# 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** [1], **August A Longino** [1], **Susan Q Lang** [2, 3], **Marvin D Lilley** [4], **Gretchen L Früh-Green** [3], **Matthew O Schrenk** [5]

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

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

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

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

[5] Department of Earth and Environmental Sciences, Michigan State University, East Lansing, Michigan, 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.

1    **Metagenomic identification of active methanogens and methanotrophs in serpentinite**

2    **springs of the Voltri Massif, Italy**

3    William J. Brazelton[1], Christopher N. Thornton[1], Alex Hyer[1], Katrina I. Twing[2], August A.

4    Longino[1], Susan Q. Lang[3,5], Marvin D. Lilley[4], Gretchen L. Früh-Green[5], Matthew O. Schrenk[2]

5    [1]Department of Biology, University of Utah, Salt Lake City, UT, USA

6    [2]Department of Earth and Environmental Sciences, Michigan State University, East Lansing, MI,

7    USA

8    [3]Department of Earth and Ocean Sciences, University of South Carolina, Columbia, SC, USA

9    [4]School of Oceanography, University of Washington, Seattle, WA, USA

10    [5]Department of Earth Sciences, ETH Zurich, Zurich, Switzerland

11    Corresponding author:

12    William J. Brazelton

13    Department of Biology

14    University of Utah

15    257 South 1400 East, Rm. 201

16    Salt Lake City, UT 84112-0840

17    (801) 587-9455

18    william.brazelton@utah.edu

## 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.

34   **Introduction**

35   Hydrothermal systems hosted in ultramafic rocks rich in serpentine minerals can support active

36   and unique microbial communities (see Schrenk, Brazelton & Lang, 201 for review). Microbial

37   ecosystems associated with serpentinization (the process by which serpentine is formed) have

38   been investigated by sampling seafloor hydrothermal chimneys (e.g. Kelley et al., 2005;

39   Quéméneur et al., 2014), alkaline springs on land (Brazelton et al., 2013; Suzuki et al., 2013;

40   Woycheese et al., 2015), and wells drilled into serpentinizing rocks (Tiago & Veríssimo, 2013;

41   Cardace et al., 2013; Crespo-Medina et al., 2014). The microbial inhabitants of these

42   environments are thought to feed on the hydrogen gas and organic compounds released as

43   byproducts of serpentinization, but representative examples of specific organisms consuming

44   specific products of serpentinization are lacking. For example, methane is typically abundant in

45   serpentinite-hosted hydrothermal systems, but only a few studies have provided evidence for

46   methanogenic or methanotrophic biological activity in such systems (Brazelton et al., 2011; Kohl

47   et al., 2016).


48   The role of methane-cycling organisms in serpentinizing systems has been enigmatic since the

49   discovery of thick biofilms of methanogens belonging to order Methanosarcinales in carbonate-

50   brucite chimneys of the Lost City hydrothermal field (Schrenk et al., 2004). Whether these

51   organisms are primarily involved in methane production or methane oxidation remains uncertain

52   (Proskurowski et al., 2008; Bradley, Hayes & Summons, 2009; Brazelton et al., 2011; Lang et al.,

53   2012; Méhay et al., 2013). Similar methanogens belonging to order Methanosarcinales have been

54   identified in Prony Bay, New Caledonia, where freshwater serpentinizing fluids are venting into

55   shallow marine waters and forming carbonate chimneys on the seafloor (Quéméneur et al., 2014).

56   16S rRNA genes associated with methanotrophic bacteria have been detected at Lost City

57   (Brazelton et al., 2006) and Prony Bay (Quéméneur et al., 2014).

58    Recent studies of non-marine alkaline springs associated with serpentinization have identified

59    16S rRNA and/or methyl coenzyme-M reductase (mcrA) genes belonging to methanogens in

60    California (The Cedars: Suzuki et al., 2013; Adobe Springs: (Blank et al., 2009), Costa Rica

61    (Sánchez-Murillo et al., 2014), and the Philippines (Woycheese et al., 2015). At the Tablelands

62    Ophiolite (Newfoundland, CA), however, experimental incubations have been unable to detect

63    methanogenesis at high pH (Morrill et al., 2014), and a metagenomic study at this site did not

64    detect any methanogenesis genes (Brazelton, Nelson & Schrenk, 2012; Brazelton et al., 2013).

65    16S rRNA and mcrA sequences belonging to the ANME-1a group of anaerobic methanotrophic

66    archaea, but not typical methanogens, were detected in serpentinizing groundwater in the Cabeço

67    de Vide Aquifer in Portugal (Tiago & Veríssimo, 2013). Biogeochemical studies at The Cedars

68    have inferred biological methanogenesis from isotopic signatures of methane and experimental

69    results, but these studies have not included any biological characterizations of the responsible

70    organisms (Morrill et al., 2013; Wang et al., 2015; Kohl et al., 2016). Thus, methanogens have

71    been detected in some continental serpentinite springs, but previous reports have not provided

72    any quantitative measurements of their environmental distributions nor any genomic or

73    metagenomic sequences from methanogens. Evidence for methanotrophic bacteria in these

74    systems is even more scarce: only a single study of non-marine alkaline springs associated with

75    serpentinization has previously reported methanotrophic bacteria (Sánchez-Murillo et al., 2014).

76    In the Voltri Massif (Italy), ultrabasic springs (pH 11-12) rich in calcium and methane are issuing

77    from serpentinites and rare lherzolites and are associated with the precipitation of large amounts

78    of carbonate (Bruni et al., 2002; Cipolli et al., 2004; Boulart et al., 2013; Schwarzenbach et al.,

79    2013). Inorganic carbon geochemistry of the springs and associated travertine deposits suggest

80    autotrophic microbial activity in the subsurface (Schwarzenbach et al., 2013). A previous study

81    has characterized the archaeal and bacterial communities within the travertine deposits formed

82    when the springs reach the surface (Quéméneur et al., 2015), but there are no previous

83    microbiological studies of the spring water. In this report we show that pH 12 springs at the Voltri

84    Massif are transporting distinct archaeal and bacterial communities, including methanogenic

85    archaea and methanotrophic bacteria, from subsurface habitats where they are likely to be

86    supported by hydrogen gas, methane, and possibly other products of serpentinization-associated

87    reactions. We provide insights into the biology and potential subsurface habitats of these

88    organisms with metagenomic and experimental studies.


89    **Materials and Methods**

90    *Location and sample collection*

91    The ultrabasic springs that were investigated in this study (BR2: 44.4512ºN, 8.7820ºE; GOR34:

92    44.5970ºN, 8.7833ºE) are located in the Voltri Massif near Genoa, Italy (Figure 1). Geochemical

93    measurements of these springs were conducted in concert with the microbiological studies

94    reported here and were reported by Schwarzenbach et al. (2013). At each of the BR2 and GOR34

95    locations, ultrabasic spring water and nearby river water were filtered through Millipore Sterivex

96    0.2 µm filter cartridges using a portable peristaltic pump. Measurements of pH and Eh were

97    obtained with an Oakton PCS Testr 35 and an Oakton Testr 10, respectively, from pumped water

98    that had passed through the peristaltic tubing but before the Sterivex filter (Table 1). Eh readings

99    were corrected for the standard Ag/AgCl electrode (+200 mV).


100    GOR34-1-spring and GOR34-3-spring (Figure 1A) were sampled from two shallow pools within

101    a large deposit of travertine (Schwarzenbach et al., 2013). Each pool appeared to be fed by a

102    point source of spring water at the bottom, into which peristaltic tubing was inserted as far as

103    possible into the host rock (a few centimeters). The GOR34-1 pool was much smaller than that of

104    GOR34-3, and it was emptied before sampling by siphoning water out of the pool. During

105    pumping and sampling of water from the apparent subsurface source, the pump rate was adjusted

106    to approximately match the rate of inflow, determined by a constant water level in the pool. The

107    volume of water in the GOR34-3 pool did not visibly change during sampling, indicating that the

108    maximum pump rate was significantly slower than the rate of spring water inflow into the pool.

109    The pumped spring water from both GOR34-1 and GOR34-3 became more basic (higher pH) and

110    more reducing (lower Eh) during pumping of the first few liters, after which the readings

111    stabilized, and filtering commenced. GOR34-river was sampled from the surface of the adjacent

112    river at a site of exposed, rapid flow ~50 m upstream of the spring.


113    BR2-spring was sampled from a metal pipe (Figure 1B) through which spring water was flowing

114    at 492 mL s$^{-1}$ into the adjacent river (Schwarzenbach et al., 2013). After filtering >500 L of spring

115    water from BR2-spring in 2012, less than 500 ng of DNA was recovered in initial extractions,

116    which was reduced to 50 ng after pooling and final purification. Even lower yields were achieved

117    in 2013 (Table 2). BR2-river was sampled from the surface of the river at a site of exposed, rapid

118    flow ~10 m upstream of the spring.


119    BR2-spring, BR2-river, and GOR34-river were sampled in June 2012 and August 2013.  GOR34-

120    1 was sampled in October 2011 and August 2013, while GOR34-3 was only sampled in June

121    2012. Field replicates were collected for all spring and river samples, but replicates were pooled

122    post-DNA-extraction when necessary to accumulate enough DNA for sequencing. Furthermore,

123    even after pooling of replicates, some samples contained only enough DNA for either 16S rRNA

124    amplicon sequencing or shotgun metagenomic sequencing, but not both. For this reason, BR2-

125    spring-2012, BR2-spring-2013, and GOR34-spring3-2012 are represented in this study by

126    shotgun metagenomic sequences but not 16S rRNA amplicon sequencing surveys (Table 2).

127    When available, field replicates are indicated with 'a' or 'b' at the end of the sample name, such as

128    GOR34-river-2013a and GOR34-river-2013b.

129    *Water chemistry*

130    Water samples were collected for measurements of hydrogen, methane, sulfate, and sulfide. Gas

131    samples were collected with 100 mL syringes by sampling 60 mL of water, minimizing any

132    incorporation of air. Subsequently, 40 mL of nitrogen or helium gas was introduced to the syringe

133    as headspace gas and the two phases were allowed to equilibrate. The gas phase was injected into

134    evacuated vacutainers for later analysis. The concentrations of gas were determined by gas

135    chromatography (GC), calibrated with commercially available gas standards of known

136    concentration. For hydrogen, samples were injected in split/splitless mode (5:1 split) onto a CP-

137    Molsieve 5Å column (50 m x 0.32 µm, 30 µm thickness) at 175°C with a helium flow rate of 1.8

138    mL/min and detected with a Pulsed Discharge Detector. Methane was determined similarly but

139    with a GS-Carbonplot column (30 m x 0.32 µm ID, 1.5 µm film thickness) held at 100°C for 3

140    minutes then ramped at 50 °C/min to 200 ºC and held for 0.6 minutes. Detection was with a

141    flame ionization detector. The detection limit was ~0.6 nM for both hydrogen and methane.

142    Dissolved inorganic carbon (DIC) measurements were performed on samples that were collected

143    by injecting 1-8 mL of 0.2 µm filtered water into Exetainers (Labco Limited, UK) that had been

144    previously prepared with ~100 µL of phosphoric acid and flushed with helium. Concentrations

145    were determined by injecting aliquots of the headspace with a calibrated gas-tight syringe into the

146    same GC set-up used for hydrogen. DIC concentrations from 2011 were reported by

147    Schwarzenbach et al. (2013).

148    Sulfide concentrations were determined by the methylene blue method and detected at 610 nm on

149    a spectrophotometer (Cline, 1969). Concentrations were compared to standard curves of freshly

150    prepared sodium sulfide. The detection limit was 3.1 μM. For sulfate, 100 mL of water was

151    collected in high density polypropylene bottles and stored cool until analysis. Concentrations

152    were determined by a DX-120 ion chromatograph equipped with an IonPac As14 column (4 x

153    250 mm). The detection limit was 1.0 μM. Organic acids were analyzed by high performance

154    liquid chromatography (HPLC) by the method of Albert & Martens (Albert & Martens,

155    1997) with minor modifications. After derivatization with 2-nitrophenyl hydrazine and 1-ethyl-3-

156    (3-dimethylaminopropyl) carbodimide hydrochloride, acids were separated on a Prevail Organic

157    Acid C18 column and detected at 400 nm. Adipic acid was used as an internal standard.

158    *Enumeration of microbial cells*

159    Fluids (50 mL per field replicate; three replicates per sample in Table 2) were preserved in the

160    field for cell abundance enumeration at a final concentration of 3.7% formaldehyde and stored at

161    4°C. In the laboratory, preserved fluids (5 to 20 ml each) were filtered through a 0.2 µm black

162    polycarbonate filter, and captured cells were stained with DAPI and counted with an

163    epifluorescence microscope. At least 30 fields were counted for each field replicate and used to

164    calculate the average cell concentration per mL of fluid for that replicate sample. The numbers

165    reported in Table 2 reflect the mean cell concentration for three field replicates from each sample.

166    *DNA extraction and sequencing*

167    Filters were placed on ice immediately and stored within a few hours at liquid nitrogen

168    temperature in a vapor shipper (MVE SC4/2V) for frozen transport to the home laboratory. Each

169    Sterivex filter was extracted according to protocols modified from those described in (Huber,

170    Butterfield & Baross, 2002; Sogin et al., 2006). Briefly, extractions were performed by lysis via

171    freeze/thaw cycles and lysozyme/Proteinase K treatment and purified with phenol-chloroform

172    extractions and precipitation in ethanol. Environmental DNA quantification was performed with

173    the Qubit fluorometer (ThermoFisher). Extraction blanks never yielded quantifiable DNA and

174    were not sequenced. Further purification of DNA was conducted with QiaAmp (Qiagen) columns

175    according to the manufacturer's instructions for purification of genomic DNA.


