., Nelson B., Schrenk MO. 2012. Metagenomic evidence for H_2 oxidation and H_2
920 production by serpentinite-hosted subsurface microbial communities. *Frontiers in*
921 *Microbiology* 2:268. DOI: 10.3389/fmicb.2011.00268.

922 Brown CT., Hug LA., Thomas BC., Sharon I., Castelle CJ., Singh A., Wilkins MJ., Wrighton
923 KC., Williams KH., Banfield JF. 2015. Unusual biology across a group comprising more
924 than 15% of domain Bacteria. *Nature* 523:208–11. DOI: 10.1038/nature14486.

925 Bruni J., Canepa M., Chiodini G., Cioni R., Cipolli F., Longinelli A., Marini L., Ottonello G.,
926 Vetuschi M. 2002. Irreversible water–rock mass transfer accompanying the generation of
927 the neutral, Mg–HCO_3 and high-pH, Ca–OH spring waters of the Genova province, Italy.
928 *Applied Geochemistry* 17:455–474. DOI: 10.1016/S0883-2927(01)00113-5.

929 Cardace D., Hoehler T., McCollom T., Schrenk M., Carnevale D., Kubo M., Twing K. 2013.
930 Establishment of the Coast Range ophiolite microbial observatory (CROMO): drilling
931 objectives and preliminary outcomes. *Scientific Drilling* 16:45–55. DOI: 10.5194/sd-16-45-
932 2013.

933 Caspi R., Altman T., Billington R., Dreher K., Foerster H., Fulcher CA., Holland TA., Keseler
934 IM., Kothari A., Kubo A., Krummenacker M., Latendresse M., Mueller LA., Ong Q., Paley
935 S., Subhraveti P., Weaver DS., Weerasinghe D., Zhang P., Karp PD. 2014. The MetaCyc
936 database of metabolic pathways and enzymes and the BioCyc collection of
937 Pathway/Genome Databases. *Nucleic Acids Research* 42:D459-71. DOI:
938 10.1093/nar/gkt1103.

939 Cipolli F., Gambardella B., Marini L., Ottonello G. 2004. Geochemistry of high-pH waters from
940 serpentinites of the Gruppo di Voltri (Genova, Italy) and reaction path modeling of CO_2

941 sequestration in serpentinite aquifers. *Applied Geochemistry* 19:787–802. DOI:
942 10.1016/j.apgeochem.2003.10.007.

943 Cline JD. 1969. Spectrophotometric determination of hydrogen sulfide in natural waters.
944 *Limnology and Oceanography* 14:454–458. DOI: 10.4319/lo.1969.14.3.0454.

945 Colby J., Stirling DI., Dalton H. 1977. The soluble methane mono-oxygenase of *Methylococcus*
946 *capsulatus* (Bath). Its ability to oxygenate n-alkanes, n-alkenes, ethers, and alicyclic,
947 aromatic and heterocyclic compounds. *The Biochemical Journal* 165:395–402.

948 Consortium TU. 2015. UniProt: a hub for protein information. *Nucleic Acids Research* 43:D204–
949 D212. DOI: 10.1093/nar/gku989.

950 Crespo-Medina M., Twing KI., Kubo MDY., Hoehler TM., Cardace D., McCollom T., Schrenk
951 MO. 2014. Insights into environmental controls on microbial communities in a continental
952 serpentinite aquifer using a microcosm-based approach. *Frontiers in Microbiology* 5:604.
953 DOI: 10.3389/fmicb.2014.00604.

954 Darling AE., Jospin G., Lowe E., Matsen FA., Bik HM., Eisen JA. 2014. PhyloSift: phylogenetic
955 analysis of genomes and metagenomes. *PeerJ* 2:e243. DOI: 10.7717/peerj.243.

956 Dick GJ., Andersson AF., Baker BJ., Simmons SL., Thomas BC., Yelton AP., Banfield JF. 2009.
957 Community-wide analysis of microbial genome sequence signatures. *Genome Biology*
958 10:R85. DOI: 10.1186/gb-2009-10-8-r85.

959 Edgar RC., Haas, BJ., Clemente JC., Quince C., Knight R. 2011. UCHIME improves sensitivity
960 and speed of chimera detection. *Bioinformatics* 27:2194.

961 Ferry JG. 2010. How to make a living by exhaling methane. *Annual Review of Microbiology*
962 64:453–73. DOI: 10.1146/annurev.micro.112408.134051.

963 Fricke WF., Seedorf H., Henne A., Kruer M., Liesegang H., Hedderich R., Gottschalk G., Thauer
964 RK. 2006. The genome sequence of *Methanosphaera stadtmanae* reveals why this human
965 intestinal archaeon is restricted to methanol and H_2 for methane formation and ATP
966 synthesis. *Journal of Bacteriology* 188:642–658. DOI: 10.1128/JB.188.2.642-658.2006.

967 Gomez-Alvarez V., Teal TK., Schmidt TM. 2009. Systematic artifacts in metagenomes from
968 complex microbial communities. *The ISME Journal* 3:1314–7. DOI: 10.1038/ismej.2009.72.

969 Hu P., Tom L., Singh A., Thomas BC., Baker BJ., Piceno YM., Andersen GL., Banfield JF. 2016.
970 Genome-Resolved Metagenomic Analysis Reveals Roles for Candidate Phyla and Other
971 Microbial Community Members in Biogeochemical Transformations in Oil Reservoirs.
972 *mBio* 7:e01669-15. DOI: 10.1128/mBio.01669-15.

973 Huber JA., Butterfield DA., Baross JA. 2002. Temporal changes in archaeal diversity and
974 chemistry in a mid-ocean ridge subseafloor habitat. *Applied and Environmental*
975 *Microbiology* 68:1585–1594. DOI: 10.1128/AEM.68.4.1585–1594.2002.

976 Huse SM., Young VB., Morrison HG., Antonopoulos DA., Kwon J., Dalal S., Arrieta R., Hubert
977 NA., Shen L., Vineis JH., Koval JC., Sogin ML., Chang EB., Raffals LE. 2014a.
978 Comparison of brush and biopsy sampling methods of the ileal pouch for assessment of
979 mucosa-associated microbiota of human subjects. *Microbiome* 2:5. DOI: 10.1186/2049-
980 2618-2-5.

981 Huse SM., Mark Welch DB., Voorhis A., Shipunova A., Morrison HG., Eren A., Sogin ML.
982 2014b. VAMPS: a website for visualization and analysis of microbial population structures.
983 *BMC Bioinformatics* 15:41. DOI: 10.1186/1471-2105-15-41.

