Microbial Biogeochemistry of Boiling Springs Lake: a Physically Dynamic, Oligotrophic, Low-pH Geothermal Ecosystem

Patricia Lynne Siering
Humboldt State University

Gordon V. Wolfe
The California State University

Mark S. Wilson
Humboldt State University

A. N. Yip
Humboldt State University

Cynthia M. Carey
Humboldt State University

See next page for additional authors

Let us know how access to this document benefits you.

Follow this and additional works at: https://pdxscholar.library.pdx.edu/bio_fac

Part of the Biogeochemistry Commons

Citation Details

This Article is brought to you for free and open access. It has been accepted for inclusion in Biology Faculty Publications and Presentations by an authorized administrator of PDXScholar. For more information, please contact pdxscholar@pdx.edu.
Authors

This article is available at PDXScholar: https://pdxscholar.library.pdx.edu/bio_fac/101
Microbial biogeochemistry of Boiling Springs Lake: a physically dynamic, oligotrophic, low-pH geothermal ecosystem

P. L. SIERING,1 G. V. WOLFE,2 M. S. WILSON,1 A. N. YIP,1 C. M. CAREY,1 C. D. WARDMAN,1 R. S. SHAPIRO,3 K. M. STEDMAN,4 J. KYLE,4 T. YUAN,5 J. D. VAN NOSTRAND,5 Z. HE5 AND J. ZHOU5,6

1Department of Biological Sciences, Humboldt State University, Arcata, CA, USA
2Department of Biological Sciences, California State University, Chico, Chico, CA, USA
3Department of Geosciences & Environmental Sciences, California State University, Chico, Chico, CA, USA
4Department of Biology and Center for Life in Extreme Environments, Portland State University, Portland, OR, USA
5Institute for Environmental Genomics and Department of Microbiology and Plant Biology, University of Oklahoma, Norman, OK, USA
6Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

ABSTRACT

Boiling Springs Lake (BSL) in Lassen Volcanic National Park, California, is North America’s largest hot spring, but little is known about the physical, chemical, and biological features of the system. Using a remotely operated vessel, we characterized the bathymetry and near-surface temperatures at sub-meter resolution. The majority of the 1.2 ha, pH 2.2 lake is 10 m deep and 50–52 °C, but temperatures reach 93 °C locally. We extracted DNA from water and sediments collected from warm (52 °C) and hot (73–83 °C) sites separated by 180 m. Gene clone libraries and functional gene microarray (GeoChip 3.0) were used to investigate the BSL community, and uptake of radiolabeled carbon sources was used to assess the relative importance of heterotrophic vs. autotrophic production. Microbial assemblages are similar in both sites despite the strong temperature differential, supporting observations of a dynamic, convectively mixed system. Bacteria in the Actinobacteria and Aquificales phyla are abundant in the water column, and Archaea distantly related to known taxa are abundant in sediments. The functional potential appears similar across a 5-year time span, indicating a stable community with little inter-annual variation, despite the documented seasonal temperature cycle. BSL water-derived DNA contains genes for complete C, N, and S cycles, and low hybridization to probes for N and S oxidation suggests that reductive processes dominate. Many of the detected genes for these processes were from uncultivated bacteria, suggesting novel organisms are responsible for key ecosystem services. Selection imposed by low nutrients, low pH, and high temperature appear to result in low diversity and evenness of genes for key functions involved in C, N, and S cycling. Conversely, organic degradation genes appear to be functionally redundant, and the rapid assimilation of radiolabeled organic carbon into BSL cells suggests the importance of allochthonous C fueling heterotrophic production in the BSL C cycle.

Received 12 February 2013; accepted 12 April 2013

Corresponding author: P. L. Siering. Tel.: +1 707 826 3120; fax: +1 707 826 3201; e-mail: pls13@humboldt.edu

INTRODUCTION

Hot springs, acid mine drainages (AMDs), and other environments at the extreme of temperature and pH offer a unique vantage for understanding the limits of life on Earth and can be used to model early life ecosystems and potential extraterrestrial targets. To date, biological and geological inventories have been made for many of these sites, most notably in locations such as Yellowstone National Park and the former Iron Mountain mine. However, nearly all of the research focuses specifically on either the composition of the microbial ecosystem or the
environmental constraints (geology, chemistry, hydrology) without significant linkages between the geology and the biology. Herein, we present an interdisciplinary, multi-year study that quantifies the environmental characteristics of Boiling Springs Lake (BSL) and links its unique characteristics (e.g. high temperature, very low pH, high turbidity, low metals, and low nutrients) to a multifaceted study of the microbial components.

Boiling Springs Lake, located at 1805 m elevation in the wilderness of California’s Lassen Volcanic National Park (LVNP; Fig. 1A,B), has an average summer temperature of c. 52 ± 3 °C and an average pH value of 2.0 ± 0.4 (Wilson et al., 2008). BSL is North America’s largest hot spring (http://bslmo.research.pdx.edu/), and unlike many terrestrial hydrothermal features which are ephemeral, BSL is known to pre-date 1864 (Whitney, 1865). BSL is approximately one-third the size, but significantly more acidic and older than New Zealand’s Frying Pan Lake, the world’s largest hot spring [pH 3.5, 50 °C, formed in 1918 (Lloyd & Keam, 1965)]. When LVNP’s thermal features were extensively studied by USGS geologists during the 1970s to mid-1990s (Friedman & Frank, 1978; Sorey & Ingebritsen, 1984; Ingebritsen & Sorey, 1985; Sorey & Colvard, 1994), BSL was determined to be the largest and most concentrated single heat source in the Lassen geothermal system (Sorey & Colvard, 1994). The hydrothermal system supplying BSL has recently been re-evaluated by Janik & McLaren (2010). In the revised model, the region around BSL is fed by a unique circulation system distinct from the majority of the park. Meteoric waters percolate to 5 km depth along the arcuate faults where they are heated in the reservoir to 220–240 °C. As the waters boil, they become gas-depleted before reaching the surface. BSL appears to be formed by thermal vents along a NW–SE trending transtensional fault that lies within a fault zone that defines the northern margin of the basalt and basaltic andesites of Sifford Mountain (Clynne & Muffler, 2010).

Geochemically, BSL occupies a unique niche among acidic and/or geothermal environments, as it is very low in organic C (oligotrophic), as well as inorganic salts and metals (Siering et al., 2006). Analysis of the emitted gasses from features within the BSL thermal area in 1993 and 1998 indicated that the primary constituent is CO₂ (88.3%), followed by N₂ (7.2%), H₂S (3.7%), H₂ (0.5%), Ar (0.14%), and CH₄ (0.08%), with O₂ and NH₃ at near negligible amounts (Janik & McLaren, 2010). BSL is geologically and chemically distinct from caldera-type volcanic lakes typified by the Uzon Caldera (Kamchatka, Russia) and features within the Rotoura geothermal field in New Zealand (McColl & Forsyth, 1973; Mroczek et al., 2004) and from high chloride, low sulfate features characteristic of the liquid-dominated geothermal systems in Yellowstone National Park (YNP; Thompson, 1985; McCleskey et al., 2004; Inskeep & McDermott, 2005; Janik & McLaren, 2010). Compared to well-studied AMD sites

![Fig. 1 Map of study site. (A) Lassen Volcanic National Park and its approximate location in Northern California. Black circles are geothermal sites (1, Sulfur Works; 2, Bumpass Hell; 3, Devil’s Kitchen; 4, Boiling Springs Lake). Large gray circle denotes former Brokeoff Mountain caldera. Adapted from National Park Service web site. (B) Aerial photo of lake (Google Earth, 7/10). (C): Boiling Springs Lake 30-cm resolution GPS map. Water sampling sites A and D are shown, along with major near-shore geothermal features (black diamonds).](image-url)
such as the Rio Tinto (Justice et al., 2012) or Iron Mountain Mine (Edwards et al., 1999), BSL is a much more dilute sulfuric acid solution, and contains extremely low burdens of heavy metals, as well as low nutrients (Siering et al., 2006). However, it is much warmer than coal mining lakes of more similar water chemistry (Travisany et al., 2012) restricting the eukaryotic community to entirely microbial taxa, and BSL’s basin depth and structure are obscured by suspended sediments that limit light penetration. Additionally, as an ‘aquatic island’ in an alpine forest, BSL receives allochthonous organic material from the surrounding vegetation (conifer needles and woody debris).