176    Purified DNA was submitted to the Josephine Bay Paul Center, Marine Biological Laboratory

177    (MBL) at Woods Hole for amplicon sequencing of archaeal and bacterial 16S rRNA genes with

178    an Illumina MiSeq platform. The V4-V5 hypervariable region of the bacterial 16S rRNA gene

179    was targeted with one forward (518F; CCAGCAGCYGCGGTAAN) and three reverse primers

180    (926R; CCGTCAATTCNTTTRAGT; CCGTCAATTTCTTTGAGT;

181    CCGTCTATTCCTTTGANT). The V4-V5 hypervariable region of the archaeal 16S rRNA gene

182    was targeted by a combination of five forward primer variants (517F;

183    GCCTAAAGCATCCGTAGC; GCCTAAARCGTYCGTAGC; GTCTAAAGGGTCYGTAGC;

184    GCTTAAAGNGTYCGTAGC; GTCTAAARCGYYCGTAGC) and a single reverse primer

185    (958R; CCGGCGTTGANTCCAATT). Additional methodological details and description of

186    primer development are published in (Topçuoğlu et al. in revision; Nelson et al., 2014) Amplicon

187    sequences were screened for quality, including chimera-checking with UCHIME (Edgar et al.,

188    2011), by the MBL as previously described (Huse et al., 2014a), and high-quality merged

189    sequences were distributed  via the Visualization and Analysis of Microbial Population Structures

190    (VAMPS) website (Huse et al., 2014b).


191    The MBL also conducted shotgun metagenomic sequencing of a subset of these samples (see

192    Table 2). Metagenomic libraries were constructed with the Nugen Ultralow Ovation kit according

193    to the manufacturer's instructions. Paired-end sequencing with a 100 cycle Illumina HiSeq run

194     generated partial ~30 bp overlaps, and six libraries were multiplexed per lane. 16S rRNA

195     amplicon sequences are publicly available at the VAMPS database (http://vamps.mbl.edu) under

196     the project code DCO_BRZ and sample code Serp_LIG. Shotgun metagenomic data is publicly

197     available in MG-RAST under IDs 4545477.3, 4545478.3, 4545479.3, 4545480.3, 4537863.3,

198     4537864.3, 4537868.3, and 4537869.3. All sequence data related to this study are also available

199     via the SRA identifier SRP049438 and BioProject PRJNA265986.

200     *Quantitative PCR of archaeal and bacterial 16S rRNA gene copies*

201     The 16S rRNA gene copies of Bacteria and Archaea were quantified via quantitative PCR on a

202     BioRad CFX Connect Optics Module with the BioRad SsoAdvanced SybrGreen assay and

203     domain-specific primers targeting the V6 region of the 16S rRNA gene. Primers 958F and 1048R

204     were used for Archaea, and primers 967F and 1064R were used for Bacteria (previously

205     published in Sogin et al., 2006 and described on the VAMPS website

206     https://vamps.mbl.edu/resources/primers.php). Gene copy numbers were calculated by plotting

207     quantification values from environmental samples onto standard curves generated by

208     amplification of DNA from *Methanocaldococcus jannaschii* (for Archaea) and *Escherichia coli*

209     (for Bacteria) with the domain-specific primers. Bacteria:Archaea ratios were then calculated

210     with gene copy numbers normalized to sample size, i.e. the volume of fluid filtered for that

211     sample.

212     *Analysis of 16S rRNA amplicon data*

213     Additional quality screening of 16S rRNA amplicon sequences from VAMPS was conducted with

214     the mothur (v.1.34.2) software platform (Schloss et al., 2009) to remove sequences with >9

215     homopolymers and >0 ambiguous bases. This screening step removed only 124 bacterial

216     sequences and 5 archaeal sequences. The sequences were then pre-clustered with the mothur

217    command pre.cluster (diffs=1), which reduced the number of unique sequences from 410,517 to

218    322,879 for bacteria and from 138,629 to 95,867 for archaea. This pre-clustering step removes

219    rare sequences most likely created by sequencing errors (Schloss, Gevers & Westcott, 2011). The

220    final unique sequence types were considered to be the operational taxonomic units (OTUs) for

221    this study. We chose not to cluster sequences any more broadly (e.g. 97% sequence similarity)

222    because additional clustering inevitably results in a loss of biological information without reliably

223    improving sequencing error and because no arbitrary sequence similarity threshold can be

224    demonstrated to correspond to consistent species-like units across all taxa. Taxonomic

225    classification of all OTUs was performed with mothur using the SILVA reference alignment

226    (SSURefv119) and taxonomy outline (Pruesse, Peplies & Glöckner, 2012). Differences in the

227    relative abundances of OTUs between groups of samples were measured with the aid of the R

228    package edgeR v.3.6.8 (Robinson, McCarthy & Smyth, 2010) as recommended by (McMurdie &

229    Holmes, 2014). Results were visualized with the aid of the R package phyloseq v.1.9.13

230    (McMurdie & Holmes, 2013).


231    *Analysis of metagenomic data*

232    Detailed documentation of all computational analyses in this study, including actual software

233    commands, are provided on the Brazelton lab's website (https://baas-

234    becking.biology.utah.edu/data/category/18-protocols), and all custom software and scripts are

235    available at https://github.com/Brazelton-Lab. Quality-filtering of shotgun metagenomic

236    sequences included identification and removal of artifactual sequences with cutadapt v.1.9

237    (Martin, 2011) as follows: reads found to have Illumina adapters starting at their 5'-end were

238    discarded, and reads containing Illumina adapters towards the 3'-end were trimmed where the

239    first adapter began. Identical and 5'-prefix replicates were also removed, as suggested by

240    (Gomez-Alvarez, Teal & Schmidt, 2009). Nucleotides (0-2) at the beginning and end of reads

241     were also cropped from all reads in that sample if those positions exhibited nucleotide

242     frequencies inconsistent with the nucleotide frequency distribution for the rest of the read. Low-

243     quality bases were removed from the ends of the reads, and the remaining sequence was scanned

244     20 base pairs at a time and trimmed where the mean quality score fell below a score of 8. Reads

245     that did not pass a minimum length threshold of 62 bp after quality and adapter trimming were

246     removed from the dataset. Phylogenetic affiliations of the unassembled, quality-checked paired

247     reads were determined with PhyloSift v.1.0.1 (Darling et al., 2014).


248     Each metagenomic dataset was processed individually as described above. The four

249     metagenomes from ultrabasic springs and the four metagenomes from adjacent rivers were then

250     combined for two pooled assemblies with Ray Meta v.2.3.1 (Boisvert et al., 2012). A kmer of 41

251     was chosen after manual inspection of assemblies with kmer values of 31, 41, 51, and 61. High-

252     quality reads from each sample were mapped onto the pooled assembly with Bowtie2 v.2.2.6

253     (Langmead & Salzberg, 2012) to obtain sample-specific coverages. The Prokka pipeline

254     (Seemann, 2014) was used for gene prediction and functional annotation. The arguments

255     --metagenome and --proteins were used with Prokka v.1.12 to indicate that genes should be

256     predicted with the implementation of Prodigal v.2.6.2 (Hyatt et al., 2010) optimized for

257     metagenomes and then searched preferentially against a custom protein database. The database

258     provided was the last free version (2011) of the Kyoto Encyclopedia of Genes and Genomes

259     (Ogata et al., 1999), obtained from MG-RAST (Meyer et al., 2008). Abundances of metabolic

260     pathways were obtained by mapping KEGG protein IDs and their normalized counts calculated

261     with HTSeq v.0.6.1 (Anders, Pyl & Huber, 2015) onto the FOAM ontology (Prestat et al.,

262     2014) with MinPath (Ye & Doak, 2009) as implemented in HUMAnN2 v.0.6.0 (Abubucker et al.,

263     2012). Similar results were also achieved by annotating both pooled assemblies with the

264     UniProtKB protein database (Consortium, 2015) and MetaCyc pathways (Caspi et al., 2014).

265    Contigs from the pooled spring assembly were binned according to their tetranucleotide

266    frequencies with the aid of an emergent self-organized map (ESOM) as implemented by

267    Databionic ESOM tools v.1.1 (Ultsch & Mörchen, 2005) and as described by (Dick et al., 2009).

268    The resulting ESOM was annotated with phylogenetic markers classified by PhyloSift and

269    selected predicted protein functions of interest identified by KEGG ID. Completeness and

270    contamination of ESOM bins were evaluated with CheckM v.1.0.4 (Parks et al., 2015).

271    Contamination levels as measured by CheckM do not always agree with the percentage of

272    minority taxa reported by PhyloSift because CheckM only identifies contamination when single-

273    copy markers are present as multiple, divergent sequences and because PhyloSift results are

274    abundance-weighted.


275    *Metabolic Activity Assays*

276    During each sampling expedition, water samples collected from each of the BR2 and GOR34

277    springs were incubated with one of several $^{13}$C-labeled carbon sources (obtained from Cambridge

278    Isotope Laboratories, Tewksbury, MA): bicarbonate (NaH$^{13}$CO3), formate (Na$^{13}$CHOO), acetate-

279    $^{13}$C1 (NaCH$_3$$^{13}$COO), acetate-$^{13}$C2 (Na$^{13}$CH$_3$COO), propionate (NaCH$_3$CH$_2$$^{13}$COO), or methane

280    ($^{13}$CH$_4$). Each 99% $^{13}$C-labeled carbon compound was diluted 1:100 with the corresponding non-

281    $^{13}$C-labeled compound to produce a final $\delta^{13}$C of +691 ‰, and this compound was present in each

282    incubation at a final concentration of 0.1 M after addition of 3 mL of water sample. No buffering

283    nor additional amendments were applied to the incubations, and the experiments are assumed to

284    have proceeded at the in situ pH of the spring water. 'Dead' control incubations were conducted

285    by filter-sterilizing the spring water through a 0.2 µm syringe filter in the field and prior to

286    incubation. Three live replicates and three dead replicates from each spring were incubated for

287    each treatment. Methanogenesis experiments received an addition of 5 cc H$_2$ gas to each

288    replicate. Methanotrophy experiments received an addition of 5 cc $^{13}CH_4$ that had been diluted

289    1:10 with non-$^{13}$C-labeled methane. All incubations were conducted in gas-tight glass Exetainers

290    (Labco Limited, UK) flushed with $N_2$ and 2-5% $H_2$ in a Coy anoxic chamber and flushed again

291    outside the chamber and slightly over-pressured for transport with $N_2$ gas. Oxic methanotrophy

292    experiments were conducted after exposing the contents to room air by opening the lids for a few

293    minutes prior to adding $^{13}CH_4$. Some methanogenesis experiments were treated with a reducing

294    agent in the form of dithiothreitol (DTT) or sodium sulfide, but no differences were observed

295    with or without the additional reducing agent. After incubation at ambient temperature (20-35 ºC)

296    in the dark for 4-8 days, Exetainers were injected with 0.2 mL of 50% phosphoric acid to convert

297    all dissolved inorganic carbon into $CO_2$, and the $^{13}C/^{12}C$ ratios of $CO_2$ and $CH_4$ were measured by

298    gas chromatography isotope ratio mass spectrometry using a Thermo Fisher Delta V Plus Isotope

299    Ratio Mass Spectrometer, interfaced with a Trace gas chromatograph, a GC IsoLink interface and

300    a ConFlo IV at the Stable Isotope Laboratory of the ETH in Zurich.


301    **Results**

302    *Biogeochemical contrasts between ultrabasic springs and surface rivers*

303    At both the BR2 and GOR34 sites, the spring water had much higher pH, lower Eh (i.e. more

304    reducing), fewer particulates, and lower biomass than the adjacent river (Tables 1-2).  The high

305    pH (up to 12.3) and low Eh are consistent with the spring water being sourced from ultramafic

306    rocks where serpentinization-like reactions are occurring now or have occurred in the past (Alt et

307    al., 2013). The difference in particulates was noticeable due to the much greater volumes of

308    spring water, compared to surface river water, that could be filtered through a single Sterivex

309    cartridge before clogging (more than 20-50 L for springs, only 2-5 L for surface rivers). The

310    difference in biomass was also evident from the DNA yield: the GOR34 springs yielded 3-5-fold

311    less DNA per liter of water compared to the adjacent river, and the BR2 spring yielded >1,000-

312    fold less DNA per liter of water compared to the adjacent river. The cell densities (as measured

313    by enumeration of DAPI-stained cells under a microscope) of the BR2 spring ($10^2 - 10^3$ cells per

314    mL of water) were 2-40-fold lower than the cell densities of the adjacent river (Table 2). The cell

315    densities of the GOR34 springs were 2-20-fold lower than the adjacent river.


316    BR2 spring waters contained higher methane and lower hydrogen concentrations compared to the

317    GOR34 springs (Table 1). At both sites, spring water contained much lower dissolved inorganic

318    carbon (DIC) than in the adjacent rivers, likely due to the precipitation of calcium carbonate at

319    high pH (Alt et al., 2013). Sulfate was also very low in all springs, though detectable at BR2.

320    Sulfide was elevated in all springs compared to the adjacent rivers.


321    *Taxonomic survey of bacterial 16S rRNA genes*

322    The bacterial taxonomic composition of the pH 12 spring GOR34-spring1 was distinct from that

323    of surface rivers, as measured by high-coverage Illumina MiSeq sequencing of bacterial 16S

324    rRNA gene amplicons. Comamonadaceae (a member of order Burkholderiales, class

325    Betaproteobacteria) was the most abundant family in GOR34-spring1 in 2013 (35-40% of all

326    bacterial 16S rRNA amplicon sequences, Supplementary File S1). The most abundant OTU

327    within this family is 100% identical to the most abundant OTU recovered from ultrabasic

328    serpentinite springs at the Tablelands Ophiolite in Newfoundland, Canada (Brazelton et al., 2013)

329    and to the most abundant OTU in borehole fluids from the Coast Range Ophiolite Microbial

330    Observatory in California (Crespo-Medina et al., 2014). This same bacterial taxon has been

331    isolated from serpentinite springs at The Cedars (northern California) by Suzuki et al. (Suzuki et

332    al., 2014), who have proposed the novel genus *Serpentinomonas* for these organisms. Curiously,

333    Comamonadaceae sequences only comprised 14% of all bacterial sequences in GOR34-spring1

334    in 2011. Instead, the most abundant sequences were those classified as Candidate Division OD1

335    (21% of sequences in GOR34-spring1 in 2011 but only 4% in 2013).