984 Hyatt D., Chen G-L., LoCascio PF., Land ML., Larimer FW., Hauser LJ. 2010. Prodigal:
985 prokaryotic gene recognition and translation initiation site identification. *BMC*
986 *Bioinformatics* 11:119. DOI: 10.1186/1471-2105-11-119.

987 Kelley DS., Karson JA., Blackman DK., Fruh-Green GL., Butterfield DA., Lilley MD., Olson
988 EJ., Schrenk MO., Roe KK., Lebon GT., Rivizzigno P. 2001. An off-axis hydrothermal vent
989 field near the Mid-Atlantic Ridge at 30^{o} N. *Nature* 412:145–149. DOI: 10.1038/35084000.

990 Kelley DS., Karson JA., Früh-Green GL., Yoerger DR., Shank TM., Butterfield DA., Hayes JM.,
991 Schrenk MO., Olson EJ., Proskurowski G., Jakuba M., Bradley A., Larson B., Ludwig K.,
992 Glickson D., Buckman K., Bradley AS., Brazelton WJ., Roe K., Elend MJ., Delacour A.,
993 Bernasconi SM., Lilley MD., Baross JA., Summons RE., Sylva SP. 2005. A serpentinite-
994 hosted ecosystem: the Lost City hydrothermal field. *Science* 307:1428–34. DOI:
995 10.1126/science.1102556.

996 Kohl L., Cumming E., Cox A., Rietze A., Morrissey L., Lang SQ., Richter A., Suzuki S.,
997 Nealson KH., Morrill PL. 2016. Exploring the metabolic potential of microbial
998 communities in ultra-basic, reducing springs at The Cedars, CA, USA: Experimental
999 evidence of microbial methanogenesis and heterotrophic acetogenesis. *Journal of*
1000 *Geophysical Research: Biogeosciences* 121:1203–1220. DOI: 10.1002/2015JG003233.

1001 Lang SQ., Butterfield DA., Schulte M., Kelley DS., Lilley MD. 2010. Elevated concentrations of
1002 formate, acetate and dissolved organic carbon found at the Lost City hydrothermal field.
1003 *Geochimica et Cosmochimica Acta* 74:941–952. DOI: 10.1016/j.gca.2009.10.045.

1004 Lang SQ., Früh-Green GL., Bernasconi SM., Lilley MD., Proskurowski G., Méhay S.,
1005 Butterfield DA. 2012. Microbial utilization of abiogenic carbon and hydrogen in a
1006 serpentinite-hosted system. *Geochimica et Cosmochimica Acta* 92:82–99. DOI:
1007 10.1016/j.gca.2012.06.006.

1008 Langmead B., Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods*
1009 9:357–9. DOI: 10.1038/nmeth.1923.

1010 Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads.
1011 *EMBnet.journal* 17:10. DOI: 10.14806/ej.17.1.200.

1012 McCollom T., Seewald J. 2001. A reassessment of the potential for reduction of dissolved CO_2
1013 to hydrocarbons during serpentinization of olivine. *Geochimica et Cosmochimica Acta*
1014 65:3769–3778. DOI: 10.1016/S0016-7037(01)00655-X.

1015 McCollom TM., Seewald JS. 2003. Experimental study of the hydrothermal reactivity of organic
1016 acids and acid anions: II. Acetic acid, acetate, and valeric acid. *Geochimica et*
1017 *Cosmochimica Acta* 67:3645–3664. DOI: 10.1016/S0016-7037(03)00135-2.

1018 McMurdie PJ., Holmes S. 2013. phyloseq: an R package for reproducible interactive analysis
1019 and graphics of microbiome census data. *PloS ONE* 8:e61217. DOI:
1020 10.1371/journal.pone.0061217.

1021 McMurdie PJ., Holmes S. 2014. Waste not, want not: why rarefying microbiome data is
1022 inadmissible. *PLoS Computational Biology* 10:e1003531. DOI:
1023 10.1371/journal.pcbi.1003531.

1024 Méhay S., Früh-Green G., Lang S., Bernasconi S., Brazelton W., Schrenk M., Schaeffer P.,
1025 Adam P. 2013. Record of archaeal activity at the serpentinite-hosted Lost City
1026 Hydrothermal Field. *Geobiology* 11:570–592. DOI: 10.1111/gbi.12062.

1027 Mei N., Zergane N., Postec A., Erauso G., Ollier A., Payri C., Pelletier B., Fardeau ML., Ollivier
1028 B., Quéméneur M. 2014. Fermentative hydrogen production by a new alkaliphilic
1029 Clostridium sp. (strain PROH2) isolated from a shallow submarine hydrothermal chimney
1030 in Prony Bay, New Caledonia. *International Journal of Hydrogen Energy* 39:19465–19473.
1031 DOI: 10.1016/j.ijhydene.2014.09.111.

1032 Morrill PL., Kuenen JG., Johnson OJ., Suzuki S., Rietze A., Sessions AL., Fogel ML., Nealson
1033 KH. 2013. Geochemistry and geobiology of a present-day serpentinization site in California:
1034 The Cedars. *Geochimica et Cosmochimica Acta* 109:222–240. DOI:
1035 10.1016/j.gca.2013.01.043.

1036 Morrill PL., Brazelton WJ., Kohl L., Rietze A., Miles SM., Kavanagh H., Schrenk MO., Ziegler
1037 SE., Lang SQ. 2014. Investigations of potential microbial methanogenic and carbon
1038 monoxide utilization pathways in ultra-basic reducing springs associated with present-day
1039 continental serpentinization: the Tablelands, NL, CAN. *Frontiers in Microbiology* 5:613.
1040 DOI: 10.3389/fmicb.2014.00613.

1041 Nelson MC., Morrison HG., Benjamino J., Grim SL., Graf J. 2014. Analysis, Optimization and
1042 Verification of Illumina-Generated 16S rRNA Gene Amplicon Surveys. *PLoS ONE*
1043 9:e94249. DOI: 10.1371/journal.pone.0094249.

1044 Nelson WC., Stegen JC. 2015. The reduced genomes of Parcubacteria (OD1) contain signatures
1045 of a symbiotic lifestyle. *Frontiers in Microbiology* 6:713. DOI: 10.3389/fmicb.2015.00713.

1046 Nolla-Ardèvol V., Strous M., Tegetmeyer HE. 2015. Anaerobic digestion of the microalga
1047 *Spirulina* at extreme alkaline conditions: biogas production, metagenome, and
1048 metatranscriptome. *Frontiers in Microbiology* 6:597. DOI: 10.3389/fmicb.2015.00597.