Previous 16S rRNA gene analyses of BSL sediments revealed that the majority of detected phylotypes (PTs) represented novel genera of Archaea and Bacteria that shared <85% identity with cultured isolates (Wilson et al., 2008). Eukaryotic diversity is limited, with Cyanidiales, ascomycete fungi, and heterolobose anaerobes dominating the planktonic and benthic communities (Brown & Wolfe, 2006; Siering et al., 2006; Reeder, 2011). Virus-like particles are at low densities (G. Diemer and J. Kyle, pers. comm.), consistent with what has been observed in other high-temperature acidic waters (Ortmann et al., 2006; Prangishvili et al., 2006). We also used terminal-restriction fragment length polymorphism (T-RFLP) analysis to investigate the microbial diversity of BSL over a 7-year time span (Wilson et al., 2008), and we found a stable community of micro-organisms in water and sediment samples collected from the northern end of the lake.

In this paper, we characterize the previously unknown bathymetry of BSL, and detail dynamic changes over sub-hour time scales due to thermal mixing. Geothermal lake bathymetry has been mapped for several sites, including Lake Yellowstone (Morgan et al., 1977, 2003; Otis et al., 1977; Kaplinski, 1991), and sites in New Zealand (McColl & Forsyth, 1973; Fish, 1979; Pickrill, 1993), Indonesia (Sriwana et al., 2000; Takano et al., 2004), and Japan (Miyabuchi & Terada, 2009). However, only a few studies have attempted to link lake structure to microbial populations and dynamics (Clingenpeel et al., 2011). Additionally, to complement geochemical and clone library data, and to gain insight into possible nutrient and energy cycling in BSL, we performed functional gene microarray analysis using GeoCHIP 3.0 (He et al., 2010). This and earlier versions of the GeoCHIP have been used to investigate community structure and functional potential in various ecosystems and extreme habitats, including hydrothermal vents (Wang et al., 2009), hypersaline systems (Parnell et al., 2010), permafrost (Hinsa-Leasure et al., 2010), and most notably, AMD sites (Xie et al., 2011). Additionally, it has been employed in numerous beta-diversity studies comparing anthropogenically or geogenically impacted sites (Xiong et al., 2010; Liang et al., 2011; Van Nostrand et al., 2011; Beazley et al., 2012; Ding et al., 2012; Lu et al., 2012; Reith et al., 2012). We use the data collected in this study along with previous 16S rRNA gene analysis and cultivation studies to assemble a qualitative model of nutrient cycling in BSL.

MATERIALS AND METHODS

Physical characterization

GPS mapping

High-resolution (<0.6 m at 95% confidence) differential GPS mapping (Fig. 1C) was used to create a database of specific shoreline features using an SX Blue II (Geneq, Inc., Montreal, Quebec, Canada) LI GPS receiver integrating the public Wide Area Augmentation System with proprietary COAST algorithms to smooth the C/A carrier phase. Data integrity was maintained at a maximum dilution of precision (DOP) of 2.0 (average 1.1) by limiting mapping to optimal satellite configurations. Data locations were held for a minimum of 10 epochs to increase resolution. Locations, along with collection parameters (DOP, epochs, resolution), and metadata (geographic feature, pH, temperature), were collected using SOLO Field (Tripod Data Systems, Corvallis, OR, USA) on a TDS Nomad field computer. Data were analyzed and plotted using Arc GIS (ESRI) and MapWindow GIS (open source accessed from www.mapwindow.org) software.

Bathymetry

The study employed a student-designed radio-controlled vehicle (ROV) to characterize bathymetry. The ROV included a logging sonar/GPS system (Lowrance LMS-480M; Lowrance Electronics, Tulsa, OK, USA), and a 4-channel data logger (Madgetech, Warner, NH, USA) for water temperature. The ROV was outfitted with a sub-meter resolution SX Blue II GPS receiver (Geneq, Inc.) that transmitted via Bluetooth to a data collector on shore to verify position on the base map during operation; both position systems gave excellent agreement (Fig. S2). Sonar depth data were analyzed by ROV X-Y position and synchronized with temperature data by time stamps. Depth soundings were inspected visually and manually edited to remove gas bubble artifacts (see Data S1 for details), and contour plots were generated with Origin 8 (OriginLab Corp., Northampton, MA, USA) using Renka–Cline gridding, suitable for randomly spaced X-Y data.

Temperature monitoring and time-lapse photography of hydrothermal convection

During winters (October–May) 2008–2009, 2009–2010, and 2011–2012, we deployed a tethered PVC buoy with a 4-channel data logger (Madgetech Inc.) to measure water temperatures at 0–8 m depths every 30 min. In fall 2010, time-lapse photography was used to document lake con-
vection over hourly scales, using a Canon EOS Rebel T1i camera and GTTimelapse software (Granite Bay Software, Granite Bay, CA, USA), set for 15-min photo intervals. For observations of winter months, a DigiSnap 2000 controller (Harbortronics, Inc., Fort Collins, CO, USA) was powered by a 12V RV battery, running a Nikon Coolpix 900 (Nikon Inc, Melville, NY, USA), and taking three photos per day for 6 months.

Geochemical analysis

Water samples were filtered through a 0.2 μm membrane (Durapore; EMD Millipore, Billerica, MA, USA), except for dissolved organic carbon (DOC), which was filtered through two muffled GF/F filters (Whatman, Piscataway, NJ, USA), and total organic carbon (TOC), which was not pre-filtered. Both DOC and TOC were stored in glass amber bottles containing sulfuric acid. All other samples were stored in polypropylene bottles containing nitric acid (for cation determination) or sulfuric acid (total P). Samples collected for anion and alkalinity determination were not chemically pre-treated. Samples from 2008 were collected at sites A and D on July 5, 2008, and from site A on September 14, 2008, and stored at 4 °C until analysis. DOC and TOC were determined by combustion (Shimadzu Scientific Instruments, Columbia, MD, USA). Ions were analyzed with inductively coupled plasma mass spectrometry (ICP-MS) at the TEAL laboratory at Portland State University (http://web.pdx.edu/~rperkins/Laboratory.htm). Samples from 2011 were collected from sites A and D (Fig. 1) on July 23, 2011, stored at 4 °C and analyzed by Test America, Inc. (Beaverton, OR, USA) using EPA standard methods. Cations were analyzed with ICP-MS or atomic emission spectroscopy (ICP-AES) using EPA 200.8 and 200.7, respectively. Total P and nitrate/nitrite were analyzed using colorimetry (EPA SM 4500P-F and 353.2, respectively), sulfate was analyzed using ion chromatography (EPA 300.0), alkalinity by titration (EPA SM 2320B), DOC by UV-promoted persulfate oxidation (EPA 415.2), and TOC by persulfate-UV or heated persulfate oxidation (EPA SM 5310C).

Biological characterization

Cell enumeration

Samples for bacterial cell counts were fixed on site in 3.7% formalin for acridine orange direct cell counts that were performed as previously described (Siering & Ghiors, 1997).

DNA sample collection and processing

Near-shore water samples from sites A and D, and sediment samples at site A water-sediment interface were collected in July 2004 and 2009 from BSL as previously described (Wilson et al., 2008). Within each site, three 5 L water samples were taken c. 1 m apart and independent nucleic acid extractions were obtained from three subsamples of each sample. Cells were collected aseptically by filtering 500 mL water samples through 0.2 μm Durapore Millipore filters for site A samples, and 50 mL for site D samples. The volume of water put through each filter was limited by the amount of suspended solids in the sample. Filters were immediately frozen in liquid nitrogen and stored at ~80 °C, and nucleic acids were extracted from collected cells and purified by previously reported methods (Siering et al., 2006) except with the addition of RNAs prior to phenol-chloroform extraction and recovery. Pooled DNA extracts were prepared for both sites by combining DNA from three separate subsamples. The same A and D pools used for clone library preparation of site A and D samples were also used for the 2004 site A GEOCHIP analysis and functional gene PCR screens. Site A sediments were aseptically subsampled into 0.75 mL volumes in individual tubes. Glycerol was added to a final concentration of 15% (v/v) and mixed prior to immersing sediment tubes in liquid nitrogen. Nucleic acids were extracted from site A sediments by a previously described bead-beating method (Bond et al., 2000). All purified nucleic acids were quantified on an ethidium bromide-stained agarose gels by comparison with known molecular weight standards, and concentrations were also confirmed by use of NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies Inc., Wilmington, DE, USA).