336    Comamonadaceae sequences were also well represented in surface river samples (Supplementary

337    File S1), but the Comamonadaceae OTUs from river samples were distinct from the

338    Comamonadaceae OTUs from ultrabasic spring samples (only 94% sequence identity over 373

339    bases between the most abundant Comamonadaceae OTU in GOR34-spring1-2013b compared to

340    the most abundant Comamonadaceae OTU in GOR34-river-2013). Neither of these sequences

341    could be classified at a lower taxonomic rank than family. This result highlighted the need to

342    conduct comparisons of bacterial community compositions at the level of individual OTUs

343    (Supplementary File S2), rather than taxonomic classifications.

344    *Differential abundance of bacteria*

345    Although the bacterial community compositions of ultrabasic springs and the adjacent rivers were

346    generally distinct, the most abundant OTUs in rivers were also present, at low relative

347    abundances, in ultrabasic springs, and the converse was also true. In other words, the most

348    abundant 16S rRNA OTUs at each site were not completely exclusive to that site. For example,

349    the most abundant OTU in GOR34-river-2013b (classified as Sphingomonadaceae) occurred

350    8515 times in that sample and 0 and 4 times in GOR34-spring1-2013a and GOR34-spring1-

351    2013b, respectively. Conversely, the most abundant OTU in GOR34-spring1-2013b (the

352    Comamonadaceae *Serpentinomonas* OTU described above) occurred 32,866 times in that sample

353    and 71 and 0 times in GOR34-river-2013a and GOR34-river-2013b, respectively (Supplementary

354    File S2). As with any environmental study, these sequences could have appeared in multiple

355    locations because of natural mixing in the environment or due to accidental contamination during

356    sampling or sequencing. Regardless of the cause, the source of shared sequences can be inferred

357    by making the parsimonious assumption that the site of higher abundance is closer to that

358    organism's true habitat. Therefore, we looked for potential subsurface-specific organisms by

359    testing for significant differences in the abundances of individual OTUs between ultrabasic

360    spring water and the adjacent river.


361    For this analysis, we contrasted the counts of OTUs in GOR34-spring1 to those in GOR34-river.

362    Each site was represented by two field replicates collected in 2013 (Table 2). In Figure 2, each

363    data point of the plot represents the difference in abundance of a single OTU between GOR34-

364    spring1 and GOR34-river. The plot's x-axis displays the mean abundance (in units of $\log_2$ counts

365    per million) of each OTU across all four samples (two GOR34-spring1 replicates and two

366    GOR34-river replicates). Red data points indicate OTUs whose differential abundance between

367    GOR34-spring1 and GOR34-river passed a significance test (false discovery rate < 0.05)

368    implemented by the edgeR package (Robinson, McCarthy & Smyth, 2010). Spring-enriched

369    OTUs include the Comamonadaceae described above as well as a group of uncultured

370    Bacteroidetes known as aquatic group ML635J-40, Candidate Division OD1, an uncultured group

371    within the Thermoanaerobacterales, and Methylococcaceae (Figure 2A). Sequences that were not

372    consistently enriched in one sample type compared to the other are represented by black points in

373    Figure 2. These sequences could represent potential contaminants or truly cosmopolitan species,

374    or their abundances may be correlated to factors not captured by the study design. The data

375    visualization approach in Figure 2 highlights the key differences between the spring and river

376    microbial communities by acknowledging the reality that all environmental samples are mixtures

377    formed by multiple sources of organisms and without making the naive and extreme assumption

378    that all shared sequences must be some kind of contamination.


379    *Taxonomic survey of archaeal 16S rRNA genes*

380    As with the bacterial communities, the archaeal communities in the ultrabasic spring waters were

381    distinct from those in the surface rivers, as measured by high-coverage Illumina MiSeq

382    sequencing of archaeal 16S rRNA gene amplicons. More than 88% of total archaeal sequences in

383    GOR34-spring1-2013, for example, belong to the Methanobacteriaceae family of methanogens.

384    The closest match (~97% sequence similarity) in the GenBank, VAMPS, and SILVA databases to

385    these Methanobacteriaceae sequences is a clone from subsurface water collected in a deep South

386    Africa mine (GenBank DQ230925). The taxonomic composition of GOR34-river was nearly

387    completely different; ~72% of sequences in both 2012 and 2013 were assigned to Marine Group

388    I, within the Thaumarchaeota (Supplementary File S3). As with the bacterial sequences, the most

389    abundant archaeal OTUs in GOR34-river were also present, at low relative abundances, in

390    GOR34-spring1, and the converse was also true. For example, the most abundant OTU in

391    GOR34-spring1-2013b occurs 89,407 times in that sample, 127 times in GOR34-river-2013, and

392    24 times in GOR34-river-2012 (Supplementary File S4).


393    We identified potential subsurface-specific archaeal 16S rRNA OTUs by contrasting the archaeal

394    sequences from GOR34-spring1 to those of GOR34-river. Unfortunately, GOR34-spring1-2013b

395    was the only spring sample to yield sufficient DNA and sufficient read coverage for this analysis

396    of archaeal 16S rRNA sequences. GOR34-river is represented in this analysis by two samples:

397    GOR34-river-2012a and GOR34-river-2013a. Not surprisingly, considering the differential

398    abundances reported above, the Methanobacteriaceae sequences were identified as the most

399    abundant OTUs that were significantly enriched in GOR34-spring1 (Figure 2B). Other families

400    of methanogens in addition to Methanobacteriaceae were identified in the 16S rRNA sequences,

401    but none were significantly more abundant in the ultrabasic springs compared to the adjacent

402    river. One sequence belonging to family Methanosaetaceae (order Methanosarcinales) was

403    enriched in the GOR34 spring compared to the adjacent river in 2013, but Methanosaetaceae

404    sequences were equally abundant in the river in 2012 (Supplementary Files S3-S4). The only

405    OTUs that could be considered significantly enriched in GOR34-river belong to the

406    Thaumarchaeota Marine Group I family. The variability of less abundant OTUs between the two

407    samples from GOR34-river prevented them from passing our significance test, even though many

408    of them were extremely rare or absent in GOR34-spring1.


409    *Taxonomic survey of shotgun metagenomic sequences*

410    Phylogenetic classifications (from PhyloSift) of unassembled metagenomic sequences were

411    largely consistent with the 16S rRNA amplicon results described above.  For example, even

412    though Comamonadaceae was the most common family identified in metagenomic sequences

413    from the ultrabasic springs as well as the river at GOR34 (Supplementary File S5), the sample

414    types differed at the genus level: *Hydrogenophaga* was the most common member of

415    Comamonadaceae in the GOR34-spring1 and GOR34-spring3 metagenomes, whereas

416    *Acidovorax* was the predominant representative of Comamonadaceae in GOR34-river

417    (Supplementary File S6). Comamonadaceae sequences were also abundant in metagenomic data

418    from the ultrabasic spring at BR2, although this site appears to be dominated by

419    Desulfovibrionales, a member of Deltaproteobacteria (Supplementary File S5).


420    Metagenomic sequences predicted to represent methanogens (family Methanobacteriaceae)

421    comprised ~1% of shotgun sequences in BR2-spring-2012 and GOR34-spring3-2012

422    (Supplementary File S5). Curiously, no methanogen sequences were found in GOR34-spring1-

423    2013b (Supplementary File S6), despite their dominance in the archaeal 16S rRNA amplicon data

424    from this same sample (Figure 2). None of the 28 million sequence pairs in this metagenome

425    were classified as domain Archaea, so it is possible that the ratio of bacterial to archaeal DNA in

426    this sample was very high. A quantitative PCR assay of bacterial and archaeal 16S rRNA gene

427    copies indicated a 10:1 ratio of Bacteria:Archaea (Table 2) in GOR34-spring1-2013b, which

428    would predict a minority, but perhaps not a complete absence, of archaeal sequences in its

429    metagenome. The Bacteria:Archaea ratio for BR2-spring-2012 was less than 2:1, and only 3% of

430    shotgun metagenomic sequences from this sample were classified as Archaea, indicating a lack of

431    direct correspondence between quantitative PCR results and numbers of shotgun sequences.

432    When archaea were detected in metagenomic sequences from the rivers adjacent to the BR2 and

433    GOR34 springs, they were dominated by Thaumarchaeota (Supplementary File S5), consistent

434    with the archaeal 16S rRNA amplicon results (Figure 2). No Methanobacteriaceae sequences

435    were detected in the metagenomic data from either river.


436    *Metagenomic assembly and metabolic pathway prediction*

437    Shotgun metagenomic sequences were assembled into contiguous genomic fragments (contigs),

438    which were used for predictions of genes, protein functions, and metabolic pathways. Because

439    the four metagenomes from ultrabasic springs (BR2-spring-2012, BR2-spring-2013, GOR34-

440    spring1-2013b, and GOR34-spring3-2012) were similar to each other in taxonomic composition

441    (Supplementary Files S5-S6), they were combined in a pooled assembly called 'combined-

442    spring'. The N50 (length of median contig) of this pooled assembly was 2.4 kb (Table 3), and the

443    longest contig was 99 kb. Of the 104,651 predicted proteins in the combined-spring assembly,

444    49,794 could be annotated with the KEGG database. A large proportion (45%) of predicted

445    proteins had no matches in any of the protein databases and could not be assigned a function

446    (Table 3). The four metagenomes from adjacent rivers (BR2-river-2012a, BR2-river-2013,

447    GOR34-river-2012a, and GOR34-river-2013b) were also combined for a pooled assembly called

448    'combined-river', which has similar assembly statistics to the combined-spring assembly (Table

449    3).

450    The abundance of metabolic pathways was determined by mapping predicted protein functions

451    (as annotated with KEGG) and their sequencing read coverages onto the FOAM ontology.

452    Predicted proteins that could be assigned a KEGG ID represented only 23-33% of the coverage

453    for each spring sample, and 27-29% of this coverage could be integrated into a FOAM pathway

454    (Supplementary File S7). Similar results were obtained by annotating predicted proteins with the

455    UniProtKB database and mapping onto MetaCyc pathways (Supplementary File S7). Table 4 lists

456    the FOAM metabolic pathways that were significantly enriched ($p < 0.05$, two-tailed T test)

457    either in the BR2 spring compared to its adjacent river or in the GOR34 springs compared to their

458    adjacent river. Aerobic oxidation of hydrogen and methanogenesis are among the top pathways

459    enriched in the ultrabasic springs, as one would expect from the enrichment in the spring

460    metagenomes of the *Hydrogenophaga* and Methanobacteriaceae sequences described above. The

461    most abundant metabolic pathways listed in Table 4 also include several pathways involved in

462    fermentation and hydrogen production, which is consistent with the presence of Clostridiaceae,

463    Bacteroidetes, and other potential fermenters (Supplementary File S5). The bacterial methane

464    oxidation pathway (employed by Methylococcaceae, for example) was also abundant in the

465    spring metagenomes and nearly absent in the rivers, but the variable abundance of this pathway

466    among the springs (Supplementary File S7) prevented it from passing the significance test

467    defining Table 4.

468    Methanogenesis pathways using carbon dioxide and formate were among the most enriched

469    pathways in the spring metagenomes, but other FOAM-defined methanogenesis pathways

470    predicted to utilize acetate and methylated compounds were also predicted to be present in these

471    springs (Supplementary File S7). To better understand the evidence for different kinds of

472    methanogenesis pathways in these springs, we examined the presence of each key

473    methanogenesis-associated gene. All of the steps in the 'core' methanogenesis pathway were

474    represented by at least one predicted protein (as defined by KEGG) in the combined-spring

475    assembly (green highlighting in Figure 3). The only exception is the lack of sequences encoding

476    the enzyme Hmd ($H_2$-forming methylenetetrahydromethanopterin dehydrogenase), but this gene

477    is not present in many methanogens (Fricke et al., 2006), including the five *Methanobacterium*

478    genomes currently available in the Integrated Microbial Genomes database (img.jgi.doe.gov).


479    All of the genes encoding formate dehydrogenase (Fdh), acetate kinase (Ack),

480    phosphotransacetylase (Pta), AMP-forming acetyl-CoA synthetase (Acs), and carbon monoxide

481    dehydrogenase / acetyl-CoA decarbonylase/synthase (Cdh/ACDS) are present in the combined-

482    spring assembly. The presence of these genes indicates the genetic potential for both formate and

483    acetate to be used as substrates for methanogenesis, although these proteins can be utilized by

484    non-methanogens as well, which can explain why some of these proteins are also found in the

485    combined-river assembly (yellow highlighting in Figure 3). Two of the three genes encoding Mta

486    (methanol-specific methylcobalamin:coenzyme M methyltransferase) are also present, suggesting

487    methanol as another potential substrate for methanogenesis.


488    *Binning of metagenomic assembly*

489    In order to examine the distribution of predicted protein functions and metabolic pathways among

490    putative organisms within the combined-spring assembly, fragments of metagenomic contigs

491    were binned according to their tetranucleotide compositions with an emergent self-organizing

492    map (ESOM). Each point in the ESOM in Figure 4 represents a contig or a ~5 kb fragment of a

493    contig if the original contig was larger than 5 kb. In Figure 4A, contigs containing PhyloSift

494    taxonomic markers that correspond to the enriched taxa in Figure 2 are colored accordingly.

495    Figure 4B is the same map as Figure 4A, but in this case, fragments are color-coded according to

496    the presence of predicted proteins associated with metabolic functions of interest (Supplementary

497    File S8). Five bins of metagenomic fragments with relatively homogenous taxonomic

498    assignments were identified in this ESOM (Table 5). Unfortunately, none of the five bins

499    contained any 16S rRNA genes, so direct links to the 16S rRNA gene amplicon data could not be

500    made, and comparisons with PhyloSift taxonomic classifications of protein-encoding taxonomic

501    markers are reported below instead.