1049 Ogata H., Goto S., Sato K., Fujibuchi W., Bono H., Kanehisa M. 1999. KEGG: Kyoto
1050 Encyclopedia of Genes and Genomes. *Nucleic Acids Research* 27:29–34. DOI:
1051 10.1093/nar/27.1.29.

1052 Parks DH., Imelfort M., Skennerton CT., Hugenholtz P., Tyson GW. 2015. CheckM: assessing
1053 the quality of microbial genomes recovered from isolates, single cells, and metagenomes.
1054 *Genome Research* 25:1043–1055. DOI: 10.1101/gr.186072.114.

1055 Pikuta E V., Hoover RB., Bej AK., Marsic D., Whitman WB., Cleland D., Krader P. 2003.
1056 *Desulfonatronum thiodismutans* sp. nov., a novel alkaliphilic, sulfate-reducing bacterium
1057 capable of lithoautotrophic growth. *International Journal of Systematic and Evolutionary*
1058 *Microbiology* 53:1327–32. DOI: 10.1099/ijs.0.02598-0.

1059 Prestat E., David MM., Hultman J., Taş N., Lamendella R., Dvornik J., Mackelprang R., Myrold
1060 DD., Jumpponen A., Tringe SG., Holman E., Mavromatis K., Jansson JK. 2014. FOAM
1061 (Functional Ontology Assignments for Metagenomes): a Hidden Markov Model (HMM)
1062 database with environmental focus. *Nucleic Acids Research* 42:e145. DOI:
1063 10.1093/nar/gku702.

1064 Proskurowski G., Lilley MD., Seewald JS., Früh-Green GL., Olson EJ., Lupton JE., Sylva SP.,
1065 Kelley DS. 2008. Abiogenic hydrocarbon production at Lost City Hydrothermal Field.
1066 *Science* 319:604–7. DOI: 10.1126/science.1151194.

1067 Pruesse E., Peplies J., Glöckner FO. 2012. SINA: accurate high-throughput multiple sequence
1068 alignment of ribosomal RNA genes. *Bioinformatics* 28:1823–9. DOI:
1069 10.1093/bioinformatics/bts252.

1070 Purkamo L., Bomberg M., Kietäväinen R., Salavirta H., Nyyssönen M., Nuppunen-Puputti M.,
1071 Ahonen L., Kukkonen I., Itävaara M. 2016. Microbial co-occurrence patterns in deep
1072 Precambrian bedrock fracture fluids. *Biogeosciences* 13:3091–3108. DOI: 10.5194/bg-13-
1073 3091-2016.

1074 Quéméneur M., Bes M., Postec A., Mei N., Hamelin J., Monnin C., Chavagnac V., Payri C.,
1075 Pelletier B., Guentas-Dombrowsky L., Gérard M., Pisapia C., Gérard E., Ménez B., Ollivier

B., Erauso G. 2014. Spatial distribution of microbial communities in the shallow submarine alkaline hydrothermal field of the Prony Bay, New Caledonia. *Environmental Microbiology Reports* 33. DOI: 10.1111/1758-2229.12184.

Quéméneur M., Palvadeau A., Postec A., Monnin C., Chavagnac V., Ollivier B., Erauso G. 2015. Endolithic microbial communities in carbonate precipitates from serpentinite-hosted hyperalkaline springs of the Voltri Massif (Ligurian Alps, Northern Italy). *Environmental Science and Pollution Research* 22:13613–24. DOI: 10.1007/s11356-015-4113-7.

Rahalkar M., Bussmann I., Schink B. 2007. *Methylosoma difficile* gen. nov., sp. nov., a novel methanotroph enriched by gradient cultivation from littoral sediment of Lake Constance. *International Journal of Systematic and Evolutionary Microbiology* 57:1073–1080. DOI: 10.1099/ijs.0.64574-0.

Robinson MD., McCarthy DJ., Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26:139–40. DOI: 10.1093/bioinformatics/btp616.

Sánchez-Murillo R., Gazel E., Schwarzenbach EM., Crespo-Medina M., Schrenk MO., Boll J., Gill BC. 2014. Geochemical evidence for active tropical serpentinization in the Santa Elena Ophiolite, Costa Rica: An analog of a humid early Earth? *Geochemistry Geophysics Geosystems* 15:1783–1800. DOI: 10.1002/2013GC005213.

Schloss PD., Westcott SL., Ryabin T., Hall JR., Hartmann M., Hollister EB., Lesniewski RA., Oakley BB., Parks DH., Robinson CJ., Sahl JW., Stres B., Thallinger GG., Van Horn DJ., Weber CF. 2009. Introducing mothur: Open Source, Platform-independent, Community-Supported Software for Describing and Comparing Microbial Communities. *Applied and Environmental Microbiology* 75:7537–7541. DOI: 10.1128/AEM.01541-09.

Schloss PD., Gevers D., Westcott SL. 2011. Reducing the Effects of PCR Amplification and Sequencing Artifacts on 16S rRNA-Based Studies. *PLoS ONE* 6:e27310. DOI: 10.1371/journal.pone.0027310.

Schrenk MO., Kelley DS., Bolton SA., Baross JA. 2004. Low archaeal diversity linked to subseafloor geochemical processes at the Lost City Hydrothermal Field, Mid-Atlantic Ridge. *Environmental Microbiology* 6:1086–1095. DOI: 10.1111/j.1462-2920.2004.00650.x.

Schrenk MO., Brazelton WJ., Lang SQ. 2013. Serpentinization, Carbon, and Deep Life. *Reviews in Mineralogy and Geochemistry* 75:575–606. DOI: 10.2138/rmg.2013.75.18.

Schwarzenbach EM., Lang SQ., Früh-Green GL., Lilley MD., Bernasconi SM., Méhay S. 2013. Sources and cycling of carbon in continental, serpentinite-hosted alkaline springs in the Voltri Massif, Italy. *Lithos* 177:226–244. DOI: 10.1016/j.lithos.2013.07.009.

Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–9. DOI: 10.1093/bioinformatics/btu153.

Seewald JS., Zolotov MY., McCollom T. 2006. Experimental investigation of single carbon compounds under hydrothermal conditions. *Geochimica et Cosmochimica Acta* 70:446–460. DOI: 10.1016/j.gca.2005.09.002.

Shock E. 1992. Stability of peptides in high-temperature aqueous solutions. *Geochimica et Cosmochimica Acta* 56:3481–3491. DOI: 10.1016/0016-7037(92)90392-V.

Shock EL., Schulte MD. 1998. Organic synthesis during fluid mixing in hydrothermal systems. *Journal of Geophysical Research* 103:28,513-28,527. DOI: 199810.1029/98JE02142.