16S rRNA gene clone libraries

Three 16S rRNA gene clone libraries were constructed separately for site A and for site D pooled planktonic DNA extracts using PCR primers U341F (Hansen et al., 1998) with U1406R (Reysenbach & Pace, 1995), or universal primer U515F with either U1406R or the prokaryotic-specific primer P1525R (Reysenbach & Pace, 1995) as previously described (Wilson et al., 2008). Additionally, a 16S rRNA gene clone library was prepared from sediment-derived DNA using primers U341 and U1406. The site D water clone library from primers U341F and U1406R was made and sequenced as previously described (Siering et al., 2006): 163 of these clones were sequenced on one strand, and 34 clones [representatives of each phytype (PT) and singletons] were sequenced in both directions. All other PCR products were cloned and sequenced by the Washington University Genome Sequencing Center (St. Louis, MO). Sequence data were processed using SEQUENCER, Version 4.0 (Gene Codes Corporation, Ann Arbor, MI, USA). The MOTHUR v.1.9.0 platform (Schloss et al., 2009) was used for PT assignment using a 97% ID cutoff value. Sequences were checked for potential chimeras using the Chimera Check program from the RIBOSOMAL DATABASE PROJECT v.9 (Cole et al., 2009) and Bellerophon (Huber et al., 2004). By this analysis, c. 5–10% of the sequences were chimeric and were removed from further analysis. The most conserved sequences within each PT and the sin-
gle clones that did not initially group with other clones during cluster analysis were chosen as reference sequences and submitted to GenBank. Clones from planktonic libraries constructed with the U341F and U1406R primers were submitted under the accession numbers FJ358714–FJ358730 (site A) and EF558669–EF558700 (site D), and 16S rRNA gene reference clones from the remaining libraries were submitted under accession numbers JX428588–JX428655. Species effort analyses were performed using ANALYTIC RAREFACTION 1.3 (Holland, 2003).

Functional gene clone libraries
We used several sets of PCR primers to screen for genes in four of the known carbon fixation pathways, including the Calvin–Benson–Bassham and reductive TCA cycles, and the reductive acetyl CoA and 3-hydroxypropionate pathways (Table S1). In addition, we screened for a number of N-cycle processes, including N₂ fixation, denitrification, and nitrification (Table S2). PCR reactions were performed according to methods specified in primer references (Tables S1 and S2). When amplifications were unsuccessful, additional attempts were made using touchdown PCR and gradient PCR procedures. Three independent PCR reactions were run on the site A water pooled extracts from 2004 (described above), and where successful, amplicons were pooled in mass-normalized amounts for construction of clone libraries. Sequences were aligned using the San Diego Supercomputer Biology Workbench (http://workbench.sdsc.edu; Subramanian, 1998) using CLUSTALW. A distance matrix was constructed using CLUSTAL DIST and DNADIST with the Maximum Likelihood model of substitution using empirical base frequencies. The most conserved sequences in each identified PT were used for BLAST and blastx analysis (Altschul et al., 1997) and submitted in GenBank under the accession numbers JX428656–JX428665.

GEOChip 3.0 analysis
GEOChip 3.0 employs c. 28 000 50-mer oligonucleotides selected for their wide representation of functional genes and pathways. The oligonucleotides are designed to have low (>0.1%) false-positive and false-negative hybridization signals, but, as with any microarray, the genes that can be detected are limited by the oligonucleotides chosen. Moreover, hybridization signals depend on both sequence identity to the oligonucleotide probes and gene abundance. In order to limit the effect of false positives and negatives, our analysis depends on hybridization to multiple DNA samples from 2 years. DNA samples used for GEOChip analysis had 260:230 absorbance ratios of 1.90–1.95, with concentrations of 30 and 39 ng μL⁻¹ for 2004 and 2009, respectively. As we were unable to extract DNA of sufficient A260:230 ratios for GEOChip analysis from site A sediments, or site D samples, only Site A water DNA was subject to analysis.

Genomic DNA samples for each year were subjected to whole genome amplification (WGA) (in triplicate) using the Templiphi kit (GE Healthcare) as previously described (Wu et al., 2006). Samples of unamplified (one sample per year), and whole genome amplified (three replicates per year) were subjected to Cy-5 labeling and hybridization using GEOChip 3.0 (He et al., 2010; Trivedi et al., 2012). This provided four hybridizations for each of the 2004 and 2009 samples (eight hybridizations total). Raw hybridization data were pre-processed using a data analysis pipeline to remove poor quality hybridization spots, normalize intensities, and remove replicates whose signal intensity was >2 standard deviations from the mean (He & Zhou, 2008; Trivedi et al., 2012). Only gene sequences that were detected in at least two of the four hybridizations for both years, or at least three of the four hybridizations for a single year, hereby called detected genes (or gene sequences), were analyzed. Detected gene sequences were analyzed for phylogenetic affiliation, gene category, and specific processes within individual gene categories as defined by GEOChip 3.0. Signal intensities (SI) were summed for all of the detected gene sequences and used to determine normalized signal intensities (NSI). NSI for phylogenetic affiliation and gene category were determined by dividing the total SI for a particular lineage or gene category by the total SI for all detected gene sequences. NSI for individual processes within each category were determined by dividing the total SI for that process by the total SI for that category (e.g. total SI for all detected nif genes was divided by total SI for nitrogen cycling to determine NSI for nif genes). This was carried out separately for the 2004 and 2009 hybridizations.

Primary and bacterial secondary production
Rates were estimated by radiotracer methodology under simulated in situ conditions in the laboratory. Single samples were collected in July 2010, November 2010, and January 2011. Water samples were brought to the laboratory within 5–8 h, and stored at 4°C overnight to try to maintain in situ conditions and avoid disturbing microbial activities. For primary production, samples were incubated in 20 mL scintillation vials with NaH¹⁴CO₂ (Moravek Biochemicals, Brea, CA, USA) under simulated full-light conditions, or in foil-wrapped vials to simulate dark conditions. Vials were filled to avoid outgassing of CO₂ under acidic conditions. Killed controls were pre-treated with 0.5% glutaraldehyde. Samples were incubated for 2–6 h and then filtered onto 0.2 μm-pore size polycarbonate filters (Poretics Corporation, Livermore, CA, USA), rinsed, and counted in ScintiVerse cocktail (Thermo-Fisher, Waltham, MA, USA). Heterotrophic potential was measured by incorporation of ¹⁴C[U]-acetate or glucose (Moravek Biochemicals, Brea, CA, USA). ‘Bacterial’ production (likely all prokaryotes) was measured by ³H-thymi-
dine incorporation (Gyure et al., 1987). Following incubation, samples were fixed with glutaraldehyde (0.5% v/v final) and ice-cold trichloroacetic acid (TCA; 5% v/v final). After 30 min incubation, these were filtered as with other samples, and filters were rinsed with ice-cold TCA and counted as for $^{14}$C. All additions were used at 0.1–1 µCi per sample, and label did not exceed 8 ppm. Typical counting efficiencies were 62% ($^3$H) and 95% ($^{14}$C) as determined by the internal standards method. In some instances, we pre-incubated samples for 2–12 h with 10–1000 µM of various C sources to investigate the impact on bacterial production (see Results).

RESULTS

Site characterization

The BSL thermal area includes prominent areas of hot, dry soil that is rich in iron oxides, especially on the north and east sides of the lake (Fig. 1B), where we have observed temperatures of up to 96 °C at 1.5 m depth. The lake basin lies along an unnamed fault (Clyne & Muffler, 2010). Seasonal (October–May) water inputs include snow, snowmelt, and runoff, and a stream that flows into the south end. Winter lake levels are limited by outflow through a channel near site A, and winter water elevations can be inferred along the shoreline from white areas where acid has leached out iron and left behind silicates (Fig. 1B). These weathered soils also deposit into the lake acid has leached out iron and left behind silicates (Fig. 1B). These weathered soils also deposit into the lake through a channel near site A, and winter water elevations measured in this study were 4–10 °C cooler (Fig. 3). Two temperature phenomena are evident: a seasonal temperature depression from 52 °C to about 48 °C, interspersed by rapid temperature drops that occur over day–week time intervals. The drops in temperature correlate with episodic snow input from storms, as deduced from weather station data located 15 km distant and at Chester, CA (1380 m elevation). Daily time-lapse photography revealed that snow could be seen at the lake’s edge following storms, but melt-back was evident within hours (not shown).