502    ESOM Bin 1 (44% complete, 1% contamination) corresponds to the Methanobacteriaceae 16S

503    rRNA sequences that were enriched in the ultrabasic springs (Figure 2) because 88% of its

504    sequences encoding taxonomic markers were classified as Methanobacteriaceae by PhyloSift

505    (Supplementary File S6). Furthermore, genomes belonging to genus *Methanobacterium* were the

506    most common source of best hits when sequences belonging to ESOM Bin 1 were competitively

507    recruited to a set of all methanogen genomes available from GenBank. ESOM Bin 1 sequences

508    were recruited to many other types of methanogens as well, indicating that this bin does not have

509    a single nearly identical whole-genome match in GenBank. ESOM Bin 1 includes many of the

510    methanogenesis-associated sequences in the combined-spring assembly (black borders in Figure

511    3), including genes involved in the utilization of carbon dioxide, formate, acetate, and methanol

512    as substrates for methanogenesis. Two of the genes encoding nitrogenase (nifDK) were also

513    identified in this bin, suggesting the potential for nitrogen fixation (Supplementary File S8).


514    ESOM Bin 2 (42% complete, 12% contamination) includes sequences that were classified as

515    Comamonadaceae (43% of taxonomic markers) and Xanthomonadaceae (27% of taxonomic

516    markers) (Supplementary File S6). Both of these bacterial families were also among the most

517    abundant taxa identified in the bacterial 16S rRNA data (Figure 2). Repeated attempts to partition

518    this bin to separate the Comamonadaceae and Xanthomonadaceae sequences were unsuccessful,

519    suggesting shared sequence compositions and/or mis-assembly. ESOM Bin 2 encodes RuBisCo

520    (ribulose-1,5-bisphosphate carboxylase/oxygenase), the key enzyme in the Calvin-Benson-

521    Bassham pathway of carbon fixation. The predicted protein sequence of the large subunit has

522    83% identities and 91% similarities with the Comamonadaceae RuBisCo sequence identified in

523    metagenomic sequences from a serpentinite-hosted ultrabasic spring at the Tablelands Ophiolite

524    in Newfoundland, Canada (Brazelton, Nelson & Schrenk, 2012). Nearly all predicted proteins in

525    the combined-spring assembly involved in the degradation of chlorinated aromatic compounds

526    are also found in ESOM Bin 2 (Supplementary File S7), and these sequences have close matches

527    in other Comamonadaceae genomes. This bin also includes sequences (Supplementary File S8)

528    encoding methane/phenol/toluene hydroxylase (KEGG ID K08684), formate dehydrogenase

529    (K00123), and phosphate acetyltransferase (K00625), each of which has high similarity to

530    proteins encoded by Comamonadaceae or other families within order Burkholderiales. This bin

531    does not include any sequences encoding a NiFe-hydrogenase, but there is at least one large

532    contig in the combined-spring assembly that encodes a NiFe-hydrogenase with 84% amino acid

533    identities with the group 1 NiFe-hydrogenase from *Serpentinomonas* strain H1, a

534    Comamonadaceae isolate from a serpentinite spring at The Cedars, California (Suzuki et al.,

535    2014).


536    ESOM Bin 3 (28% complete, 2% contamination) is a collection of sequences with PhyloSift

537    classifications consistent with that of the Bacteroidetes uncultured aquatic group 'ML635J-40'

538    identified as the second-most abundant bacterial taxon in the ultrabasic springs (Figure 2). 92%

539    of all taxonomic markers in ESOM Bin 3 were classified as Bacteroidetes (Supplementary File

540    S6), although these sequences were somewhat evenly distributed among several different families

541    within the Bacteroidetes, which is consistent with the undetermined phylogenetic placement of

542    the ML635J-40 aquatic group (Nolla-Ardèvol, Strous & Tegetmeyer, 2015). This bin includes

543    50% of all combined-spring metagenomic sequences assigned to the FOAM pathway 'Pyruvate

544    fermentation to acetate III' and is also rich in sugar transporters (Supplementary File S7). ESOM

545    Bin 3 encodes multiple [FeFe]-hydrogenases with high similarity to those encoded by other

546    Bacteroidetes including *Lentimicrobium saccharophilum* (GenBank GAP44922.1) and *Alistipes*

547    sp. ZOR0009 (GenBank WP_047449271).


548    ESOM Bin 4 (48% complete, 6% contamination) corresponds to the Methylococcaceae 16S

549    rRNA sequences identified in Figure 2. 60% of taxonomic markers in this bin were classified as

550    family Methylococcaceae, and 82% were classified as order Methylococcales (Supplementary

551    File S6). Bacteroidetes sequences comprised 12% of taxonomic markers in this bin. Predicted

552    protein sequences for soluble methane monooxygenase (mmoX, K08684) were prominent in

553    ESOM Bin 4 (Supplementary File S8), and they shared 94-96% amino acid identities with mmoX

554    protein sequences in *Methylovulum miyakonense* and *Methylomonas methanica*. Curiously, no

555    particulate methane monooxygenase sequences (pmoABC, K10944-46) were found in the

556    combined-spring assembly. A BLASTP search with known pmoA protein sequences against the

557    combined-spring assembly's predicted proteins yielded only the previously mentioned mmoX

558    sequences.


559    ESOM Bin 5 (60% complete, 6% contamination) appears to capture much of the genomic content

560    associated with the Desulfovibrionales 16S rRNA sequences that dominate the BR2-spring-2012

561    metagenome. 83% of taxonomic markers in this bin were classified as Desulfovibrionales

562    (Supplementary File S6). Furthermore, [NiFe]-hydrogenase sequences in ESOM Bin 5 have high

563    similarity to other Desulfobivrionales genomes; for example, one predicted protein sequence has

564    90% identity with the hydrogenase 2 large subunit of *Desulfonatronum lacustre* (Pikuta et al.,

565    2003). The *Desulfonatronum* genus includes several alkaliphilic, hydrogen-oxidizing species,

566    although there are no reports of *Desulfonatronum* growth above pH 10. In addition to [NiFe]-

567     hydrogenase (hyaAB; K06281-82), most of the combined-spring metagenomic sequence

568     coverage of carbon monoxide dehydrogenase (cooS; K00198), acetyl-coa synthase (acsB;

569     K14128), three subunits of acetyl-CoA decarbonylase/synthase complex (Cdh/ACDS; K00194,

570     K00197, K00198), dissimilatory sulfite reductase (dsrAB; K11180-81), and phosphate

571     transacetylase (K13788) were included in this bin (Supplementary File S8).


572     *Metabolic activity assays*

573     The activity of methanogenesis, methanotrophy, and oxidation of organic acids in ultrabasic

574     springs was tested by measuring the conversion of $^{13}$C-labeled carbon compounds during

575     incubations of spring water samples at ambient temperatures (~20-25 ºC) in the dark. A single

576     plus symbol in Table 6 indicates that the $^{13}$C ratio of the product for that reaction was higher in all

577     three live replicates compared to all three dead replicates, which had been filter-sterilized prior to

578     incubation. Double plus symbols indicate that the live replicates were 10-60 ‰ heavier than the

579     corresponding dead replicates, and triple plus symbols indicate that the live replicates were 100-

580     700 ‰ heavier than the dead replicates.


581     Methanogenesis from acetate was detected in GOR34-spring3-2011, but only from the carboxyl

582     group (C1 atom) of acetate. Acetoclastic methanogenesis generates methane by reduction of the

583     methyl group (C2 atom), so this result suggests that the carboxyl group of acetate was first

584     oxidized to inorganic carbon (carbon dioxide, bicarbonate, or carbonate) and then utilized by

585     methanogens. Indeed, oxidation of acetate to carbon dioxide was observed in GOR34-spring3-

586     2011. Methane oxidation to carbon dioxide was also observed at least once in all three springs.

587     Methane oxidation signals were greatly increased when incubations were exposed to oxygen

588     (indicated by the third plus symbol in GOR34-spring1-2012 in Table 6) and were completely

589     eliminated by addition of a reducing agent. This result strongly indicates that the measured

590     methane oxidation was aerobic, although some activity was also detected even without oxygen

591     exposure, leaving open the possibility that alternative electron acceptors can be used as well.


592     **Discussion**

593     *Methanogenesis*

594     This study of ultrabasic springs at the Voltri Massif provides the first metagenomic evidence for

595     methanogens and methanogenesis pathways in a continental serpentinizing system. All genes

596     required for methanogenesis are present in the combined-spring assembly (Figure 3), and many

597     of these were collected in a single genomic bin corresponding to the Methanobacteriaceae taxa

598     that are enriched in the ultrabasic springs (Figure 2). The combined-spring assembly includes the

599     genetic potential for the utilization of carbon dioxide, formate, acetate, and methanol as

600     methanogenic substrates. Some of the genes involved in carbon dioxide, formate, and acetate

601     utilization were also found in genomic bins for non-methanogens and in the adjacent rivers, but

602     the two proteins required for the utilization of methanol as a methanogenic substrate were unique

603     to the Methanobacteriaceae bin. There are no previous examples of organisms within family

604     Methanobacteriaceae that can use methanol as a substrate, so this result should be confirmed with

605     additional genomic and experimental data. Nevertheless, methanol and formate are among the

606     single carbon compounds predicted to be abiogenic products of serpentinization (Shock, 1992;

607     Shock & Schulte, 1998; Seewald, Zolotov & McCollom, 2006), so these metagenomic data are

608     consistent with the possibility that Methanobacteriaceae methanogens at this site are directly

609     supported by subsurface serpentinization reactions.


610     We detected active methanogenesis in GOR34-spring3 (Table 6), which is assumed to have

611     occurred at close to the in situ pH of the spring (pH 11.8). To our knowledge, this would

612     represent the highest pH known to support methanogenesis, although no pH measurements were

613   made during or after the incubation experiments. Methane produced during this experiment was

614   derived from the carboxyl group (C1 atom) of acetate, and no methane was produced from the

615   methyl group (C2 atom). This is contrary to the expectation for acetate-based methanogenesis,

616   which produces methane by reduction of the methyl group. Furthermore, Methanobacteriaceae

617   methanogens are not known to be able to use acetate as a substrate for methanogenesis.

618   Therefore, it is more likely that the carboxyl group was oxidized to inorganic carbon (activity for

619   which was also detected in this spring; see Table 6) and subsequently utilized for

620   methanogenesis. Unfortunately, methanogenesis from inorganic carbon (carbon dioxide,

621   bicarbonate, or carbonate) was not measured during that experiment, and subsequent experiments

622   failed to detect methanogenesis from any substrates.


623   Thus, the genomic evidence for the potential of inorganic carbon, formate, acetate, and methanol

624   to support methanogenesis by Methanobacteriaceae organisms remains only incompletely

625   confirmed by metabolic activity experiments. Future experimental studies designed to test the

626   utilization of specific carbon compounds during active methanogenesis at high pH would provide

627   stronger evidence for the link between subsurface carbon sources and methanogens in

628   serpentinite springs.


629   Methanobacteriaceae (as well as orders Methanosarcinales and Methanocellales) have been

630   previously detected in travertine deposits formed by ultrabasic springs of the Voltri Massif

631   nearby, but not the same as, the springs described here (Quéméneur et al., 2015). The presence of

632   Methanobacteriaceae in the carbonate deposits raises the question of whether they primarily

633   inhabit surface or subsurface environments. Methanobacteriaceae and predicted proteins in the

634   methanogenesis pathway are most abundant in the BR2 spring (Table 4 and Supplementary File

635   S8), which is not submerged by an overlying pool, does not appear to come into contact with

636    surface travertine deposits, and can be sampled directly from the spring's source (Figure 1).

637    Therefore, the high abundance of Methanobacteriaceae in this spring suggests a subsurface

638    habitat for these methanogens, although it is possible that these organisms could be active in both

639    subsurface springs and surface deposits where biofilms may create anoxic micro-environments.


640    *Hydrogen metabolism*

641    In addition to hydrogen gas consumption by methanogens, the Voltri Massif springs encode

642    abundant genes for hydrogen oxidation (Table 4). [NiFe]-hydrogenases associated with

643    Comamonadaceae and Desulfovibrionales, which dominate the bacterial communities of the

644    ultrabasic springs, were abundant in the metagenomic data from these springs. The

645    Comamonadaceae 16S rRNA, [NiFe]-hydrogenase, and Rubisco sequences are highly similar to

646    those reported for Comamonadaceae from other continental sites of serpentinization including the

647    Cabeço de Vide Aquifer (Tiago & Veríssimo, 2013), the Tablelands Ophiolite (Brazelton, Nelson

648    & Schrenk, 2012), The Cedars (Suzuki et al., 2014), and the Coast Range Ophiolite Microbial

649    Observatory (Crespo-Medina et al., 2014). The presence of these organisms in alkaline springs

650    appears to be a consistent indicator of serpentinization (Schrenk, Brazelton & Lang, 2013), and

651    the genus *Serpentinomonas* has been proposed for them (Suzuki et al., 2014). *Serpentinomonas*

652    16S rRNA sequences have also been recovered from surface carbonate deposits formed by a

653    spring at the Voltri Massif near to those reported in this study (Quéméneur et al., 2015), which is

654    consistent with the interpretation that these organisms live at the surface or the shallow

655    subsurface where serpentinizing fluids mix with the oxygenated atmosphere (Schrenk, Brazelton

656    & Lang, 2013; Brazelton et al., 2013).


657    By contrast, the Desulfovibrionales represented by ESOM Bin 5 appear to be anaerobic,

658    hydrogen-utilizing, sulfate-reducing bacteria and may live in anoxic, deep subsurface habitats

659    underlying the springs. These bacteria, like the methanogens described above, were most

660    abundant in the BR2 spring, which also contains higher sulfide levels. Therefore, this spring

661    appears to provide more direct access to organisms that may have been recently active in anoxic,

662    methanogenic, and sulfate-reducing subsurface habitats.

663    [FeFe]-hydrogenase sequences were also abundant, and many were associated with the

664    unclassified Bacteroidetes group enriched in the ultrabasic springs (Figure 2). These genes are

665    typically involved in fermentative hydrogen production and appear to have a broad taxonomic

666    distribution in our study, as illustrated by their wide distribution in the ESOM (Figure 4B).

667    Clostridiales, which were also enriched in the springs, are expected to encode some of the [FeFe]-

668    hydrogenase sequences (e.g. Mei et al., 2014), but no ESOM bins with taxonomic markers

669    consistently classified as Clostridiales could be identified.