Sogin ML., Morrison HG., Huber JA., Mark Welch D., Huse SM., Neal PR., Arrieta JM., Herndl GJ. 2006. Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proceedings of the National Academy of Sciences of the United States of America* 103:12115–12120. DOI: 10.1073/pnas.0605127103.

Suzuki S., Ishii S., Wu A., Cheung A., Tenney A., Wanger G., Kuenen JG., Nealson KH. 2013. Microbial diversity in The Cedars, an ultrabasic, ultrareducing, and low salinity serpentinizing ecosystem. *Proceedings of the National Academy of Sciences of the United States of America* 110:15336–41. DOI: 10.1073/pnas.1302426110.

Suzuki S., Kuenen JG., Schipper K., van der Velde S., Ishii S., Wu A., Sorokin DY., Tenney A., Meng X., Morrill PL., Kamagata Y., Muyzer G., Nealson KH. 2014. Physiological and genomic features of highly alkaliphilic hydrogen-utilizing Betaproteobacteria from a continental serpentinizing site. *Nature Communications* 5:3900. DOI: 10.1038/ncomms4900.

Tiago I., Veríssimo A. 2013. Microbial and functional diversity of a subterrestrial high pH groundwater associated to serpentinization. *Environmental Microbiology* 15:1687–1706. DOI: 10.1111/1462-2920.12034.

Topçuoğlu BD., Stewart LC., Morrison HG., Butterfield DA., Huber JA., Holden JF. 2016. Hydrogen limitation and syntrophic growth among natural assemblages of thermophilic methanogens at deep-sea hydrothermal vents. *Frontiers in Microbiology* 7:1240. DOI: 10.3389/fmicb.2016.01240.

Ultsch A., Mörchen F. 2005. *ESOM-Maps: tools for clustering, visualization, and classification with Emergent SOM*. Technical Report Dept. of Mathematics and Computer Science, University of Marburg, Germany, No. 46.

Wang DT., Gruen DS., Lollar BS., Hinrichs K., Stewart LC., Holden JF., Hristov AN., Pohlman JW., Morrill PL., Könneke M., Delwiche KB., Reeves EP., Sutcliffe N., Ritter DJ., Seewald JS., Jennifer C., Hemond HF., Kubo MD., Hoehler TM., Ono S. 2015. Nonequilibrium

clumped isotope signals in microbial methane. *Science* 348:428–431. DOI: 10.1126/science.aaa4326.

Ward N., Larsen Ø., Sakwa J., Bruseth L., Khouri H., Durkin AS., Dimitrov G., Jiang L., Scanlan D., Kang KH., Lewis M., Nelson KE., Methé B., Wu M., Heidelberg JF., Paulsen IT., Fouts D., Ravel J., Tettelin H., Ren Q., Read T., DeBoy RT., Seshadri R., Salzberg SL., Jensen HB., Birkeland NK., Nelson WC., Dodson RJ., Grindhaug SH., Holt I., Eidhammer I., Jonasen I., Vanaken S., Utterback T., Feldblyum T V., Fraser CM., Lillehaug JR., Eisen JA. 2004. Genomic Insights into Methanotrophy: The Complete Genome Sequence of *Methylococcus capsulatus* (Bath). *PLoS Biology* 2:e303. DOI: 10.1371/journal.pbio.0020303.

Woycheese KM., Meyer-Dombard DR., Cardace D., Argayosa AM., Arcilla CA. 2015. Out of the dark: transitional subsurface-to-surface microbial diversity in a terrestrial serpentinizing seep (Manleluag, Pangasinan, the Philippines). *Frontiers in Microbiology* 6:44. DOI: 10.3389/fmicb.2015.00044.

Ye Y., Doak TG. 2009. A Parsimony Approach to Biological Pathway Reconstruction/Inference for Genomes and Metagenomes. *PLoS Computational Biology* 5:e1000465. DOI: 10.1371/journal.pcbi.1000465.

Figure Captions

Figure 1. Ultrabasic springs of the Voltri Massif in Italy were sampled at two locations: A) GOR34, including two springs and an adjacent river and B) BR2, including one spring and an adjacent river. All water samples were collected by peristaltic pumping through a 0.2 μm filter cartridge. Photo credits: W.J. Brazelton (A) and B. Nelson (B).

Figure 2. Taxonomic classifications of bacterial (**A**) and archaeal (**B**) operational taxonomic units (OTUs, defined as unique 16S rRNA sequences) that were identified as significantly enriched in the ultrabasic serpentinite spring GOR34-spring1 compared to the adjacent river. Each data point in the plots represents one OTU's mean abundance across all samples in the analysis and its differential abundance in the comparison GOR34-spring1 versus GOR34-river.

Mean abundance is reported as $\log_2$ counts per million; a value of 16 corresponds to 23,000 sequences for that OTU. Red data points represent OTUs with differential abundances that are significantly enriched (false discovery rate < 0.05) in either GOR34-spring1 or GOR34-river. The relative abundances of these significantly enriched OTUs are reported as a percentage of all significantly enriched sequences and grouped into taxonomic classifications at the family level in the colored bar charts. Sections of the bar chart were labeled with the corresponding family where possible, and the abbreviations are defined by bold font in the right-hand legends.

Figure 3. Diagram of methanogenesis pathways from carbon dioxide, formate, acetate, methanol, and methylamines with associated protein homologs identified with KEGG IDs. Green-highlighted proteins are predicted to occur in the combined-spring metagenomic assembly, and yellow-highlighted proteins are predicted to occur in both the combined-spring and combined-river assemblies. The black border indicates proteins identified in the Methanobacteriaceae bin (ESOM Bin 1 in **Figure 4** and **Table 4**). The diagram is modified from (Ferry, 2010).

Figure 4. An emergent self-organizing map (ESOM) constructed with the combined-spring metagenomic assembly. Each point represents a fragment of metagenomic assembly, and points are arranged in space according to the similarities of their tetranucleotide compositions. Dark gray areas are valleys where highly similar sequences cluster together. Light areas of the map are high topographic ridges that separate dissimilar sequences. The same ESOM is shown here twice: metagenomic fragments are color-coded in A) to indicate PhyloSift classifications of taxonomic markers and in B) to show which fragments are predicted to encode protein functions of

particular interest to this study. Based on this information, five bins of sequences were selected from the map for further evaluation (**Table 4**).

Table 1(on next page)

Chemical characteristics of ultrabasic springs and adjacent rivers of the Voltri Massif, Italy

Chemical characteristics of ultrabasic springs and adjacent rivers of the Voltri Massif, Italy

Table 1. Chemical characteristics of ultrabasic springs and adjacent rivers of the Voltri Massif, Italy.