Geochemistry and cell counts

Geochemical analyses of BSL water collected from sites A and D from the summer and fall of 2008, and summer of 2011 were performed to provide data complementary to that collected in 2001 and 2004 (Snyder, 2005; Siering et al., 2006) and to extend the period of collection to 10 years. These measurements indicate fairly high sulfate concentrations (1060 ± 147 ppm), and relatively low Cl$^-$ (1.3 ± 0.3 ppm; Table 1). Iron and Al are present at 30.3 ± 11.6 ppm, and 22.9 ± 10.7 ppm, respectively, while all other metals were found at low to extremely low concentrations (<0.1 ppm). NO$_3^-$ and PO$_4^{3-}$ were very low, c. 0.05 and c. 0.02 ppm, respectively. Measurements of TOC in 2008 were c. 1 ppm, with DOC near or below the detection limit (c. 0.3 ppm). Measurements of TOC and DOC in 2011 were below reported detection limits (1 ppm).

Water samples from sites A and D (Fig. 1C) had prokaryote cell concentrations of 2.49 × 10$^7$ (SD 2.72 × 10$^5$) and 5.66 × 10$^6$ (SD 7.22 × 10$^6$) cells mL$^{-1}$, respectively, comparable to mesothermic, neutral pH, aquatic environments, and with counts from site A waters and sediments in previous years (Wilson et al., 2008).

Detected genes on GeoChip 3.0

We detected 275 GeoChip 3.0 gene sequences in all eight hybridizations of BSL water-derived DNA (see Materials...
Phylogenetic affiliation of detected genes/organisms

Archaebacteria and Eukarya accounted for only minor fractions (4% and 5%, respectively) of the normalized signal intensity (NSI) of genes detected on the GEOCHIP 3.0, with 90% of all detected genes from Bacteria. Proteobacteria accounted for half of the bacterial NSI, while Firmicutes accounted for c. 13%, and Actinobacteria <3%. Of 1164 gyrB (gyrB) probes used as a phylogenetic marker on the GEOCHIP 3.0 (He et al., 2010), BSL samples hybridized to 45, representing c. 5% of NSI (Fig. 4A). Almost half (22) was from Proteobacteria classes, with the highest gyrB signal (9.8% of all gyrB) from Thiocapsa rosea, a mesophilic, neutrophilic, anoxygenic phototroph capable of anoxic photolithotrophic growth on sulfide, thiosulfate and sulfur. Other gyrB gene probes included Actinobacteria (1), Acidobacteria (1), Firmicutes (6), Cyanobacteria (1), Bacteroidetes (1), Planctomycetes (1), and Tenericutes (4). Six were derived from uncultivated bacteria from AMD samples, with one (GenBank ABG78306) accounting for c.
6.5% NSI for all gyrB genes detected. Most of the hybridization-positive gyrB probes, including the one for *Thiocapsa*, correspond to organisms that have not been isolated or detected in 16S rRNA gene clone libraries made from BSL (Siering et al., 2006; Wilson et al., 2008), but the sole archaeal gyrB sequence detected, the euryarchaeote *Thermoplasma acidophilum*, was previously detected in BSL sediments (Wilson et al., 2008) and also recently isolated from BSL sediment/water slurry (P. L. Siering and M. S. Wilson, unpubl. data).

Rarefaction analysis indicated that the PT accumulation curves from the 16S rRNA gene clone libraries (Table 2) had begun to flatten out (data not shown), which suggests that the most abundant PTs in the clone libraries were identified. Bacterial PTs were more abundant than were archaeal PTs, but in contrast with the GeoChip results, *Proteobacteria* did not dominate the clone libraries. The
Hybridization to carbon-cycling gene probes accounted for 16.5% of the NSI for all detected gene sequences (Table 4, Fig. 4). Both autotrophic and heterotrophic processes were well represented.

### Autotrophy

Hybridization to carbon fixation gene probes corresponded to 19.4% of the NSI for all detected carbon-cycling genes, with the hydroxypropionate pathway representing 40% of this signal (Table 4). Approximately, 41% of the RuBisCo probe signal intensity was due to hybridization to a single probe designed from an uncultivated bacterium in pyrite mine acid drainage (Kamimura et al., 2010). The protein sequence from which this probe was designed has 100% amino acid identity (over 158 aa) to the encoded RuBisCo from Acidithiobacillus ferrooxidans, and it is likely that this detected gene is very closely related to our ST-2 from the 595F/1387R cbbL clone library (Table 3).

C1 metabolism was also evident; we detected hybridization to 10 CO dehydrogenase probes (for carboxydrotrophs; Table 4) and 60% of the NSI for this gene was contributed by probes whose sequences were derived from uncultured *Archaebacteria and Bacteria*. We also detected hybridization to a soluble methane monoxygenase gene (*mnaX*) probe from *Novosphingobium* spp. JSP614, and the particular methane monoxygenase gene (*mnaA*) probes from an uncultured bacterium in a soda lake, a gas hydrate region in the Gulf of Mexico, and from a hydrothermal vent.
study. The single methane reduction gene probe on the chip (mcrA, encoding methane coenzyme A reductase) hybridized only at low SI (<0.1% NSI).

We successfully amplified two carbon fixation genes from BSL DNA extractions (Table 3). We detected three cbbL alleles, and we also detected a single allele for the gene encoding FTHFS; only one of the cbbL alleles was identified from the GeoChip results (see Autotrophy, above). Both of the cbbL primer sets (targeting the large subunit of Rubisco I) amplified a sequence that had 86–88% nucleotide (and 97% amino acid) identity to cbbL from A. ferrooxidans. The $595F/1387R$ primer set (also targeting the large subunit of Rubisco I) amplified an additional set of sequences that was c. 82% identical to the first set of clones, and shared 86–88% nucleotide (98% amino acid) identity to the cbbL gene from *Acidithiobacillus caldus*. The primer set for FTHFS (formyltetrahydrofolate synthetase gene used in reductive Acetyl CoA pathway) amplified an c. 1100 nt region from clones that were related (64% nucleotide used in reductive Acetyl CoA pathway) amplified an sequence that had 86–88% nucleotide (and 97% amino acid) identity to FTHFS from *A. ferrooxidans*. The $595F/1387R$ primer set (also targeting the large subunit of Rubisco I) amplified an additional set of sequences that was c. 82% identical to the first set of clones, and shared 86–88% nucleotide (98% amino acid) identity to the cbbL gene from *Acidithiobacillus caldus*. The primer set for FTHFS (formyltetrahydrofolate synthetase gene used in reductive Acetyl CoA pathway) amplified an c. 1100 nt region from clones that were related (64% nucleotide identity) to the sequence of FTHFS from *Thermoanaerobacterium brockii*, and BLASTX analyses showed that the inferred amino acid sequence was 50% identical across 326 amino acids to the FTHFS protein sequences from several *Thermoplasma* and *Thermoanaerobacterium* isolates. No amplicons were detected for other genes and primer sets employed.

**Heterotrophy**

Approximately, 65% of the NSI for all C-cycling processes corresponded to probes for heterotrophic C degradation (Table 4). Of particular interest were genes for breakdown of terrestrial plant polymers, likely a major C source to the BSL microbial community. These were observed for both prokaryotes, as well as fungi. Hybridization to starch degradation gene probes represented c. 16% of the NSI for C cycling, although only three of eight different probes were detected in BSL DNA (Table 4). Hybridization to the *arA* probe from *Thermoanaerobacter ethanolicus* had a particularly high signal intensity compared to the others, and represented 12–16% of the NSI for all detected starch degradation genes. Other species and strains of this genus were detected in sediment clone libraries (Table 2) and in a previous study (Wilson *et al.*, 2008). No gene sequences targeting peptin degradation were detected, but three genes for hemicellulose degradation were detected from five bacterial phyla, and an additional gene was detected from five species within the *Ascomycota* phylum (*Eukaryota*) accounting for 11.5% of NSI for all detected C-cycling genes (Table 4). Three of the four cellulose degradation genes assayed on the chip were detected from four bacterial phyla and one from the *Ascomycota*. Hybridization to a cellulase gene probe from *Thermococcus* and *Eukaryota* sp. were also detected. Other genes involved in lignin breakdown may be present but are possibly classified as ‘organicism remediation’, the largest fraction of probes on the entire GeoChip 3.0 (He *et al.*, 2010). The NSI for this category was high (Table 4), but like other gene categories, only a small fraction of all probes for this category hybridized to DNA from BSL water.