670    These results are consistent with previous observations of abundant hydrogenases in

671    metagenomic data from marine and continental serpentinite springs (Brazelton, Nelson &

672    Schrenk, 2012). Phylogenetic analyses of those sequences indicated that [NiFe]-hydrogenases

673    associated with Comamonadaceae and [FeFe]-hydrogenases associated with Firmicutes are likely

674    to be involved in hydrogen consumption and production, respectively. The co-occurence of

675    Betaproteobacteria (often order Burkholderiales, including family Comamonadaceae) and

676    Firmicutes (in particular order Clostridiales) is a common feature of serpentinite springs and

677    perhaps the mixing of deep subsurface and surface fluids in general (reviewed by Schrenk,

678    Brazelton & Lang, 2013; see also Purkamo et al., 2016). This report contributes to that emerging

679    trend by providing another example of Comamonadaceae organisms with [NiFe]-hydrogenases

680    and also reports [NiFe]-hydrogenases in Desulfovibrionales and [FeFe]-hydrogenases in

681    Bacteroidetes in serpentinite springs for the first time.

682   *Methanotrophy*

683   Aerobic methane oxidation was detected in all three of the ultrabasic springs in this study (Table

684   6), and the Methylococcaceae family of aerobic methanotrophs was among the most enriched

685   taxa in the ultrabasic springs (Figure 2). The genomic bin corresponding to this family (ESOM

686   Bin 4) includes sequences encoding the soluble form of methane monooxygenase (mmoX).

687   Although methane seems to be the most likely substrate for mmoX here, this enzyme can also

688   utilize a variety of larger organic compounds (Colby, Stirling & Dalton, 1977). Curiously, no

689   sequences encoding the particulate form of methane monooxygenase (pmoABC) were detected in

690   the combined-spring assembly, as would be expected for members of Methylococcaceae (Ward et

691   al., 2004). The absence of sequences in a metagenomic assembly cannot be used to draw any firm

692   conclusions, but the complete absence of even a single pmoA sequence in an assembly containing

693   89 Mb is striking. The *Methylocella* and *Methyloferula* genera lack pmoA, but these

694   methanotrophs are Alphaproteobacteria, highly divergent from the gammaproteobacterial

695   Methylococcaceae detected here.


696   Being aerobic, these bacteria are unlikely to be active in anoxic subsurface environments.

697   However, a completely surface-exposed habitat also seems unlikely because they were not

698   detected in a travertine deposit formed by one of these springs (Quéméneur et al., 2015).

699   Therefore, they may inhabit a shallow subsurface mixing zone where low levels of oxygen reach

700   methane-rich subsurface water, similar to the inferred habitat for the aerobic, hydrogen-oxidizing

701   Comamonadaceae described above.


702   *Fermentation*

703    Multiple fermentation pathways were enriched in the ultrabasic springs (Table 4), and

704    fermentation genes were widely distributed among metagenomic fragments in the ESOM (Figure

705    4). Candidate Division OD1 (now called Parcubacteria) was enriched in the ultrabasic springs

706    here and has also been found in serpentinite springs at The Cedars that are expected to represent

707    deeper, subsurface habitats (Suzuki et al., 2013). These bacteria are expected to be involved with

708    fermentation (Hu et al., 2016), but we were unable to verify this due to the lack of any ESOM

709    bins with a strong representation of Candidate Division OD1. Additional work will be required to

710    better characterize the genomes of these organisms, which have been reported to be small and

711    lacking genes typically thought to be essential for basic cellular processes (Nelson & Stegen,

712    2015; Brown et al., 2015).


713    An uncultured group of Bacteroidetes known as 'ML635J-40 aquatic group' was also enriched in

714    the ultrabasic springs and was well-represented in the ESOM. The bin of metagenomic sequences

715    associated with this group (ESOM Bin 3) encoded proteins predicted to be involved in pyruvate

716    fermentation, sugar uptake, and hydrogen production. A member of ML635J-40 aquatic group

717    was recently identified in a pH 10 anaerobic reactor in which it was predicted to be primarily

718    responsible for hydrolysis of organic matter and supply of hydrogen gas to methanogens (Nolla-

719    Ardèvol, Strous & Tegetmeyer, 2015). The Bacteroidetes in the BR2 and GOR34 ultrabasic

720    springs could have a similar syntrophic relationship with methanogens, although the source of

721    hydrolyzable organic matter in this case is unclear. The ML635J-40 aquatic group was not

722    detected in a previous study of surface travertine deposits (Quéméneur et al., 2015), indicating

723    that they do not inhabit a completely surface-exposed environment. Although great care was

724    taken to avoid surface contamination during field sampling, including measures to sample the

725    subsurface source of the spring rather than the overlying pool (see Methods), it remains possible

726    that the Bacteroidetes sequences in the GOR34 springs represent pool-dwelling organisms that

727    are dependent on terrestrial organic matter that falls into the surface-exposed pools. This

728    interpretation is consistent with the lower abundance of Bacteroidetes in the BR2 spring

729    (Supplementary File S6), which is not submerged by an overlying pool (Figure 1). It is also

730    possible that they inhabit a shallow subsurface transition zone, similar to that inferred above for

731    the Comamonadaceae and Methylococcaceae. Future studies should investigate the carbon source

732    for these organisms in order to determine whether they are supported by compounds synthesized

733    by subsurface serpentinization-associated reactions.


734    *Formate and acetate metabolism*

735    Both formate and acetate are present at elevated concentrations in fluids venting from the

736    chimneys of the Lost City hydrothermal field (Lang et al., 2010), and formate is expected to be

737    an abiogenic product of serpentinization-associated reactions (Shock, 1992; Shock & Schulte,

738    1998; McCollom & Seewald, 2001, 2003; Seewald, Zolotov & McCollom, 2006). Metagenomic

739    sequences predicted to encode proteins associated with formate and acetate metabolism were

740    abundant (Supplementary File S8) and widespread among taxa (green and blue points,

741    respectively, in Figure 4B) in the combined-spring assembly, including the ESOM bins

742    representing Methylococcaceae and Desulfovibrionales (Table 5). The wide taxonomic

743    distribution of these genes (formate c-acetyltransferase (Pfl), formate dehydrogenase (Fdh),

744    acetate kinase (Ack), phosphotransacetylase (Pta), AMP-forming acetyl-CoA synthetase (Acs),

745    and carbon monoxide dehydrogenase / acetyl-CoA decarbonylase/synthase (Cdh/ACDS)) clearly

746    indicates that they are not restricted to methanogens in this system and that this ecosystem

747    features a broad metabolic potential for utilizing formate and acetate. We measured organic acid

748    concentrations in Voltri Massif springs in 2010, before the microbiological studies described here.

749    Formate and acetate were present at low but detectable concentrations (0.4 and 2.7 µM,

750    respectively) at GOR34-spring3 and were below detection in all other springs. It is unclear

751    whether the low levels indicate a lack of production or rapid consumption, but our initial

752    metabolic activity assays have demonstrated that acetate can be oxidized to carbon dioxide at in

753    situ conditions characterized by extremely high pH. Additional work is required to test whether

754    the availability of formate and acetate are correlated with the activity of specific genes, including

755    those identified by this study. Furthermore, future studies should attempt to demonstrate in the

756    case of each group of organisms reported here whether formate and acetate are consumed or

757    produced, whether they are cycled intracellularly, and whether they are obtained from or released

758    into the environment.


759    **Conclusions**

760    The potential significance of serpentinization-supported microbial ecosystems has been widely

761    recognized since the discovery of the Lost City hydrothermal field (Kelley et al., 2001), but we

762    are still in the early stages of characterizing the organisms and pathways that may benefit from

763    the subsurface geochemical reactions associated with serpentinization. This study of ultrabasic

764    springs at the Voltri Massif, Italy provides the first evidence for active methanogenesis and

765    aerobic methanotrophy in continental serpentinite springs as well as genomic information about

766    the specific organisms likely to be responsible for these processes. Both methanogenesis and

767    methanotrophy were active at very high pH, potentially raising the upper pH limit known to

768    support both processes. These potential activity experiments were not intended to quantify in situ

769    rates, however, and do not preclude the possibility that the vast majority of methane in these

770    springs is produced by non-biological processes. The data reported here also contribute additional

771    metagenomic evidence for the importance of hydrogenases involved in both the consumption and

772    production of hydrogen gas in serpentinite systems.

773   The unavailability of inorganic carbon at the extremely high pH of these springs is likely to be a

774   limiting factor for autotrophic activity, and small organic compounds such as formate and acetate

775   may serve as the primary carbon sources in such ecosystems. Accordingly, we found evidence for

776   inorganic carbon as a substrate for methanogenesis only after it had been liberated by the

777   oxidation of acetate, perhaps by a syntrophic partner. The genetic potential for metabolism of

778   formate and acetate is widespread among taxa in this environment, which is consistent with an

779   ecosystem being at least partially supported by organic carbon synthesized by subsurface

780   serpentinization-associated reactions.


781   These genomic and metabolic data were obtained despite the extremely low biomass of these

782   springs, suggesting that our results are the first clues into the few organisms capable of survival,

783   and potentially growth, in the harsh conditions of the pH 12 springs. Because many of these

784   organisms were only found in the spring waters and not in surface travertine deposits

785   (Quéméneur et al., 2015), they are likely to represent inhabitants of subsurface environments in

786   the serpentinite rocks underlying the springs. We anticipate that the metagenomic inventories and

787   data analysis tools described here will provide a foundation for future studies to investigate how

788   these organisms make a living in these conditions and to test whether such ecosystems can be

789   supported solely by subsurface serpentinization-associated reactions.

802 **References**

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

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

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

815 Anders S., Pyl PT., Huber W. 2015. HTSeq--a Python framework to work with high-throughput
816     sequencing data. *Bioinformatics (Oxford, England)* 31:166–9. DOI:
817     10.1093/bioinformatics/btu638.

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

821 Boisvert S., Raymond F., Godzaridis E., Laviolette F., Corbeil J. 2012. Ray Meta: scalable de
822     novo metagenome assembly and profiling. *Genome biology* 13:R122. DOI: 10.1186/gb-
823     2012-13-12-r122.

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

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

830     Brazelton WJ., Schrenk MO., Kelley DS., Baross JA. 2006. Methane- and sulfur-metabolizing
831         microbial communities dominate the Lost City hydrothermal field ecosystem. *Appl Environ*
832         *Microbiol* 72:6257–6270. DOI: 10.1128/AEM.00574-06.

833     Brazelton WJ., Mehta MP., Kelley DS., Baross JA. 2011. Physiological Differentiation within a
834         Single-Species Biofilm Fueled by Serpentinization. *mBio* 2:e00127–11. DOI:
835         10.1128/mBio.00127-11.

836     Brazelton WJ., Morrill PL., Szponar N., Schrenk MO. 2013. Bacterial communities associated
837         with subsurface geochemical processes in continental serpentinite springs. *Applied and*
838         *environmental microbiology* 79:3906–16. DOI: 10.1128/AEM.00330-13.

839     Brazelton WJ., Nelson B., Schrenk MO. 2012. Metagenomic evidence for $H_2$ oxidation and $H_2$
840         production by serpentinite-hosted subsurface microbial communities. *Frontiers in*
841         *microbiology* 2:268. DOI: 10.3389/fmicb.2011.00268.

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

845     Bruni J., Canepa M., Chiodini G., Cioni R., Cipolli F., Longinelli A., Marini L., Ottonello G.,
846         Vetuschi M. 2002. Irreversible water–rock mass transfer accompanying the generation of the
847         neutral, Mg–HCO3 and high-pH, Ca–OH spring waters of the Genova province, Italy.
848         *Applied Geochemistry* 17:455–474. DOI: 10.1016/S0883-2927(01)00113-5.

849     Cardace D., Hoehler T., McCollom T., Schrenk M., Carnevale D., Kubo M., Twing K. 2013.
850         Establishment of the Coast Range ophiolite microbial observatory (CROMO): drilling
851         objectives and preliminary outcomes. *Scientific Drilling* 16:45–55. DOI: 10.5194/sd-16-45-
852         2013.

853     Caspi R., Altman T., Billington R., Dreher K., Foerster H., Fulcher CA., Holland TA., Keseler
854         IM., Kothari A., Kubo A., Krummenacker M., Latendresse M., Mueller LA., Ong Q., Paley
855         S., Subhraveti P., Weaver DS., Weerasinghe D., Zhang P., Karp PD. 2014. The MetaCyc
856         database of metabolic pathways and enzymes and the BioCyc collection of
857         Pathway/Genome Databases. *Nucleic acids research* 42:D459–71. DOI:
858         10.1093/nar/gkt1103.

859 Cipolli F., Gambardella B., Marini L., Ottonello G. 2004. Geochemistry of high-pH waters from
860      serpentinites of the Gruppo di Voltri (Genova, Italy) and reaction path modeling of $CO_2$
861      sequestration in serpentinite aquifers. *Applied Geochemistry* 19:787–802. DOI:
862      10.1016/j.apgeochem.2003.10.007.

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

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

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

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

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

876 Dick GJ., Andersson AF., Baker BJ., Simmons SL., Thomas BC., Yelton AP., Banfield JF. 2009.
877      Community-wide analysis of microbial genome sequence signatures. *Genome biology*
878      10:R85. DOI: 10.1186/gb-2009-10-8-r85.

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

881 Ferry JG. 2010. How to make a living by exhaling methane. *Annual review of microbiology*
882      64:453–73. DOI: 10.1146/annurev.micro.112408.134051.

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

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

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

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

896    Huse SM., Young VB., Morrison HG., Antonopoulos DA., Kwon J., Dalal S., Arrieta R., Hubert
897          NA., Shen L., Vineis JH., Koval JC., Sogin ML., Chang EB., Raffals LE. 2014a.
898          Comparison of brush and biopsy sampling methods of the ileal pouch for assessment of
899          mucosa-associated microbiota of human subjects. *Microbiome* 2:5. DOI: 10.1186/2049-
900          2618-2-5.