Sample Name	pH	E_h (mV)	CH_4 (µM)	H_2 (µM)	DIC (µM)	Sulfate (µM)	Sulfide (µM)
BR2-spring-2012	12.1	-195	689	0.5	29	6.5	16.5
BR2-spring-2013	12.3	-75	733	1.8	7.8	nm	16.2
BR2-river-2012	8.1	450	bdl	0.2	3031	130	bdl
BR2-river-2013	8.0	311	bdl	bdl	2678	nm	bdl
GOR34-spring1-2011	12.2	-60	155	3.9	14	3.0	bdl
GOR34-spring1-2013	12.3	-156	213	9.2	17.2	nm	18.6
GOR34-spring3-2012	11.8	-202	201	26.8	29	bdl	17.7
GOR34-river-2012	9.3	360	0.1	bdl	1752	49	6.3
GOR34-river-2013	9.5	240	bdl	bdl	1670	nm	bdl

nm = not measured; bdl = below detection limit

Table 2(on next page)

Water samples collected from ultrabasic springs for metagenomic analyses.

Water samples collected from ultrabasic springs for metagenomic analyses.

Table 2. Water samples collected from ultrabasic springs for metagenomic analyses.

Sample Name	Sample Date	Volume Filtered (L)	ng DNA per L water	Cells per mL[1]	Bact: Arch ratio	Bacterial 16S rDNA Sequences	Archaeal 16S rDNA Sequences	Metagenomic Sequences[2]
BR2-spring-2012	30-Jun-2012	527	0.9	7.6E+02	1.6	nm	nm	13,774,495
BR2-spring-2013	22-Aug-2013	82	0.1	6.9E+03	nm	nm	nm	8,047,152
BR2-river-2012a	30-Jun-2012	1	136	2.4E+04	nm	52,045	36,474	17,953,252
BR2-river-2012b	30-Jun-2012	2.2	214	2.8E+04	0.8	45,787	29,855	nm
BR2-river-2013	22-Aug-2013	10	69	1.7E+04	5	nm	nm	21,308,793
GOR34-spring1-2011	18-Oct-2011	5	47	2.3E+04	nm	15,056	724[3]	nm
GOR34-spring1-2013a	24-Aug-2013	12	41	2.2E+04	270	73,829	267[3]	nm
GOR34-spring1-2013b	24-Aug-2013	35	4	1.4E+04	10	107,493	101,232	27,882,181
GOR34-spring3-2012	1-Jul-2012	24	< 0.1	3.1E+03	nm	nm	nm	14,967,449
GOR34-river-2012a	1-Jul-2012	2	274	2.4E+04	6800	71,285	33,999	22,118,087
GOR34-river-2012b	1-Jul-2012	2	294	7.1E+04	0.7	79,012	115[3]	nm
GOR34-river-2013a	24-Aug-2013	3.3	122	5.3E+04	16	91,952	94,369	nm
GOR34-river-2013b	24-Aug-2013	9.2	56	4.2E+04	48	80,683	851[3]	36,276,889

nm = not measured

[1]*Cell concentrations are reported as the mean of three field replicates for each sample.*

[2]*Metagenomic sequences are reported as the number of quality-filtered read pairs.*

[3]*Samples with very low numbers of sequences were not included in any downstream analyses.*

Table 3(on next page)

Assembly and annotation statistics for the pooled metagenomic assemblies and the five ESOM bins.

Assembly and annotation statistics for the pooled metagenomic assemblies and the five ESOM bins.

Table 3. Assembly and annotation statistics for the pooled metagenomic assemblies and the five ESOM bins.

	combined-spring	**combined-river**	**ESOM Bin 1**	**ESOM Bin 2**	**ESOM Bin 3**	**ESOM Bin 4**	**ESOM Bin 5**
# contigs (>= 500 bp)	55,894	40,355	228	157	123	169	88
# contigs (>= 25000 bp)	122	186	7	0	11	9	6
Total length (>= 500 bp)	89,619,646	73,320,758	1,261,593	1,129,542	1,117,180	1,686,028	1,124,653
Total length (>= 25000 bp)	4,783,632	9,953,584	330,655	0	342,191	277,131	191,679
Largest contig (bp)	99,214	304,939	93,351	24,214	37,024	41,177	34,915
N50 (length of median contig)	2,367	3,515	4,908	9,513	12,857	15,069	16,218
Predicted genes	108,150	91,693	1,280	974	820	1,453	965
Predicted protein-encoding genes	104,651	89,606	1,256	963	805	1,423	933
Predicted rRNA-encoding genes	77	67	0	0	0	0	0
Predicted tRNA-encoding genes	2,267	1,349	16	11	12	20	27
Predicted other RNA-encoding genes	1,155	671	8	0	3	10	5
Predicted proteins annotated with KEGG[1]	71,856	67,170	1,014	874	714	1,271	855
Predicted proteins annotated with UniprotKB	34	19	0	0	0	0	0
Predicted proteins annotated with Pfam	482	367	4	7	2	3	3
Predicted proteins annotated with HAMAP	58	30	0	0	0	0	0
Predicted proteins annotated with CLUSTERS	212	151	1	0	0	1	0
Predicted proteins with no annotation	32,009	21,869	237	82	89	148	75

[1]*Predicted proteins were first mapped to the KEGG database, and any remaining proteins without a predicted function were then mapped to the other databases.*

Table 4(on next page)

Predicted metabolic pathways significantly enriched in ultrabasic spring metagenomes compared to adjacent rivers.

Predicted metabolic pathways significantly enriched in ultrabasic spring metagenomes compared to adjacent rivers.

Table 4. Characteristics of ESOM bins identified in **Figure 4**.

	Completeness	Contamination	Strain heterogeneity	Taxonomy	Metabolic Potential
Bin 1	44.11%	1.2%	75%	Methanobacteriaceae	Methanogenesis; nitrogen fixation
Bin 2	42.16%	11.72%	3.23%	Comamonadaceae / Xanthomonadaceae	Carbon fixation (Rubisco); organic carbon degradation
Bin 3	27.59%	1.72%	0%	Bacteroidetes	Hydrogen production; fermentation
Bin 4	47.98%	6.24%	34.88%	Methylococcaceae	Methylotrophy
Bin 5	60.34%	6.03%	100%	Desulfovibrionales	Hydrogen oxidation; dissimilatory sulfate reduction

Table 5(on next page)

Characteristics of ESOM bins identified in Figure 4.

Characteristics of ESOM bins identified in **Figure 4**.