**Table 3** Functional C- and N-cycling genes detected by PCR screens. Nucleotide identity was determined using MegaBlast, and amino acid identity was determined using BLASTX (Altschul *et al.*, 1997). A total of 12 primer sets were used to screen for various genes involved in nitrogen-cycling processes, and thirteen primer sets were used for genes involved in five known pathways for carbon fixation (Tables S1 and S2). Only positive results are shown.

| Primers used | Gene/activity detected (amplon size) | Phylotype (no. of sequences in phylotype) | Nucleotide identity (% coverage) [accession no.] | Amino acid identity (number of a.a.) [accession no.]
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>nifH3/nifH4 (outside)</td>
<td>nifH/N₂ fixation (363 bp)</td>
<td>ST-1 (n = 45)</td>
<td>98–99% (98%) <em>Methylacidiphilum fumarophilum</em> SoL [GU299762.1]</td>
<td>100% (119 aa) <em>M. fumarophilum</em> SoL [GU299762.1]</td>
</tr>
<tr>
<td>nifH1/nifH2 (inside) (Izquierdo &amp; Nüsslein, 2006)</td>
<td>(792 bp)</td>
<td>ST-2 (n = 19)</td>
<td>82–84% (98–99%) <em>Anabaena variabilis</em> ATCC29413 [CP000117.1]</td>
<td>93% (120 aa) <em>Thermocrinis albus</em> [YP_003473186.1]</td>
</tr>
<tr>
<td>$595F/1387R$ (Eliaied &amp; Naganuma, 2001)</td>
<td>cbbL/CO₂ fixation (792 bp)</td>
<td>ST-1 (n = 8)</td>
<td>88–89% (100%) <em>Acidithiobacillus caldus</em> SM-1 [CP002573.1]</td>
<td>97% (271 aa) <em>A. caldus</em> SM-1 [CP002573.1]</td>
</tr>
<tr>
<td>l2F/v2R (Nanba, King &amp; Dunfield, 2004)</td>
<td>cbbL/CO₂ fixation (495 bp)</td>
<td>ST-2 (n = 6)</td>
<td>86% (100%) <em>A. ferrooxidans</em> DSM 10331 [CP001631.1]</td>
<td>94% (271 aa) <em>A. ferrooxidans</em> DSM 10331 [CP001631.1]</td>
</tr>
<tr>
<td>FTHFS (Lovell &amp; Leaphart, 2005)</td>
<td>Formyl tetrahydrofolic synthetase gene/CO₂ fixation (1100 bp)</td>
<td>ST-1 (n = 17)</td>
<td>65% (59%) <em>Thermoanaerobacter brockii</em> [CP002466.1]</td>
<td>52% (350 aa) <em>Caldicellulosiruptor obsidiansis</em> O847 [CP002164.1]</td>
</tr>
</tbody>
</table>
Nitrogen cycle

The previously reported low nitrogen (N) species measurements (c. 1.43 ppm ammonia, and <0.02 ppm nitrate; Siering et al., 2006), and the chemical data reported in this work (Table 1) correlate well with many detected N-cycle genes. Hybridization to probes for nitrogen-cycling genes account for c. 12% of the NSI of all detected genes (Table 4), and genes for scavenging N and for N2 fixation are highly represented in the top 25 detected gene sequences based on hybridization NSI (Table S3).

Table 4 Summary of GeoChip 3.0 data for C, N, and S cycles. Gene name is indicated when a single gene name exists; however, enzyme name is given when different gene names are used in different taxa. Normalized signal intensity (NSI) for each gene category (e.g. N cycle) was determined by dividing the total SI for all processes within a gene category by the total SI for all detected gene sequences on the chip. NSI for individual processes within each category was determined by dividing the total SI for that process by the total SI for that category (e.g. total SI for all detected nif genes was divided by total SI for nitrogen cycling to determine NSI for nif genes). Mean NSI reflects the mean for all eight hybridizations (four each per 2004 and 2009), and the standard error of the means was calculated based on the means for 2004 and 2009 hybridizations. Descriptions of detected alleles are provided in the text.

<table>
<thead>
<tr>
<th>Process</th>
<th>No. of detected (gene or enzyme) [no. of genes on chip]</th>
<th>No. of detected alleles</th>
<th>No. of probes on chip</th>
<th>Mean NSI (04 and 09)</th>
<th>SE of means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetogenesis</td>
<td>1 (formyltetrahydrofolate synthetase FTHFS) [1]</td>
<td>3</td>
<td>122</td>
<td>0.92</td>
<td>0.04</td>
</tr>
<tr>
<td>C deg. – cellulose</td>
<td>3 (cellobiase, endoglucanase, exoglucanase) [4]</td>
<td>10 (5, 4, 1)</td>
<td>306</td>
<td>8.00</td>
<td>0.22</td>
</tr>
<tr>
<td>C deg. – chitin</td>
<td>3 (acetylglucosaminidase, endochitinase, exochitinase) [3]</td>
<td>21 (6, 14, 1)</td>
<td>574</td>
<td>10.60</td>
<td>0.59</td>
</tr>
<tr>
<td>C deg. – hemicellulose</td>
<td>4 (ara, ara-fungi, xyA, xylanase) [5]</td>
<td>20 (5, 2, 7)</td>
<td>505</td>
<td>11.49</td>
<td>0.25</td>
</tr>
<tr>
<td>C deg. – lignin</td>
<td>3 (glx, mnp, phenol oxidase) [4]</td>
<td>9 (2, 1, 6)</td>
<td>330</td>
<td>6.89</td>
<td>0.34</td>
</tr>
<tr>
<td>C deg. – others</td>
<td>4 (aceA and B, assA, vanA, vdh) [8]</td>
<td>39 (9, 19, 3, 7, 1)</td>
<td>1258</td>
<td>24.08</td>
<td>0.12</td>
</tr>
<tr>
<td>C deg. – pectin</td>
<td>None detected [1]</td>
<td>0</td>
<td>33</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>C deg. – starch</td>
<td>3 (amyA, glucoamylase, pulA) [8]</td>
<td>27 (21, 1, 5)</td>
<td>772</td>
<td>15.78</td>
<td>0.48</td>
</tr>
<tr>
<td>C fixation – aOB</td>
<td>None detected [1]</td>
<td>0</td>
<td>19</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>C fixation – CODH</td>
<td>1 (carbon monoxide dehydrogenase) [1]</td>
<td>10</td>
<td>280</td>
<td>5.50</td>
<td>0.29</td>
</tr>
<tr>
<td>C fixation – pcc</td>
<td>1 (propionyl AOA carboxylase) [1]</td>
<td>16</td>
<td>408</td>
<td>7.87</td>
<td>0.35</td>
</tr>
<tr>
<td>C fixation – Rubisco</td>
<td>1 (cbbL) [1]</td>
<td>8</td>
<td>337</td>
<td>6.10</td>
<td>0.45</td>
</tr>
<tr>
<td>C-fixation- total</td>
<td>3 (CODH, pcc, cbbL) [4]</td>
<td>34</td>
<td>1044</td>
<td>19.47</td>
<td>0.50</td>
</tr>
<tr>
<td>Methane oxidation</td>
<td>2 (mmoX, pmoA) [2]</td>
<td>5 (1, 4)</td>
<td>118</td>
<td>2.31</td>
<td>0.15</td>
</tr>
<tr>
<td>Methanogenesis</td>
<td>1 (mcrA) [1]</td>
<td>2</td>
<td>136</td>
<td>0.46</td>
<td>0.07</td>
</tr>
<tr>
<td>C cycle total</td>
<td>27</td>
<td>170</td>
<td>5196</td>
<td>16.46</td>
<td>0.59</td>
</tr>
<tr>
<td>Aromatic carboxylic acids</td>
<td>22 (22 different genes detected) [38]</td>
<td>101</td>
<td>3302</td>
<td>29.48</td>
<td>0.78</td>
</tr>
<tr>
<td>BTEX and related</td>
<td>6 (alkB, catB, ebdABC, tuFDG, xylF, xylJ) [18]</td>
<td>19 (5, 6, 1, 2, 4, 1)</td>
<td>532</td>
<td>5.03</td>
<td>0.23</td>
</tr>
<tr>
<td>Chlorinated aromatics</td>
<td>5 (cbDA, fcbA, tfDA, tfDB, tfDH) [10]</td>
<td>16 (2, 1, 5, 1, 7)</td>
<td>460</td>
<td>4.30</td>
<td>0.44</td>
</tr>
<tr>
<td>Heterocyclic aromatics</td>
<td>2 (dbDC, hdnO) [9]</td>
<td>3 (1, 2)</td>
<td>70</td>
<td>0.54</td>
<td>0.07</td>
</tr>
<tr>
<td>Nitroaromatics</td>
<td>4 (nbaC, rnh, nmoA, nphA) [9]</td>
<td>18 (3, 4, 10, 1)</td>
<td>500</td>
<td>8.15</td>
<td>0.16</td>
</tr>
<tr>
<td>Other aromatics</td>
<td>11 (11 different genes detected) [22]</td>
<td>43</td>
<td>1087</td>
<td>10.33</td>
<td>0.07</td>
</tr>
<tr>
<td>Polycyclic aromatics</td>
<td>6 (bPHA, bPHC, naph, nidA, phdI, phdK) [18]</td>
<td>10 (2, 1, 2, 1, 1, 1)</td>
<td>241</td>
<td>2.80</td>
<td>0.14</td>
</tr>
<tr>
<td>Chlorinated solvents</td>
<td>4 (cmuA, dehH, exaA, rd) [6]</td>
<td>17 (1, 2, 4, 10)</td>
<td>355</td>
<td>7.44</td>
<td>0.32</td>
</tr>
<tr>
<td>Herbicides and related</td>
<td>8 (atZ, A, B, and C, mauA, B, pceE, phn, trzA and Ni) [12]</td>
<td>37 (6, 1, 4, 6, 8, 1, 7)</td>
<td>816</td>
<td>11.81</td>
<td>0.68</td>
</tr>
<tr>
<td>Other hydrocarbons</td>
<td>6 (alkB, H, and J, chnB, cprA, Xa monooxygenase) [15]</td>
<td>21 (4, 7, 2, 4, 1, 3)</td>
<td>454</td>
<td>8.08</td>
<td>0.21</td>
</tr>
<tr>
<td>Organic remediation-other</td>
<td>5 (alkK, cah, dmsA, xenobiotic reductase, scnABC) [12]</td>
<td>19 (9, 1, 6, 1, 2)</td>
<td>509</td>
<td>6.84</td>
<td>0.09</td>
</tr>
<tr>
<td>Organic remediation total</td>
<td>80</td>
<td>319</td>
<td>8614</td>
<td>29.46</td>
<td>0.42</td>
</tr>
<tr>
<td>Ammonification</td>
<td>1 (ureC) [2]</td>
<td>7</td>
<td>315</td>
<td>2.89</td>
<td>0.20</td>
</tr>
<tr>
<td>Anammox</td>
<td>1 (hzo) [1]</td>
<td>1</td>
<td>42</td>
<td>1.30</td>
<td>0.01</td>
</tr>
<tr>
<td>Assimilatory nitrate reduction</td>
<td>3 (nasA, nirA, and B) [3]</td>
<td>7 (3, 1, 3)</td>
<td>266</td>
<td>6.03</td>
<td>0.06</td>
</tr>
<tr>
<td>Denitrification</td>
<td>5 (narC, nirK, nirS, norB, nosZ) [5]</td>
<td>30 (8, 10, 3, 1)</td>
<td>1543</td>
<td>24.74</td>
<td>0.61</td>
</tr>
<tr>
<td>Dissimilatory nitrate reduction</td>
<td>2 (napA, rnfA) [2]</td>
<td>6 (2, 4)</td>
<td>262</td>
<td>3.55</td>
<td>0.09</td>
</tr>
<tr>
<td>Nitrification</td>
<td>1 (amoA) [2]</td>
<td>2</td>
<td>111</td>
<td>1.04</td>
<td>0.15</td>
</tr>
<tr>
<td>N fixation</td>
<td>1 (nifH) [1]</td>
<td>64</td>
<td>1224</td>
<td>60.45</td>
<td>0.25</td>
</tr>
<tr>
<td>N cycle total</td>
<td>14</td>
<td>170</td>
<td>3763</td>
<td>12.11</td>
<td>0.09</td>
</tr>
<tr>
<td>S oxidation</td>
<td>1 (sos)</td>
<td>7</td>
<td>195</td>
<td>25.69</td>
<td>2.36</td>
</tr>
<tr>
<td>S reduction</td>
<td>2 (dsr A and B)</td>
<td>40 (36, 4)</td>
<td>1123</td>
<td>69.71</td>
<td>0.32</td>
</tr>
<tr>
<td>Other</td>
<td>1 (aprA)</td>
<td>8</td>
<td>186</td>
<td>4.53</td>
<td>0.96</td>
</tr>
<tr>
<td>S cycle total</td>
<td>4</td>
<td>47</td>
<td>1504</td>
<td>6.20</td>
<td>0.46</td>
</tr>
</tbody>
</table>
$N_2$ fixation and assimilation