901    Huse SM., Mark Welch DB., Voorhis A., Shipunova A., Morrison HG., Eren A., Sogin ML.,
902          Quince C., Lanzen A., Davenport R., Turnbaugh P., Edgar R., Wang Q., Garrity G., Tiedje J.,
903          Cole J., Pruesse E., Quast C., Knittel K., Fuchs B., Ludwig W., Peplies J., Glockner F.,
904          Meyer F., Paarmann D., D'Souza M., Olson R., Glass E., Kubal M., Paczian T., Rodriguez
905          A., Stevens R., Wilke A., Schloss P., Westcott S., Ryabin T., Hall J., Hartmann M., Hollister
906          E., Lesniewski R., Oakley B., Parks D., Robinson C., Caporaso J., Kuczynski J., Stombaugh
907          J., Bittinger K., Bushman F., Costello E., Fierer N., Pena A., Goodrich J., Gordon J., Fox P.,
908          Hendler J., Huse S., Huber J., Morrison H., Sogin M., Welch DM., Huse S., Dethlefsen L.,
909          Huber J., Welch DM., Relman D., Sogin M., Eren A., Maignien L., Sul W., Murphy L., Grim
910          S., Morrison H., Sogin M., Huse S., Welch DM., Morrison H., Sogin M., Hao X., Jiang R.,
911          Chen T., Eren A., Zozaya M., Taylor C., Dowd S., Martin D., Ferris M., Bik H., Porazinska
912          D., Creer S., Caporaso J., Knight R., Thomas W., Peterson J., Garges S., Giovanni M.,
913          McInnes P., Wang L., Schloss J., Bonazzi V., McEwen J., Wetterstrand K., Young V., Raffals
914          L., Huse S., Vital M., Dai D., Schloss P., Brulc J., Antonopoulos D., Arrieta R., Kwon J.,
915          Amend A., Oliver T., Amaral-Zettler L., Boetius A., Fuhrman J., Horner-Devine M., Huse
916          S., Welch D., Martiny A., Ramette A., Sul W., Oliver T., Ducklow H., Amaral-Zettler L.,
917          Sogin M., Bowen J., Morrison H., Hobbie J., Sogin M., Gobet A., Boer S., Huse S.,
918          Beusekom J van., Quince C., Sogin M., Boetius A., Ramette A. 2014b. VAMPS: a website
919          for visualization and analysis of microbial population structures. *BMC Bioinformatics* 15:41.
920          DOI: 10.1186/1471-2105-15-41.

921    Hyatt D., Chen G-L., LoCascio PF., Land ML., Larimer FW., Hauser LJ., Delcher A., Bratke K.,
922          Powers E., Salzberg S., Lukashin A., Borodovsky M., Benson D., Karsch-Mizrachi I.,
923          Lipman D., Ostell J., Sayers E., Larsen T., Krogh A., Zhu H., Hu G., Yang Y., Wang J., She
924          Z., Ou H., Guo F., Zhang C., Tech M., Pfeifer N., Morgenstern B., Meinicke P., Hu G.,
925          Zheng X., Zhu H., She Z., Altschul S., Gish W., Miller W., Myers E., Lipman D.,
926          VerBerkmoes N., Shah M., Lankford P., Pelletier D., Strader M., Tabb D., McDonald W.,
927          Barton J., Hurst G., Hauser L., Badger J., Olsen G., Rudd K., Rutherford K., Parkhill J.,
928          Crook J., Horsnell T., Rice P., Rajandream M., Barrell B., Bellman R., Dam P., Olman V.,
929          Harris K., Su Z., Xu Y., Aivaliotis M., Gevaert K., Falb M., Tebbe A., Konstantinidis K.,
930          Bisle B., Klein C., Martens L., Staes A., Timmerman E., Hu G., Zheng X., Yang Y., Ortet P.,
931          She Z., Zhu H., Gattiker A., Dessimoz C., Schneider A., Xenarios I., Pagni M., Rougemont

932        J. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification.
933            *BMC Bioinformatics* 11:119. DOI: 10.1186/1471-2105-11-119.

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

937    Kelley DS., Karson JA., Früh-Green GL., Yoerger DR., Shank TM., Butterfield DA., Hayes JM.,
938            Schrenk MO., Olson EJ., Proskurowski G., Jakuba M., Bradley A., Larson B., Ludwig K.,
939            Glickson D., Buckman K., Bradley AS., Brazelton WJ., Roe K., Elend MJ., Delacour A.,
940            Bernasconi SM., Lilley MD., Baross JA., Summons RE., Sylva SP. 2005. A serpentinite-
941            hosted ecosystem: the Lost City hydrothermal field. *Science* 307:1428–34. DOI:
942            10.1126/science.1102556.

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

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

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

954    Langmead B., Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature methods*
955            9:357–9. DOI: 10.1038/nmeth.1923.

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

958    McCollom T., Seewald J. 2001. A reassessment of the potential for reduction of dissolved $CO_2$ to
959            hydrocarbons during serpentinization of olivine. *Geochimica et Cosmochimica Acta*
960            65:3769–3778. DOI: 10.1016/S0016-7037(01)00655-X.

961    McCollom TM., Seewald JS. 2003. Experimental study of the hydrothermal reactivity of organic
962            acids and acid anions: II. Acetic acid, acetate, and valeric acid. *GEOCHIMICA ET*
963            *COSMOCHIMICA ACTA* 67:3645–3664. DOI: 10.1016/S0016-7037(03)00135-2.

964    McMurdie PJ., Holmes S. 2013. phyloseq: an R package for reproducible interactive analysis and
965            graphics of microbiome census data. *PloS one* 8:e61217. DOI:
966            10.1371/journal.pone.0061217.

967   McMurdie PJ., Holmes S. 2014. Waste not, want not: why rarefying microbiome data is
968       inadmissible. *PLoS computational biology* 10:e1003531. DOI:
969       10.1371/journal.pcbi.1003531.

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

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

978   Meyer F., Paarmann D., D'Souza M., Olson R., Glass EM., Kubal M., Paczian T., Rodriguez A.,
979       Stevens R., Wilke A., Wilkening J., Edwards RA. 2008. The metagenomics RAST server - a
980       public resource for the automatic phylogenetic and functional analysis of metagenomes.
981       *BMC bioinformatics* 9:386. DOI: 10.1186/1471-2105-9-386.

982   Morrill PL., Kuenen JG., Johnson OJ., Suzuki S., Rietze A., Sessions AL., Fogel ML., Nealson
983       KH. 2013. Geochemistry and geobiology of a present-day serpentinization site in California:
984       The Cedars. *Geochimica et Cosmochimica Acta* 109:222–240. DOI:
985       10.1016/j.gca.2013.01.043.

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

991   Nelson MC., Morrison HG., Benjamino J., Grim SL., Graf J., Stahl D., Lane D., Olsen G., Pace
992       N., Stahl D., Lane D., Olsen G., Pace N., Rivière D., Desvignes V., Pelletier E.,
993       Chaussonnerie S., Guermazi S., Nelson M., Morrison M., Yu Z., Ley R., Bäckhed F.,
994       Turnbaugh P., Lozupone C., Knight R., Costello E., Lauber C., Hamady M., Fierer N.,
995       Gordon J., Sogin M., Morrison H., Huber J., Welch D., Huse S., Turnbaugh P., Hamady M.,
996       Yatsunenko T., Cantarel B., Duncan A., Huse S., Huber J., Morrison H., Sogin M., Welch D.,
997       Huse S., Welch D., Morrison H., Sogin M., Quince C., Lanzén A., Curtis T., Davenport R.,
998       Hall N., Reeder J., Knight R., Loman N., Misra R., Dallman T., Constantinidou C., Gharbia
999       S., Jünemann S., Sedlazeck F., Prior K., Albersmeier A., John U., Bartram A., Lynch M.,
1000      Stearns J., Moreno-Hagelsieb G., Neufeld J., Caporaso J., Lauber C., Walters W., Berg-
1001      Lyons D., Huntley J., Degnan P., Ochman H., Gloor G., Hummelen R., Macklaim J.,
1002      Dickson R., Fernandes A., Ram J., Karim A., Sendler E., Kato I., Kozich J., Westcott S.,
1003      Baxter N., Highlander S., Schloss P., Eren A., Vineis J., Morrison H., Sogin M., Caporaso J.,
1004      Lauber C., Walters W., Berg-Lyons D., Lozupone C., Claesson M., Wang Q., O'Sullivan O.,

1005    Greene-Diniz R., Cole J., Luo C., Tsementzi D., Kyrpides N., Read T., Konstantinidis K., Yu
1006    Z., Morrison M., Marteinsson V., Rúnarsson Á., Stefánsson A., Thorsteinsson T.,
1007    Jóhannesson T., Caporaso J., Kuczynski J., Stombaugh J., Bittinger K., Bushman F.,
1008    DeSantis T., Hugenholtz P., Larsen N., Rojas M., Brodie E., McDonald D., Price M.,
1009    Goodrich J., Nawrocki E., DeSantis T., Werner J., Koren O., Hugenholtz P., DeSantis T.,
1010    Walters W., Bokulich N., Subramanian S., Faith J., Gevers D., Gordon J., Reysenbach A-L.,
1011    Giver L., Wickham G., Pace N., Bergmann G., Bates S., Eilers K., Lauber C., Caporaso J.,
1012    Edgar R., Haas B., Clemente J., Quince C., Knight R., Engelbrektson A., Kunin V.,
1013    Wrighton K., Zvenigorodsky N., Chen F., Haas B., Gevers D., Earl A., Feldgarden M., Ward
1014    D., Pruesse E., Quast C., Knittel K., Fuchs B., Ludwig W., Worthen P., Gode C., Graf J.,
1015    Laufer A., Siddall M., Graf J. 2014. Analysis, Optimization and Verification of Illumina-
1016    Generated 16S rRNA Gene Amplicon Surveys. *PLoS ONE* 9:e94249. DOI:
1017    10.1371/journal.pone.0094249.

1018    Nelson WC., Stegen JC. 2015. The reduced genomes of Parcubacteria (OD1) contain signatures
1019    of a symbiotic lifestyle. *Frontiers in microbiology* 6:713. DOI: 10.3389/fmicb.2015.00713.

1020    Nolla-Ardèvol V., Strous M., Tegetmeyer HE. 2015. Anaerobic digestion of the microalga
1021    *Spirulina* at extreme alkaline conditions: biogas production, metagenome, and
1022    metatranscriptome. *Frontiers in microbiology* 6:597. DOI: 10.3389/fmicb.2015.00597.

1023    Ogata H., Goto S., Sato K., Fujibuchi W., Bono H., Kanehisa M. 1999. KEGG: Kyoto
1024    Encyclopedia of Genes and Genomes. *Nucleic acids research* 27:29–34. DOI:
1025    10.1093/nar/27.1.29.

1026    Parks DH., Imelfort M., Skennerton CT., Hugenholtz P., Tyson GW. 2015. CheckM: assessing the
1027    quality of microbial genomes recovered from isolates, single cells, and metagenomes.
1028    *Genome Research* 25:1043–1055. DOI: 10.1101/gr.186072.114.

1029    Pikuta E V., Hoover RB., Bej AK., Marsic D., Whitman WB., Cleland D., Krader P. 2003.
1030    *Desulfonatronum thiodismutans* sp. nov., a novel alkaliphilic, sulfate-reducing bacterium
1031    capable of lithoautotrophic growth. *International journal of systematic and evolutionary*
1032    *microbiology* 53:1327–32. DOI: 10.1099/ijs.0.02598-0.

1033    Prestat E., David MM., Hultman J., Taş N., Lamendella R., Dvornik J., Mackelprang R., Myrold
1034    DD., Jumpponen A., Tringe SG., Holman E., Mavromatis K., Jansson JK. 2014. FOAM
1035    (Functional Ontology Assignments for Metagenomes): a Hidden Markov Model (HMM)
1036    database with environmental focus. *Nucleic acids research* 42:e145. DOI:
1037    10.1093/nar/gku702.

1038    Proskurowski G., Lilley MD., Seewald JS., Früh-Green GL., Olson EJ., Lupton JE., Sylva SP.,
1039    Kelley DS. 2008. Abiogenic hydrocarbon production at Lost City Hydrothermal Field.
1040    *Science (New York, N.Y.)* 319:604–7. DOI: 10.1126/science.1151194.

1041    Pruesse E., Peplies J., Glöckner FO. 2012. SINA: accurate high-throughput multiple sequence
1042         alignment of ribosomal RNA genes. *Bioinformatics (Oxford, England)* 28:1823–9. DOI:
1043         10.1093/bioinformatics/bts252.

1044    Purkamo L., Bomberg M., Kietäväinen R., Salavirta H., Nyyssönen M., Nuppunen-Puputti M.,
1045         Ahonen L., Kukkonen I., Itävaara M. 2016. Microbial co-occurrence patterns in deep
1046         Precambrian bedrock fracture fluids. *Biogeosciences* 13:3091–3108. DOI: 10.5194/bg-13-
1047         3091-2016.

1048    Quéméneur M., Bes M., Postec A., Mei N., Hamelin J., Monnin C., Chavagnac V., Payri C.,
1049         Pelletier B., Guentas-Dombrowsky L., Gérard M., Pisapia C., Gérard E., Ménez B., Ollivier
1050         B., Erauso G. 2014. Spatial distribution of microbial communities in the shallow submarine
1051         alkaline hydrothermal field of the Prony Bay, New Caledonia. *Environmental Microbiology*
1052         *Reports* 0033. DOI: 10.1111/1758-2229.12184.

1053    Quéméneur M., Palvadeau A., Postec A., Monnin C., Chavagnac V., Ollivier B., Erauso G. 2015.
1054         Endolithic microbial communities in carbonate precipitates from serpentinite-hosted
1055         hyperalkaline springs of the Voltri Massif (Ligurian Alps, Northern Italy). *Environmental*
1056         *science and pollution research international* 22:13613–24. DOI: 10.1007/s11356-015-4113-
1057         7.

1058    Robinson MD., McCarthy DJ., Smyth GK. 2010. edgeR: a Bioconductor package for differential
1059         expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)*
1060         26:139–40. DOI: 10.1093/bioinformatics/btp616.