Table 5. Detection of metabolic activities in incubation experiments with spring water.

	BR2-spring-2011	**BR2-spring-2012**	**GOR34-spring1-2010**	**GOR34-spring1-2012**	**GOR34-spring3-2011**	**GOR34-spring3-2012**
CO_2 --> CH_4	-	-	-	-	nm	-
formate --> CH_4	-	-	-	-	nm	-
acetate-C_1 --> CH_4	-	-	-	-	+++	-
acetate-C_2 --> CH_4	-	-	-	-	-	-
acetate-C_1 --> CO_2	-	nm	++	nm	++	nm
acetate-C_2 --> CO_2	-	nm	+	nm	-	nm
propionate --> CO_2	+	nm	-	nm	+	nm
CH_4 --> CO_2	++	-	+++	++(+)*	++	-

nm = not measured

** Third (+) symbol indicates stronger signal when exposed to oxygen.*

Figure 1

Ultrabasic springs of the Voltri Massif in Italy were sampled at two locations.

A) GOR34, including two springs and an adjacent river and B) BR2, including one spring and an adjacent river. All water samples were collected by peristaltic pumping through a 0.2 µm filter cartridge. Photo credits: W.J. Brazelton (A) and B. Nelson (B).

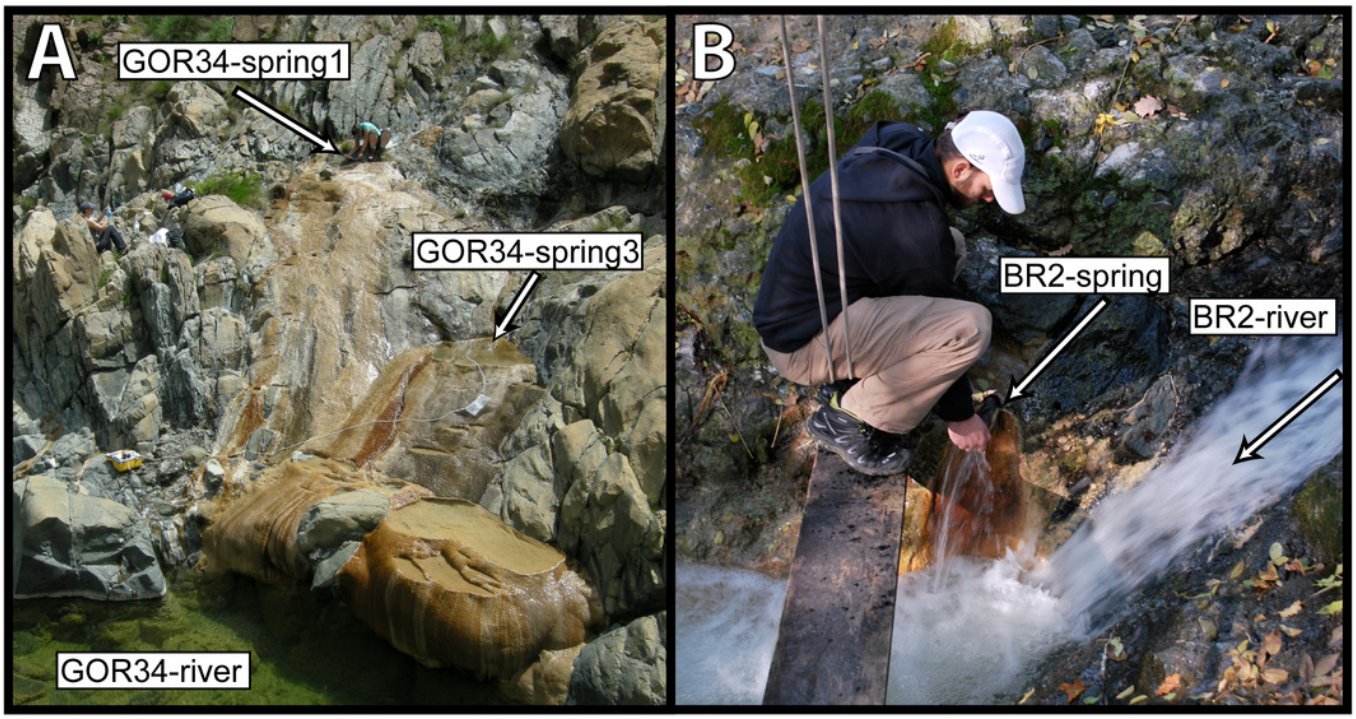


Figure 2

Bacterial and archaeal sequences enriched in ultrabasic springs.

Taxonomic classifications of bacterial (**A**) and archaeal (**B**) operational taxonomic units (OTUs, defined as unique 16S rRNA sequences) that were identified as significantly enriched in the ultrabasic serpentinite spring GOR34-spring1 compared to the adjacent river. Each data point in the plots represents one OTU's mean abundance across all samples in the analysis and its differential abundance in the comparison GOR34-spring1 versus GOR34-river. Mean abundance is reported as $\log_2$ counts per million; a value of 16 corresponds to 23,000 sequences for that OTU. Red data points represent OTUs with differential abundances that are significantly enriched (false discovery rate < 0.05) in either GOR34-spring1 or GOR34-river. The relative abundances of these significantly enriched OTUs are reported as a percentage of all significantly enriched sequences and grouped into taxonomic classifications at the family level in the colored bar charts. Sections of the bar chart were labeled with the corresponding family where possible, and the abbreviations are defined by bold font in the right-hand legends.