Hybridization to 64 of the 1224 nifH probes present accounted for 58% of NSI of all detected N-cycling genes (Table 4). Of the nif gene probes that hybridized to BSL water, only 14 were derived from cultivated and identified organisms: three from *Euarchaeota*, one from an extreme halophile, and two from methanogens. The remainder came from probes targeting primarily uncultivated and/or unclassified bacterial phyla. The highest signal intensity of any gene detected in this study, which represents 15% SI of all detected nif genes and c. 1% NSI for all detected genes is derived from a nif sequence detected in a molecular study of an oligonitrophilic rhizosphere (*Table S3*; *Hamelin et al.*, 2002). Despite numerous attempts and modifications of cycling conditions (*Table S2*), only one of the employed N-cycle primer sets (for nifH) resulted in amplicons of the expected size (*Table 3*). Analysis revealed two distinct sequence types (ST-1 and ST-2) that were related to one another by 55–59% nucleotide identity over the region sequenced. ST-1 had 98–99% nucleotide and 100% amino acid identity to the nifH gene from acidophilic verrucomicrobial methanotroph *Methylacidiphilum fumaroliicum* strain SolV. ST-2 had 95–97% nucleotide identity to sequences from uncultivated organisms from a Yellowstone hot spring (*Hamilton et al.*, 2011) and 82–84% identity (98–99% coverage) to the nifH sequence from *Anabaena variabilis* ATCC 29413. ST-2 had 93% amino acid identity to the NifH protein of *Thermococcus albus* (Accession no. YP_003473186.1) over 120 amino acids.

All three genes for assimilatory nitrate reduction probed on the GEOCHIP 3.0 hybridized to BSL DNA but at much lower SI than was observed for nifH alleles (Table 4). These included three *nasA* probes derived from uncultured bacteria, and nirB probes derived from two *Proteobacteria* and one *Verrucomicrobiae*. One *nirA* gene variant was detected from *Archaeoglobus fulgidus*, a species that has not been detected in any BSL clone libraries.

N dissimilation

As expected for a partially reduced environment, we observed hybridization to genes for anaerobic N utilization as well (*Table 4*). BSL DNA hybridized to both genes for dissimilatory nitrate reduction (*napA* and *nrfA*) that were probed on the GEOCHIP 3.0. The detected *napA* genes hybridized to probes derived from an uncultivated bacterium and a thermophilic carboxydotroph, *Carboxydothermus hydrogenoformans*. Four *nrfA* gene sequences from different *Proteobacteria* species, including *Thioalkalivibrio nitratireducens*, a haloalkaliphilic sulfur-oxidizer, were also detected. All five denitrification genes probed on the GEOCHIP 3.0 hybridized to BSL DNA, and the combined hybridization signals represented c. 24% of the NSI for all detected N-cycling genes (Table 4). Eight *narG* gene probes hybridized, and all but 1 (from *Pseudomonas* sp. MT-1) were derived from uncultivated bacteria. The signal intensity from hybridization to one *narG* probe represented 4% of the NSI for all N-cycle genes (and 16% of NSI for denitrification genes in both years). This probe corresponds to a putative dissipatory membrane-bound nitrate reductase from an uncultured bacterium in corn soils (*Philippot et al.*, 2002). Ten *nirK* and eight *nirS* probes, targeting nitrite reductase, hybridized to BSL DNA (*Table 4*), and all of these were derived from uncultivated bacteria. Probes for *norB* and *noZ*, targeting nitrous oxide reductase, also hybridized to BSL DNA. Two of the three *norB* gene probes detected were also from uncultivated bacteria. The single detected *noZ* probe was from *Rhodopseudomonas palustris*.

Other N processes

Genes for nitrogen oxidation were not well represented in our samples (*Table 4*). Only two *amoA* gene sequences for nitrification were detected in our samples, one from an uncultured marine *Crenarchaeota* (*Francis et al.*, 2005) and one from an uncultivated bacterium from a linear alkylbenzene sulfonate-degrading consortium (*Dong et al.*, 2004). A single *bzo* anammox gene, from Candidatus *Scalindua* sp. enrichment culture clone HZO10, was detected. Surprisingly, no positive hybridization signals were seen for nitrogen oxidation probes from Candidatus *Kuenenia stuttgartiensis*, though numerous other probes from this organism including, *alkK*, *copA*, *czcA*, *dsrA*, *tehB*, and two *mex* transporters did hybridize to BSL DNA.