1061    Sánchez-Murillo R., Gazel E., Schwarzenbach EM., Crespo-Medina M., Schrenk MO., Boll J.,
1062         Gill BC. 2014. Geochemical evidence for active tropical serpentinization in the Santa Elena
1063         Ophiolite, Costa Rica: An analog of a humid early Earth? *Geochem Geophys Geosyst*
1064         15:1783–1800. DOI: 10.1002/2013GC005213.

1065    Schloss PD., Westcott SL., Ryabin T., Hall JR., Hartmann M., Hollister EB., Lesniewski RA.,
1066         Oakley BB., Parks DH., Robinson CJ., Sahl JW., Stres B., Thallinger GG., Van Horn DJ.,
1067         Weber CF. 2009. Introducing mothur: Open Source, Platform-independent, Community-
1068         Supported Software for Describing and Comparing Microbial Communities. *Appl. Environ.*
1069         *Microbiol.* 75:7537–7541. DOI: 10.1128/AEM.01541-09.

1070    Schloss PD., Gevers D., Westcott SL. 2011. Reducing the Effects of PCR Amplification and
1071         Sequencing Artifacts on 16S rRNA-Based Studies. *PLoS ONE* 6:e27310. DOI:
1072         10.1371/journal.pone.0027310.

1073    Schrenk MO., Kelley DS., Bolton SA., Baross JA. 2004. Low archaeal diversity linked to
1074         subseafloor geochemical processes at the Lost City Hydrothermal Field, Mid-Atlantic
1075         Ridge. *Environ Microbiol* 6:1086–1095. DOI: 10.1111/j.1462-2920.2004.00650.x.

1076    Schrenk MO., Brazelton WJ., Lang SQ. 2013. Serpentinization, Carbon, and Deep Life. *Reviews*
1077         *in Mineralogy and Geochemistry* 75:575–606. DOI: 10.2138/rmg.2013.75.18.

1078   Schwarzenbach EM., Lang SQ., Früh-Green GL., Lilley MD., Bernasconi SM., Méhay S. 2013.
1079         Sources and cycling of carbon in continental, serpentinite-hosted alkaline springs in the
1080         Voltri Massif, Italy. *Lithos* 177:226–244. DOI: 10.1016/j.lithos.2013.07.009.

1081   Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics (Oxford,*
1082         *England)* 30:2068–9. DOI: 10.1093/bioinformatics/btu153.

1083   Seewald JS., Zolotov MY., McCollom T. 2006. Experimental investigation of single carbon
1084         compounds under hydrothermal conditions. *Geochim Cosmochim Acta* 70:446–460. DOI:
1085         10.1016/j.gca.2005.09.002.

1086   Shock E. 1992. Stability of peptides in high-temperature aqueous solutions. *Geochimica et*
1087         *cosmochimica acta* 56:3481–3491. DOI: 10.1016/0016-7037(92)90392-V.

1088   Shock EL., Schulte MD. 1998. Organic synthesis during fluid mixing in hydrothermal systems.
1089         *Journal of Geophysical Research* 103:28,513–28,527. DOI: 199810.1029/98JE02142.

1090   Sogin ML., Morrison HG., Huber JA., Mark Welch D., Huse SM., Neal PR., Arrieta JM., Herndl
1091         GJ. 2006. Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc*
1092         *Natl Acad Sci U S A* 103:12115–12120. DOI: 10.1073/pnas.0605127103.

1093   Suzuki S., Ishii S., Wu A., Cheung A., Tenney A., Wanger G., Kuenen JG., Nealson KH. 2013.
1094         Microbial diversity in The Cedars, an ultrabasic, ultrareducing, and low salinity
1095         serpentinizing ecosystem. *Proceedings of the National Academy of Sciences of the United*
1096         *States of America* 110:15336–41. DOI: 10.1073/pnas.1302426110.

1097   Suzuki S., Kuenen JG., Schipper K., van der Velde S., Ishii S., Wu A., Sorokin DY., Tenney A.,
1098         Meng X., Morrill PL., Kamagata Y., Muyzer G., Nealson KH. 2014. Physiological and
1099         genomic features of highly alkaliphilic hydrogen-utilizing Betaproteobacteria from a
1100         continental serpentinizing site. *Nature communications* 5:3900. DOI: 10.1038/ncomms4900.

1101   Tiago I., Veríssimo A. 2013. Microbial and functional diversity of a subterrestrial high pH
1102         groundwater associated to serpentinization. *Environ Microbiol* 15:1687–1706. DOI:
1103         10.1111/1462-2920.12034.

1104   Topçuoğlu BD., Stewart LC., Morrison HG., Butterfield DA., Huber JA., Holden JF. In revision.
1105         Hydrogen limitation and syntrophic growth among natural assemblages of thermophilic
1106         methanogens at deep-sea hydrothermal vents. *Frontiers in Microbiology*.

1107   Ultsch A., Mörchen F. 2005. *ESOM-Maps: tools for clustering, visualization, and classification*
1108         *with Emergent SOM*.

1109   Wang DT., Gruen DS., Lollar BS., Hinrichs K., Stewart LC., Holden JF., Hristov AN., Pohlman
1110         JW., Morrill PL., Könneke M., Delwiche KB., Reeves EP., Sutcliffe N., Ritter DJ., Seewald
1111         JS., Jennifer C., Hemond HF., Kubo MD., Hoehler TM., Ono S. 2015. Nonequilibrium
1112         clumped isotope signals in microbial methane. *Science* 348:428–431. DOI:
1113         10.1126/science.aaa4326.

1114   Ward N., Larsen Ø., Sakwa J., Bruseth L., Khouri H., Durkin AS., Dimitrov G., Jiang L., Scanlan
1115         D., Kang KH., Lewis M., Nelson KE., Methé B., Wu M., Heidelberg JF., Paulsen IT., Fouts
1116         D., Ravel J., Tettelin H., Ren Q., Read T., DeBoy RT., Seshadri R., Salzberg SL., Jensen
1117         HB., Birkeland NK., Nelson WC., Dodson RJ., Grindhaug SH., Holt I., Eidhammer I.,
1118         Jonasen I., Vanaken S., Utterback T., Feldblyum T V., Fraser CM., Lillehaug JR., Eisen JA.
1119         2004. Genomic Insights into Methanotrophy: The Complete Genome Sequence of
1120         *Methylococcus capsulatus* (Bath). *PLoS Biology* 2:e303. DOI:
1121         10.1371/journal.pbio.0020303.

1122   Woycheese KM., Meyer-Dombard DR., Cardace D., Argayosa AM., Arcilla CA. 2015. Out of the
1123         dark: transitional subsurface-to-surface microbial diversity in a terrestrial serpentinizing
1124         seep (Manleluag, Pangasinan, the Philippines). *Frontiers in Microbiology* 6:44. DOI:
1125         10.3389/fmicb.2015.00044.

1126   Ye Y., Doak TG. 2009. A Parsimony Approach to Biological Pathway Reconstruction/Inference
1127   for Genomes and Metagenomes. *PLoS Computational Biology* 5:e1000465. DOI:
1128   10.1371/journal.pcbi.1000465.

**Figure Captions**

1129

**Figure 1.** Ultrabasic springs of the Voltri Massif in Italy were sampled at two locations: A)

1130

GOR34, including two springs and an adjacent river and B) BR2, including one spring and an

1131

adjacent river. All water samples were collected by peristaltic pumping through a 0.2 µm filter

1132

cartridge. Photo credits: W.J. Brazelton (A) and B. Nelson (B).

1133

**Figure 2.** Taxonomic classifications of bacterial (**A**) and archaeal (**B**) operational taxonomic units

1134

(OTUs, defined as unique 16S rRNA sequences) that were identified as significantly enriched in

1135

the ultrabasic serpentinite spring GOR34-spring1 compared to the adjacent river. Each data point

1136

in the plots represents one OTU's mean abundance across all samples in the analysis and its

1137

differential abundance in the comparison GOR34-spring1 versus GOR34-river. Mean abundance

1138

is reported as $\log_2$ counts per million; a value of 16 corresponds to 23,000 sequences for that

1139

OTU. Red data points represent OTUs with differential abundances that are significantly enriched

1140

(false discovery rate < 0.05) in either GOR34-spring1 or GOR34-river. The relative abundances

1141

of these significantly enriched OTUs are reported as a percentage of all significantly enriched

1142

sequences and grouped into taxonomic classifications at the family level in the colored bar charts.

1143

Sections of the bar chart were labeled with the corresponding family where possible, and the

1144

abbreviations are defined by bold font in the right-hand legends.

1145

**Figure 3.** Diagram of methanogenesis pathways from carbon dioxide, formate, acetate, methanol,

1146

and methylamines with associated protein homologs identified with KEGG IDs. Green-

1147

highlighted proteins are predicted to occur in the combined-spring metagenomic assembly, and

1148

yellow-highlighted proteins are predicted to occur in both the combined-spring and combined-

1149

1150 river assemblies. The black border indicates proteins identified in the Methanobacteriaceae bin

1151 (ESOM Bin 1 in **Figure 4** and **Table 5**). The diagram is modified from (Ferry, 2010).

1152 **Figure 4.** An emergent self-organizing map (ESOM) constructed with the combined-spring

1153 metagenomic assembly. Each point represents a fragment of metagenomic assembly, and points

1154 are arranged in space according to the similarities of their tetranucleotide compositions. Dark

1155 gray areas are valleys where highly similar sequences cluster together. Light areas of the map are

1156 high topographic ridges that separate dissimilar sequences. The same ESOM is shown here twice:

1157 metagenomic fragments are color-coded in A) to indicate PhyloSift classifications of taxonomic

1158 markers and in B) to show which fragments are predicted to encode protein functions of

1159 particular interest to this study. Based on this information, five bins of sequences were selected

1160 from the map for further evaluation (**Table 5**).

# 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 o.** 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 | 1,350 | 974 | 820 | 1,454 | 965 |
| Predicted protein-encoding genes | 104,651 | 89,606 | 1,326 | 963 | 805 | 1,424 | 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] | 49,794 | 48,915 | 734 | 720 | 533 | 978 | 711 |
| Predicted proteins annotated with UniprotKB | 1,431 | 1,240 | 16 | 8 | 4 | 18 | 14 |
| Predicted proteins annotated with Pfam | 4,810 | 3,811 | 92 | 47 | 58 | 63 | 40 |
| Predicted proteins annotated with HAMAP | 285 | 162 | 4 | 1 | 6 | 4 | 1 |
| Predicted proteins annotated with CLUSTERS | 1,435 | 1,222 | 25 | 9 | 15 | 22 | 7 |
| Predicted proteins with no annotation | 46,896 | 34,256 | 455 | 178 | 189 | 339 | 160 |

[1]*Predicted proteins were first mapped to the KEGG database, and proteins without a good match 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.** Predicted metabolic pathways significantly enriched in ultrabasic spring metagenomes compared to adjacent rivers.

| FOAM Pathway | Mean Coverage BR2 Springs | Mean Coverage GOR34 Springs | Mean Coverage BR2 River | Mean Coverage GOR34 River |
|---|---|---|---|---|
| ABC transporters; cell division transporter | 0.264% | 0.082% | 0.004% | 0.043% |
| Gluconeogenesis in archaea | 0.141% | 0.072% | 0.000% | 0.000% |
| Embden Meyerhof-Parnos (EMP) in Archaea | 0.136% | 0.064% | 0.000% | 0.000% |
| Hydrogen oxidation (aerobic) | 0.089% | 0.111% | 0.000% | 0.000% |
| Pyruvate fermentation (multiple pathways) | 0.076% | 0.053% | 0.000% | 0.022% |
| Fatty acid beta-oxidation (unsaturated) | 0.064% | 0.053% | 0.000% | 0.000% |
| Amino acid metabolism; oxaloacetate to lysine | 0.054% | 0.023% | 0.003% | 0.000% |
| Methanogenesis from formate | 0.048% | 0.018% | 0.001% | 0.000% |
| Hydrogen production | 0.048% | 0.042% | 0.002% | 0.018% |
| Sulfolactate degradation | 0.045% | 0.010% | 0.000% | 0.000% |
| Methanogenesis from CO2 | 0.037% | 0.009% | 0.000% | 0.003% |
| Dissimilatory nitrate reduction to ammonia | 0.024% | 0.009% | 0.000% | 0.000% |
| Denitrification; nitrate to nitrite | 0.023% | 0.020% | 0.000% | 0.001% |
| Fermentation; acetylene degradation | 0.023% | 0.020% | 0.000% | 0.000% |
| Trehalose metabolism in response to cold stress | 0.008% | 0.005% | 0.000% | 0.000% |

**Table 5**(on next page)

Characteristics of ESOM bins identified in Figure 4.

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

**Table 5.** 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 | Methanotrophy; formate metabolism |
| **Bin 5** | 60.34% | 6.03% | 100% | Desulfovibrionales | Hydrogen oxidation; dissimilatory sulfate reduction; formate and acetate metabolism |

# Table 6(on next page)

Detection of metabolic activities in incubation experiments with spring water.