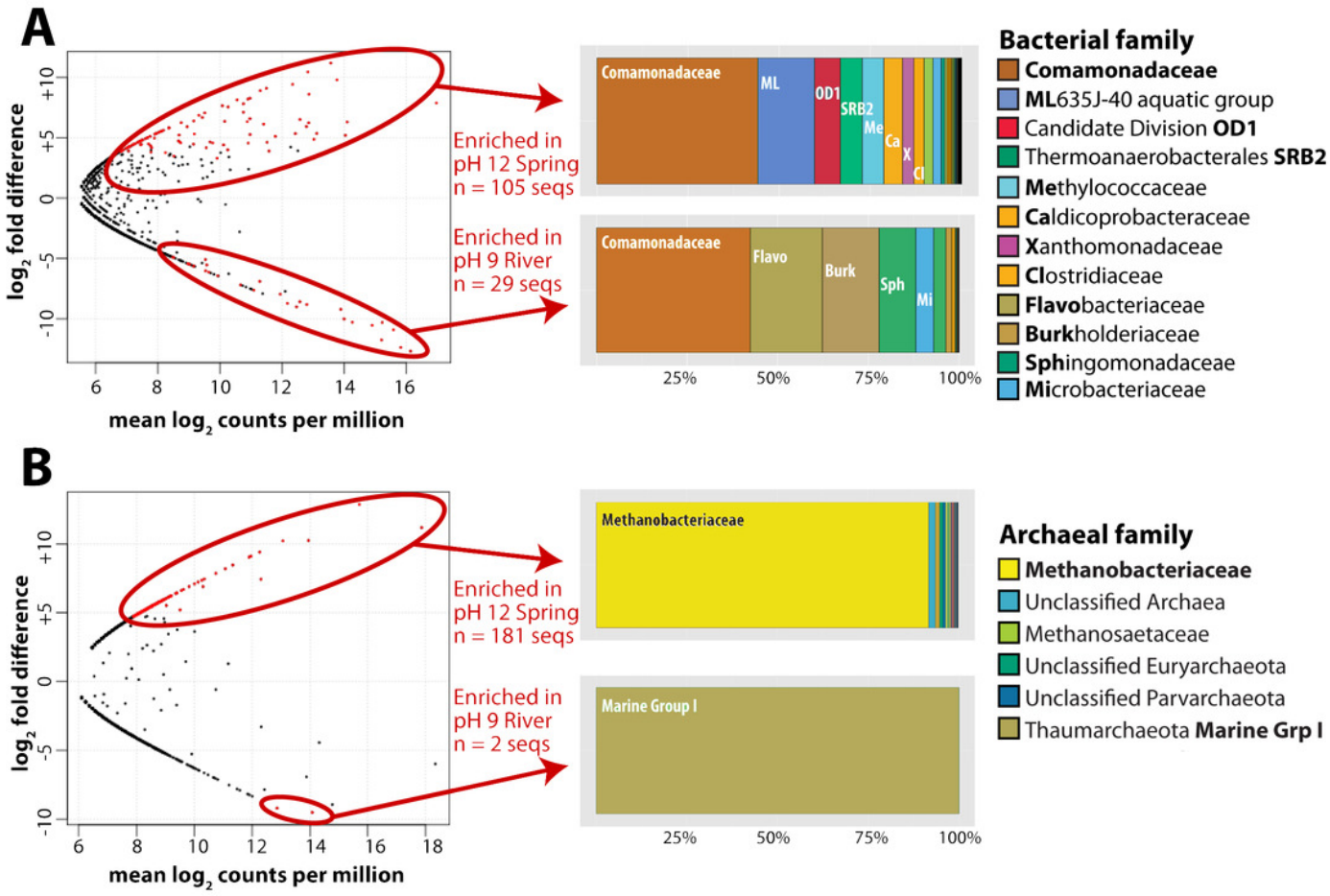

A
log2 fold difference
+10
+5
0
-5
-10
6
8
10
12
14
16
mean log2 counts per million
Enriched in pH 12 Spring n = 105 seqs
Enriched in pH 9 River n = 29 seqs
Comamonadaceae
ML
OD1
SRB2
Me
Ca
X
Cl
Comamonadaceae
Flavo
Burk
Sph
Mi
25%
50%
75%
100%
Bacterial family
Comamonadaceae
ML635J-40 aquatic group
Candidate Division OD1
Thermoanaerobacterales SRB2
Methylococcaceae
Caldicoprobacteraceae
Xanthomonadaceae
Clostridiaceae
Flavobacteriaceae
Burkholderiaceae
Sphingomonadaceae
Microbacteriaceae
B
log2 fold difference
+10
+5
0
-5
-10
6
8
10
12
14
16
18
mean log2 counts per million
Enriched in pH 12 Spring n = 181 seqs
Enriched in pH 9 River n = 2 seqs
Methanobacteriaceae
Marine Group I
25%
50%
75%
100%
Archaeal family
Methanobacteriaceae
Unclassified Archaea
Methanosaetaceae
Unclassified Euryarchaeota
Unclassified Parvarchaeota
Thaumarchaeota Marine Grp I

Figure 3

Predicted methanogenesis proteins in the combined-spring metagenomic assembly.

Diagram of methanogenesis pathways from carbon dioxide, formate, acetate, methanol, and methylamines with associated protein homologs identified with KEGG IDs. Green-highlighted proteins are predicted to occur in the combined-spring metagenomic assembly, and yellow-highlighted proteins are predicted to occur in both the combined-spring and combined-river assemblies. The black border indicates proteins identified in the Methanobacteriaceae bin (ESOM Bin 1 in **Figure 4** and **Table 5**). The diagram is modified from (Ferry, 2010).