Sulfur cycle

As expected, genes for S oxidation, reduction, and both were detected in BSL water. Hybridization to *dsrA* gene probes accounted for c. 62% of the NSI for sulfur-cycling genes, with 32 of the 36 detected sequences from uncultivated *Proteobacteria*. Only four *dsrB* gene variants were detected in BSL water, including two from methanogens, and one each from an uncultivated sulfate-reducing bacterium, and *Moorella thermaacetica*, a genus that has been detected in BSL sediments (*Wilson et al.*, 2008). The highest signal intensity for any detected gene sequence in the organic remediation category, and the third highest total signal intensity for any gene detected on the chip, corresponded to a putative membrane docking protein from *Desulfotobacterium chlororespirans* (*Clostridiales*), which is capable of using sulfite, thiosulfate, and sulfur as electron acceptors (*Sanford et al.*, 1996).

Hybridization to eight of the *aprA* gene probes represented c. 24% of NSI for detected sulfur-cycling gene probes.
The highest signal intensity (27% of total aprA signal) was from a probe for Pelodictyon clathratiforme, a green sulfur bacterium, and three additional detected gene probes were to Desulfothermus, a sulfate-respiring Firmicute. Detected sox genes accounted for 10% of the NSI for sulfur-cycling genes, with all seven detected sox probes derived from cultivated bacteria (Proteobacteria and Chlorobia).

### Primary and secondary production

Primary production, as measured by $^{14}$CO$_2$ uptake, was very low, barely above background (Fig. 5A–C), with $<$0.3% of label taken up in 2–3 h. Although DIC levels were below detection limits, assuming DIC of 0.03–0.5 ppm yields production rates of $c.$ 1–10 $\mu$g C L$^{-1}$ day$^{-1}$. $^{14}$CO$_2$ uptake in light was greater than in the dark in only one instance, in fall 2010, and then only slightly. In contrast, uptake of $^{14}$C-acetate was uniformly strong, with 30–70% of the label incorporated in 2–4 h (Fig. 5D–F). Rates were linear over 1–4 h and showed little difference between light and dark treatments in two of three incubations; in November 2010, the light bottles showed about twice as much uptake. In contrast, uptake of glucose, though linear over 4 h, was very low: only ca. 3% of label was assimilated (not shown). 'Bacterial secondary production' rates, as measured by $^3$H-thymidine incorporation into macromolecules, was also extremely low, though somewhat higher than primary production, $c.$ 6–19 $\mu$g C L$^{-1}$ day$^{-1}$, yielding division times of 2–3 day (Table 5). There was some evidence that production was repressed by 10–100 $\mu$M additions of simple sugars or xylan, though not by glucose polymers (Table 5).

### DISCUSSION

**Microbial community composition is only partially captured by either method**

16S rRNA gene library data and GeoChip results generated in this study suggest that the BSL planktonic habitat contains a relatively simple community in which Bacteria are slightly more prevalent than Archaea (Table 2). The most abundant bacterial clones are representatives of the genera Acidimicrobium, Hydrogenobaculum, and Acidocaldus, while the most abundant Archaea are from genera within the Sulfolobales (including Sulfolobus, Metallosphaera, and Stygiolobus). The low number of observed 16S rRNA PTs, and the unevenness of PT distribution are similar to what has been reported for other acidic thermal habitats (Bond et al., 2000; Siering et al., 2006; Kock &

---

Table 5 ‘Bacterial production’ estimated from $^3$H-thymidine uptake in Boiling Springs Lake water incubated in the laboratory under simulated in situ conditions. Samples were incubated with $1 \mu$Ci $^3$H-thymidine and rates were calculated from uptake over 1–4 h, as processed as described in text. Minimum inhibitory concentrations (MIC) of other C sources on $^3$H-thymidine uptake were also estimated for simple and polymeric C sources.

<table>
<thead>
<tr>
<th>Date</th>
<th>Water T (°C)</th>
<th>µg C L$^{-1}$ day$^{-1}$*</th>
<th>Acetate</th>
<th>Glucose</th>
<th>Starch, cellobiose</th>
<th>Xylan</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/22/2011</td>
<td>47†</td>
<td>16.2</td>
<td>nd</td>
<td>nd</td>
<td>nd</td>
<td>nd</td>
</tr>
<tr>
<td>7/27/2011</td>
<td>49</td>
<td>8.1</td>
<td>$&gt;$10 $\mu$M</td>
<td>$&gt;$100 $\mu$M</td>
<td>nd</td>
<td>nd</td>
</tr>
<tr>
<td>11/1/2011</td>
<td>46</td>
<td>6.9</td>
<td>nd</td>
<td>--</td>
<td>--</td>
<td>$&gt;$10 $\mu$M</td>
</tr>
</tbody>
</table>

nd, not determined; --, no inhibition to 1000 $\mu$M.*Assumes $2 \times 10^{18}$ cells per mol thymidine, and 10 fg C per cell.†Estimated T.

---

© 2013 John Wiley & Sons Ltd

---

**Fig. 5** Radioisotope uptake assays for Boiling Springs Lake for Summer, Late Fall, and Winter. (A–C) autotrophy from $^{14}$C-bicarbonate. (D–F) heterotrophy from $^{14}$C-acetate. Averages and ranges of replicate incubations are shown, except for single killed controls. Lines are interpolations between points.
Schippers, 2008; Wilson et al., 2008; Inskeep et al., 2010; Garcia-Moyano et al., 2012). However, hybridization to gyrB probes on the GeoCHIP 3.0 indicates that prokaryote diversity is likely to be greater than that estimated by 16S rRNA gene analysis.

The abundance of hybridizations to probes derived from Proteobacteria is difficult to interpret in light of the fact that no Gram-negative bacteria have ever been isolated from BSL, despite our growing collection of over eighty 16S rRNA-identified isolates. Only a few 16S rRNA gene clones from multiple libraries indicate the presence of Proteobacteria, with the most abundant being similar to facultative anaerobe Acidicaldus organivorans (Table 2), a mixotrophic sulfur-oxidizer, and heterotroph capable of anaerobic respiration using ferric iron. Some of the proteobacterial gene sequences detected in the GeoCHIP results correspond to organisms that are likely able to grow in BSL, but we suspect the high number of detected Proteobacteria genes is partly due to an over-representation of Proteobacteria on the chip (He et al., 2010). Only 3.3% of the proteobacterial probes resulted in positive hybridization signals, in contrast to 8–10% of the probes hybridizing for numerous other lineages.

Almost half of detected archaean genes on the GeoCHIP were from orders not detected in clone libraries, including the Halobacteriales, Methanobacteriales, Methanocorinales, Methanococcales, and the Methanobacterales. Twelve of these gene probes were for metal and/or antibiotic resistance genes that may have been horizontally transferred from these organisms into hosts present in our system. Given the relatively low number of probes for Crenarchaeota on the chip, and the low sequence identity of our detected archaean clones compared to their closest cultivated relatives, it is likely that the true archaean diversity of BSL is not yet defined. Hybridization to probes derived from uncultured and or uncharacterized Bacteria represented 16% of total NSI over both years. Three of the five probes exhibiting the highest total SI (and eight of the top 25) were from unclassified and uncultivated Bacteria (Table S3). The majority of probes detected from the uncultured Bacteria were for genes involved in nitrogen and sulfur cycling, suggesting that novel organisms may be responsible for these key ecosystem processes in the planktonic BSL community. This is similar to what has been found in AMD communities using GeoCHIP approaches (Xie et al., 2011).

Biases associated with WGA of DNA used in microarray experiments have been previously explored in pure culture and soil community investigations, and biases were found to be of limited consequence (Wang et al., 2011). However, it is possible that sample-specific amplification biases affected our results. We wanted to include replicate hybridizations in this study, and therefore needed to use WGA hybridizations because hybridization to non-amplified DNA required significantly higher amounts of DNA that were not available to us.

Metagenomics analysis and diverse cultivation efforts that mimic in situ conditions may allow us to better understand the BSL community. We have also recently employed tagged 16S rRNA gene pyrosequencing to identify organisms likely to be important in secondary production in BSL, and these data are currently being analyzed. Additionally, we have numerous bacterial and archaean isolates currently being characterized.