**Table 6.** 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 |
|---|---|---|---|---|---|---|
| $CO_2 \rightarrow CH_4$ | - | - | - | - | nm | - |
| formate $\rightarrow CH_4$ | - | - | - | - | nm | - |
| acetate-$C_1 \rightarrow CH_4$ | - | - | - | - | +++ | - |
| acetate-$C_2 \rightarrow CH_4$ | - | - | - | - | - | - |
| acetate-$C_1 \rightarrow CO_2$ | - | nm | ++ | nm | ++ | nm |
| acetate-$C_2 \rightarrow CO_2$ | - | nm | + | nm | - | nm |
| propionate $\rightarrow CO_2$ | + | nm | - | nm | + | nm |
| $CH_4 \rightarrow CO_2$ | ++ | - | +++ | ++(+)* | ++ | - |

*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.
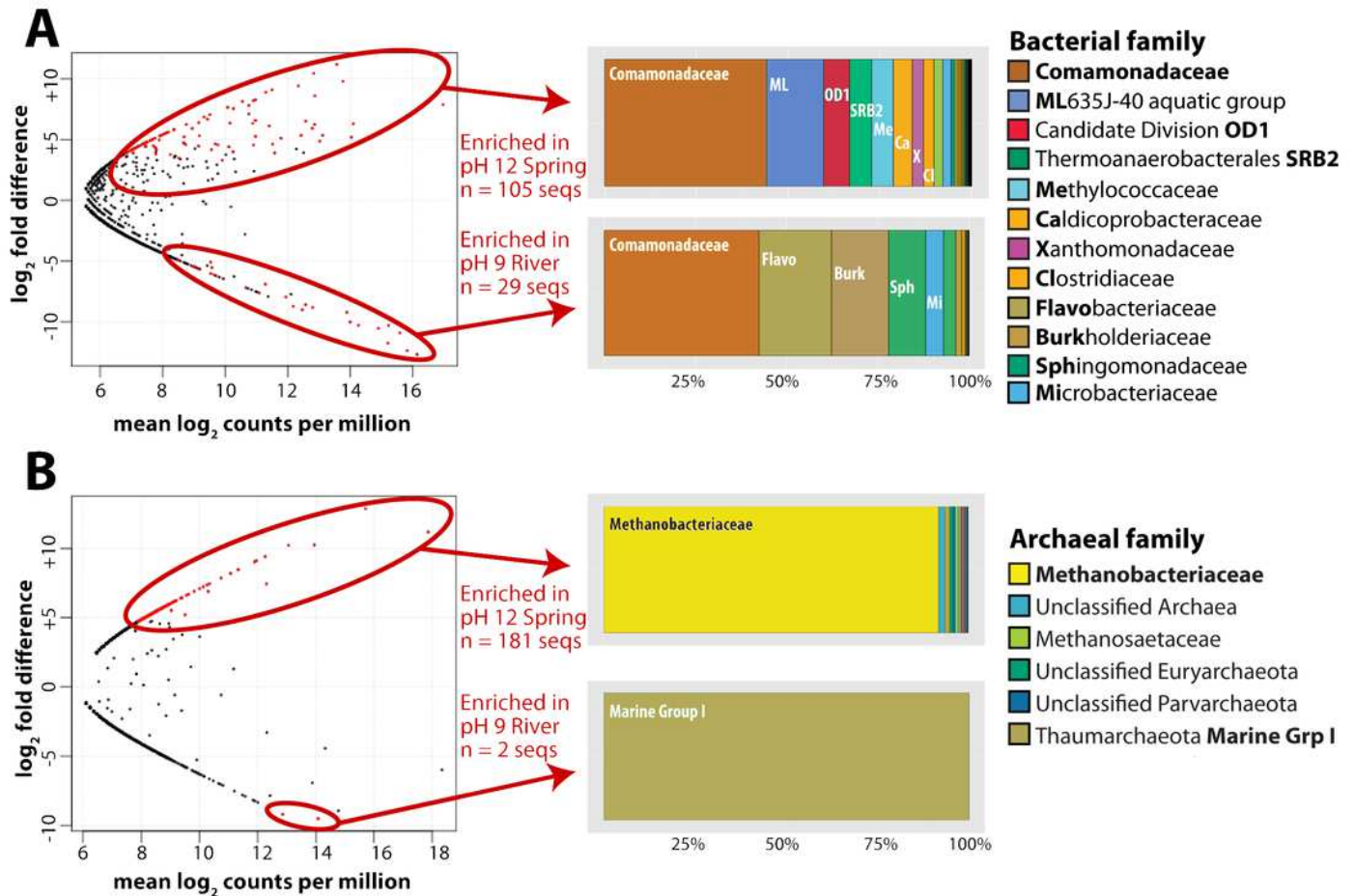
# Figure 3

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).
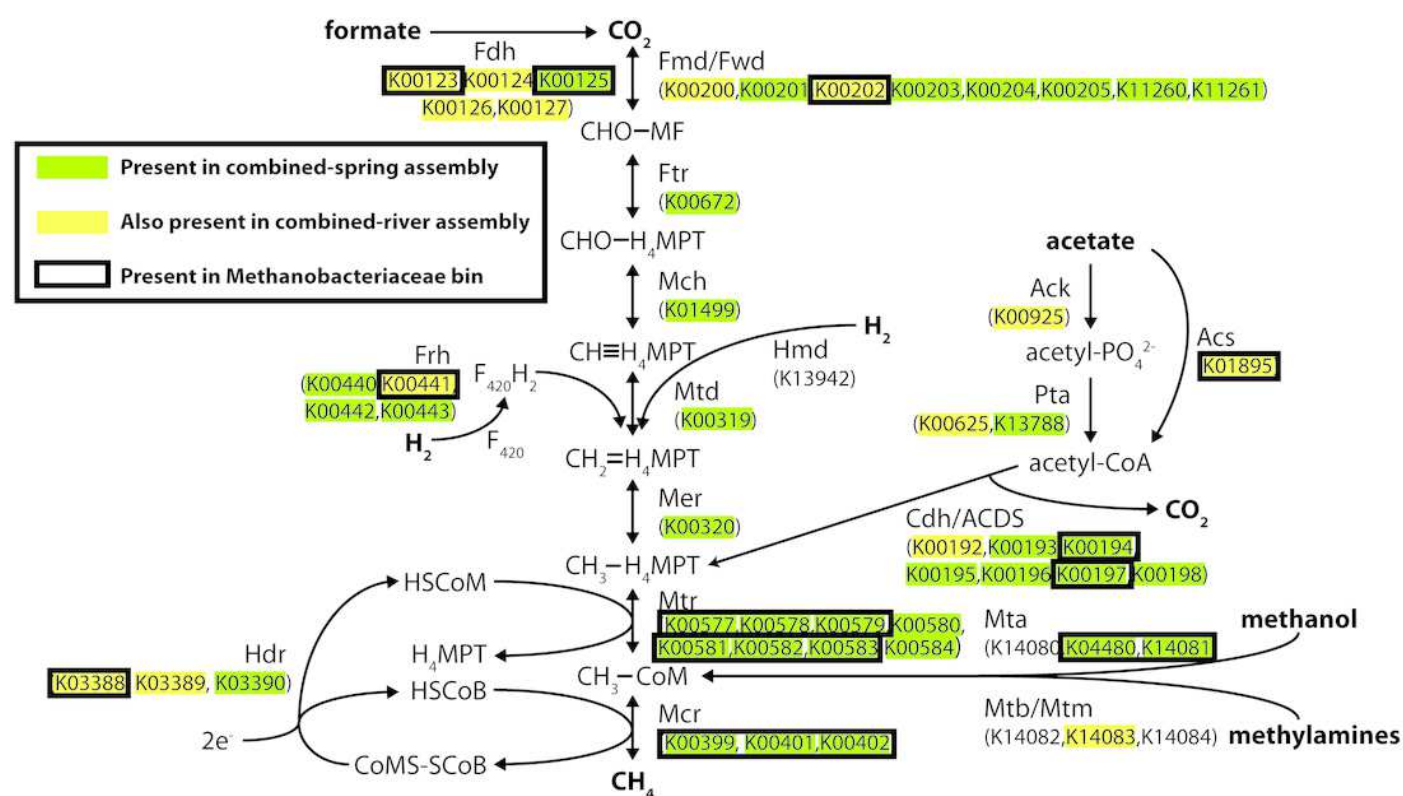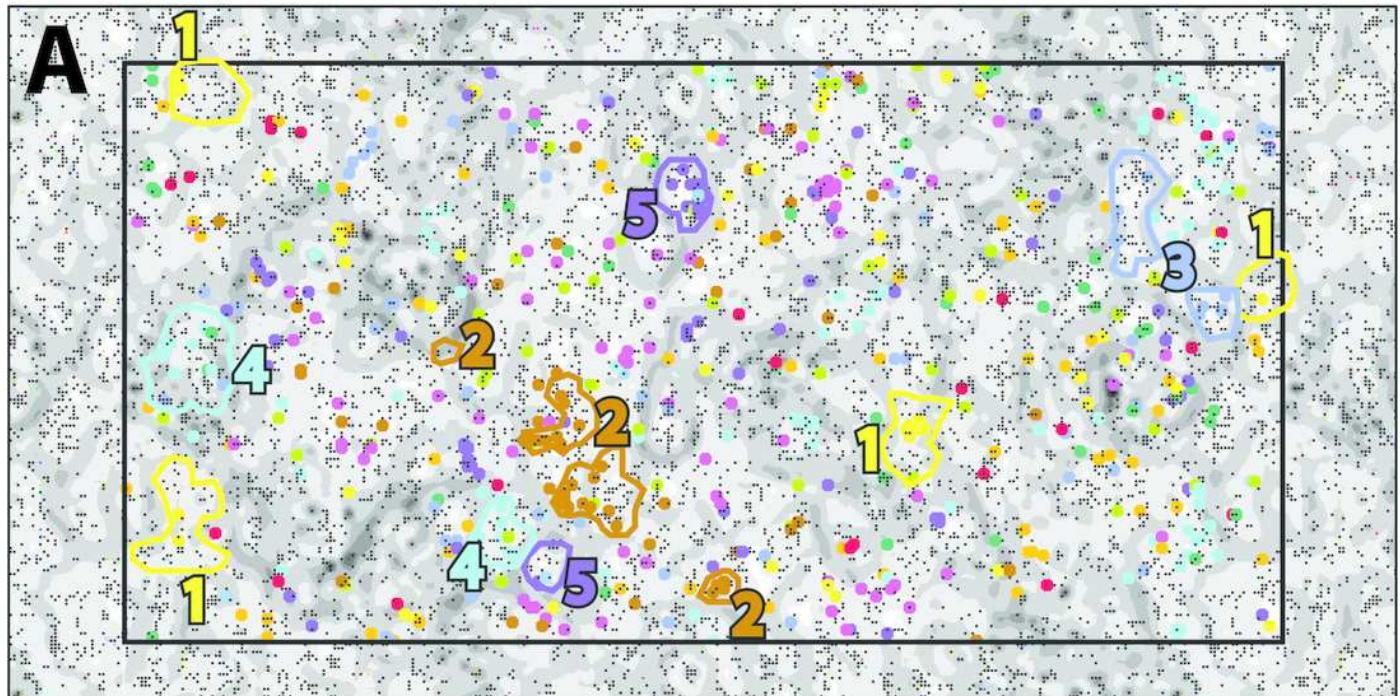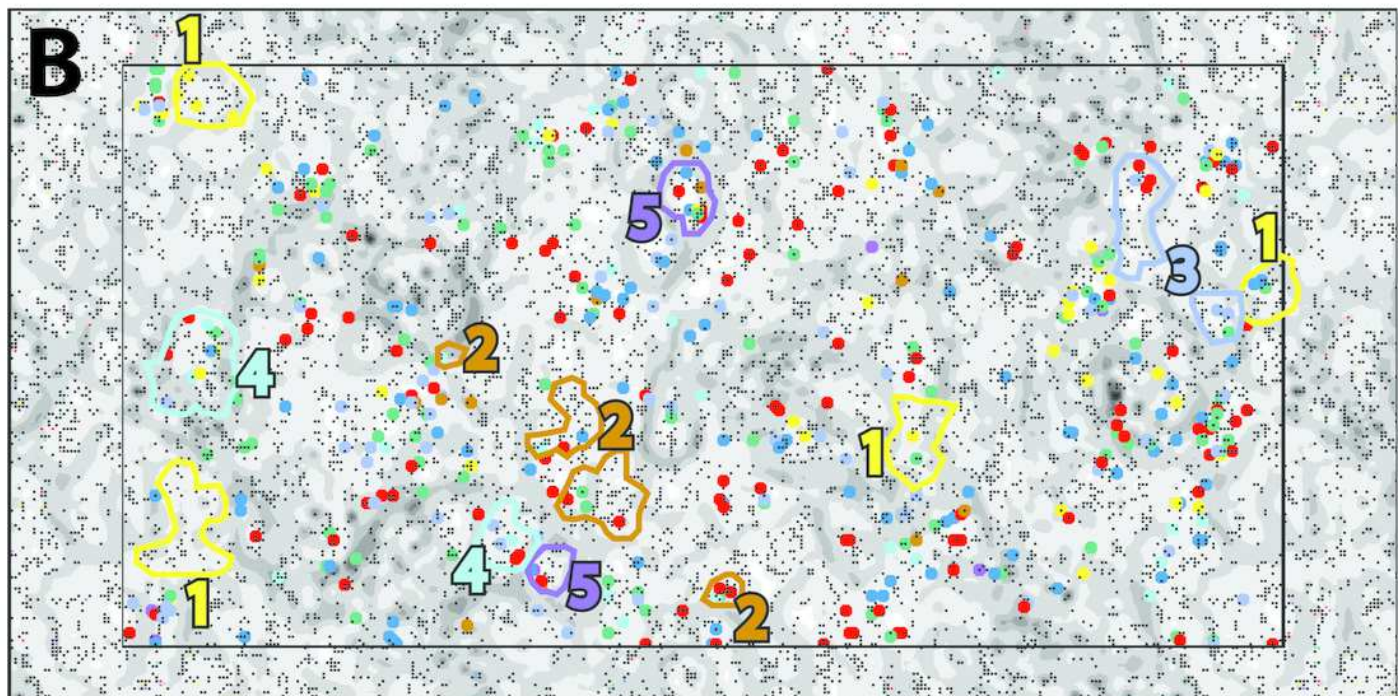
# Figure 4

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**

Legend:
- Methanobacteriaceae (yellow)
- Other methanogens (green)
- Comamonadaceae (brown)
- Bacteroidetes (light blue)
- Methylococcaceae (cyan)
- Desulfovibrionaceae (purple)
- Parcubacteria (pink)
- Thermoanaerobacterales (green)
- Clostridiales (orange)
- Xanthomonadaceae (magenta)

**B**

Legend:
- Methanogenesis (yellow)
- Acetate metabolism (blue)
- NiFe hydrogenase (brown)
- FeFe hydrogenase (light blue)
- Methylotrophy (cyan)
- Dissimilatory sulfate reduction (purple)
- Pyruvate fermentation (pink)
- Formate metabolism (green)