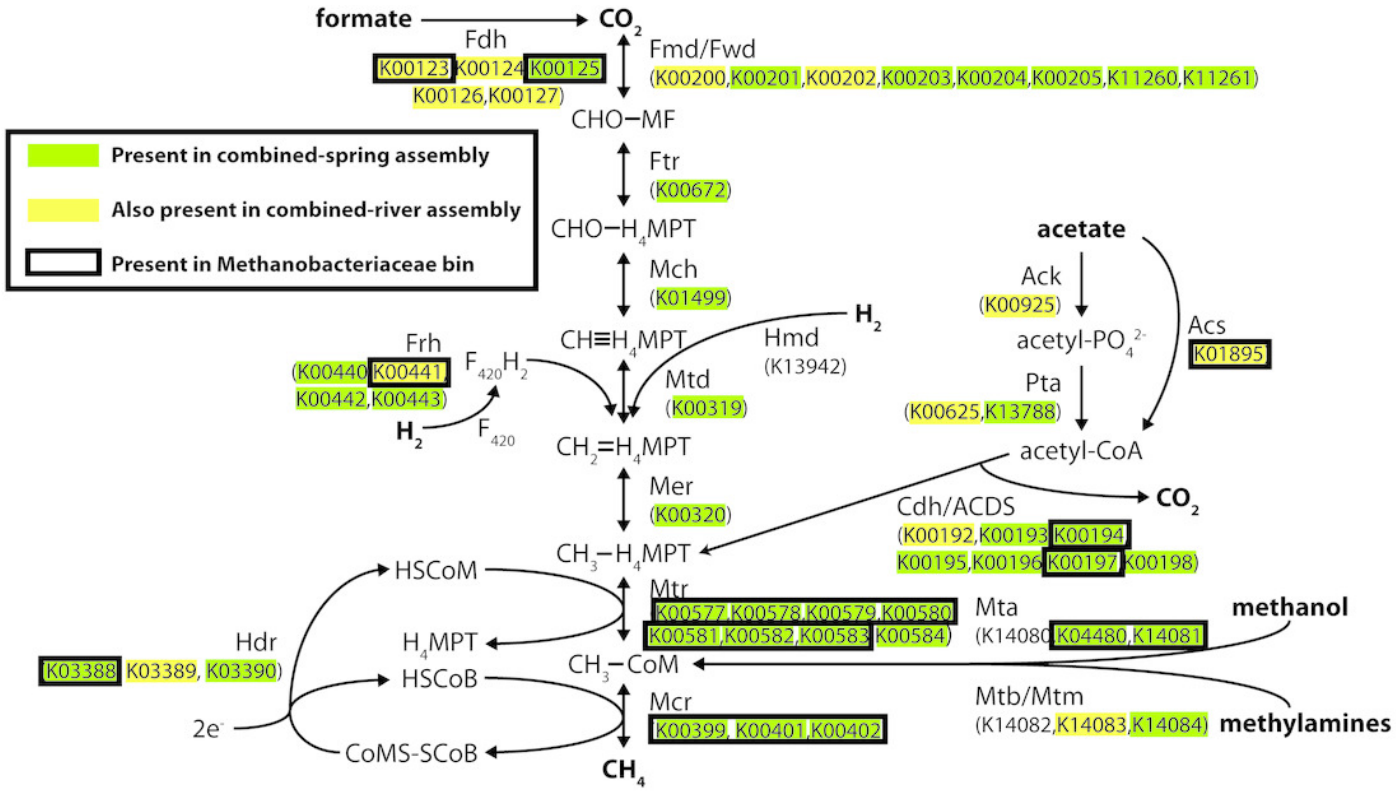
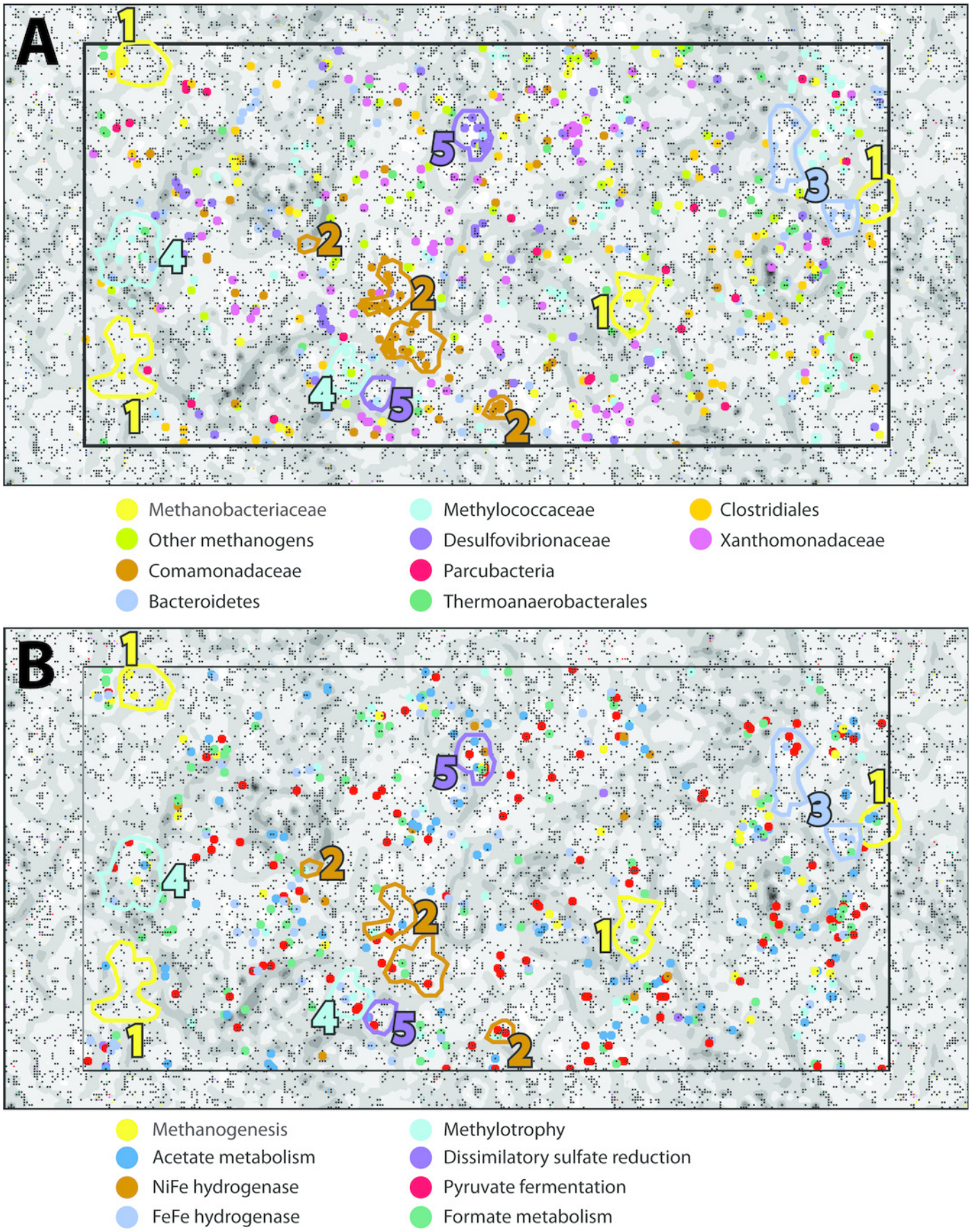


Figure 4

An emergent self-organizing map (ESOM) constructed with the combined-spring metagenomic assembly.

Each point represents a fragment of metagenomic assembly, and points are arranged in space according to the similarities of their tetranucleotide compositions. Dark gray areas are valleys where highly similar sequences cluster together. Light areas of the map are high topographic ridges that separate dissimilar sequences. The same ESOM is shown here twice: metagenomic fragments are color-coded in A) to indicate PhyloSift classifications of taxonomic markers and in B) to show which fragments are predicted to encode protein functions of particular interest to this study. Based on this information, five bins of sequences were selected from the map for further evaluation (**Table 5**).


A
1
5
3
1
4
2
2
1
1
4
5
2
Methanobacteriaceae
Other methanogens
Comamonadaceae
Bacteroidetes
Methylococcaceae
Desulfovibrionaceae
Parcubacteria
Thermoanaerobacterales
Clostridiales
Xanthomonadaceae
B
1
5
3
1
4
2
2
1
1
4
5
2
Methanogenesis
Acetate metabolism
NiFe hydrogenase
FeFe hydrogenase
Methylotrophy
Dissimilatory sulfate reduction
Pyruvate fermentation
Formate metabolism