Biogeochemical potential of the microbial community

Autotrophy is low, and photosynthesis is dominated by eukaryotes

Although the GeoCHIP suggests that C fixation in BSL water can potentially occur via the hydroxypropionate pathway, the Calvin cycle, and carboxydotrophy (Fig. 4B), the number of probes hybridized was restricted to a few/sole dominant genera or uncultivated organisms, similar to what was seen in GeoCHIP 2.0 analysis of AMD habitats (Xie et al., 2011). Similarly, our PCR screens for C-cycling genes showed very little diversity (Table 3). The SI for the detected C-fixation genes, and our primary production assays (Fig. 5), suggest that C fixation is low, and may be less important than secondary heterotrophic production in sustaining this ecosystem. While there are few productivity data for acidic geothermal environments, our results are similar to those from studies on mesothermic coal mining acid lakes in the former East Germany, which suggest that production rates are very low, and often primary production is lower than bacterial 2' production due to C limitation at acid pH (Nixdorf et al., 2003). The temperature and pH of BSL are ideal for growth of Cyanidiales, a group of primitive red algae found globally in similar environments (Doemel & Brock, 1971; Toplin et al., 2008). There are no probes for the Cyanidiales order contained on the GeoCHIP 3.0, thus, we did not expect to detect this organism in the microarray experiments. However, visible biofilms of these organisms are seen in shallow shoreline areas, and we can easily isolate these organisms under heterotrophic or photosautotrophic conditions. As few prokaryote photosynthetic members are adapted to such low pH (Steinberg et al., 1998), eukaryotes are probably the major photosynthetic primary producers.

It is not clear what role the dominant silicate minerals, smectite and kaolinite, are playing in BSL’s microbiology. Both kaolinite precipitation (Fiore et al., 2011) and bioweathering (Ehrlich & Newman, 2009) are possible and open to further study. The low planktonic virus counts both in BSL and in other low-pH ecosystems are probably partly due to virus binding to the smectite and kaolinite minerals (Syngouna & Chrysikopoulos, 2010) and not
merely virus stability (Ortmann et al., 2006; Prangishvili et al., 2006).

The low light penetration due to the fine suspended minerals suggests that C fixation by chemolithotrophy may exceed phototrophy, an observation supported by radiotracer experiments (Fig. 5). Both Acidimicrobium and Hydrogenobaculum, the two most common PTs detected in clone libraries from both sites A and D (Table 2), can grow autotrophically by oxidizing ferrous iron and pyrite, or hydrogen, respectively (Clark & Norris, 1996; Stohr et al., 2001). An Acidimicrobium sp. was also detected as the major phylotype in PCR screens for genes involved in C fixation (Table 3), and we recently isolated three strains with 99–100% rRNA gene identity to PT1 (Table 1) (Acidimicrobium sp. Y0018) and have shown them to be capable of iron and pyrite oxidation in the presence of 0.01% yeast extract (Arroyo, 2012). Two other minor PTs detected were closely related to Metallophaga hakonensis and Stygiolobus azoricus, which also grow autotrophically (Segerer et al., 1991), and Vulcanaiseta distributa may also facilitate iron and sulfur cycling in the lake (Itoh et al., 2002; Johnson & Hallberg, 2009).

GeoChip results partially supported the importance of chemolithotrophic primary production. GeoChip 3.0 is lacking probes for genes involved in Fe oxidation but does contain probes for S and N oxidation (He et al., 2010). Although hybridization to S-cycling gene probes represented a small fraction of the normalized signal intensity (Fig. 4A), we did detect sulfur oxidation genes in BSL. Additionally, a hyperthermophilic chemolithotroph (T. albus) capable of aerobically oxidizing H₂, thiosulfate, or sulfur, was detected with a nifH gene probe. It seems likely that chemolithotrophy may be limited by a lack of oxygen in the system. At BSL’s temperature, O₂ saturation is <6 ppm, and is likely quickly sparged from water by the convective overturning of reduced gases. A relatively reduced water column is supported by the many reductive dissimilatory processes detected with GeoChip, and thus, alternate electron acceptors may also be used to support chemosynthesis, possibly including novel pathways.

Allochthonous C sources

Our ¹⁴CO₂ incorporation results support observations that while GeoChip detects CO₂ fixation genes, signals for heterotrophy are much stronger (Fig. 5), and planktonic production is likely fueled by both sources. Acetate in particular was incorporated rapidly and almost completely within 1–2 h of addition (Fig. 5), while glucose was utilized at much slower rates (Table 5). Interestingly, some C amendments seem to inhibit thymidine incorporation, which is consistent with observations that numerous BSL isolates grow best on media containing very low organic C (unpubl. data). Many of the organisms described for other low-pH environments appear to use mixotrophy as a common strategy to deal with limiting C in both organic and inorganic forms (Johnson & Hallberg, 2009), and many acidophilic algae have the ability to grow photoheterotrophically, or heterotrophically depending on carbon and light availability (Gross & Schnarrenberger, 1995; Titel et al., 2005).

The abundance and diversity of GeoChip probes for organic compound degradation that hybridized to BSL DNA is similar to that observed in carbon-limited AMD communities (Xie et al., 2011) and indicate the presence of a diversity of genes that participate in degrading allochthonous carbon compounds. Conifers dominate the area surrounding BSL, including ponderosa, Jeffrey, and sugar pines, incense cedar, and the ericaceous shrub pines manzanita. Needles and leaves from these plants enter the lake especially during snow and rain events, and forest litter decomposition products would also be expected to enter the lake during run-off periods. Insect-fall may also contribute significant amounts of organic matter to the lake, and we have occasionally observed the carcasses of birds, mice, and deer floating at the edge of the lake. Efforts are currently underway to quantify allochthonous carbon inputs into the system. Degradation products of plant material might be expected to be enriched in aromatic carboxylic acids, and may explain the abundance and diversity of the detected probes in this GeoChip category. Experiments in progress suggest rapid leaching of organic carbon from coniferous and ericaceous leaves under simulated lake conditions (first order mass loss rates c. 0.01–0.04 day⁻¹), followed by slower microbial colonization and breakdown (c. 0.0007–0.0044 day⁻¹). However, microbial breakdown rates appear to be comparable or faster than those for coniferous leaves in cool, neutral pH montane lakes (Harmon et al., 1990). Fungi, especially ascomycetes, appear to be active members of the BSL microbial community, as they are easily detected in SSU rRNA gene surveys (Wilson et al., 2008) and can be cultured at 40–50 °C (G. V. Wolfe, unpubl. data). GeoChip detected a number of ascomycete genes likely involved in breakdown of hemicellulose and/or cellulose from allochthonous woody material; this suggests that fungi, as well as prokaryotes, contribute to BSL secondary production. Interestingly, the dominant grazing organism in BSL, a heterolobose amoebofilagellate closely related to Tetramitus thermacidophilus, appears to eat both prokaryotes and fungal conidia (Reeder, 2011).

Biogeochemical model of BSL ecosystem

Although no biogeochemical models exist for similar natural geothermal systems, a model was recently hypothesized for the sediment community in the Rio-Tinto AMD site (Sanchez-Andrea et al., 2011) which is cooler in temperature and much higher in metal content than BSL. Our
current data, previous studies (Brown & Wolfe, 2006; Siering et al., 2006; Reeder, 2011) and not-yet-published cultivation studies allow us to begin to assemble a picture of the BSL planktonic ecosystem (Fig. 6). Our data suggest that BSL is a system with little photosynthetic input of fixed carbon, primarily provided by thermoacidophilic algae (of the Cyanidiales order) and a thermoacid-tolerant relative of a proteobacterium Rhodovastum sp., and it is likely that photosynthesis by both groups of organisms is limited by light availability due to suspended sediments; however, both of these organisms can also grow (photo)heterotrophically. Chemosynthetic primary production is likely driven by the (micro)aerobic oxidation of reduced iron and sulfur minerals by members of the Firmicutes and Actinobacteria phyla, the γ-proteobacterium Acidithiobacillus spp., and Archaea in the Sulfolobaceae family. However, it is likely that some of this iron and sulfur oxidation occurs mixotrophically when DOC is available. The GeoChip detection of sulfur oxidizing genes suggests a diversity of unclassified sulfur-oxidizing species could also be contributing to the synthesis of organic carbon. Additionally, hydrogen-mediated primary production via Hydrogenobaculum spp. is likely, especially in oxic-anoxic transition zones and the majority of the lake water where temperature and sulfate concentrations predict significantly decreased oxygen concentrations compared to atmospheric levels (Liljeqvist et al., 2013). Carboxydotrophy also seems to play a role in organic synthesis in this ecosystem, although the identity of the carboxydotrophs is

Fig. 6 Model of Boiling Springs Lake biogeochemistry. Synthesis of data from current and previous studies, and as yet unpublished cultivation studies referred to in this manuscript. Mean depth and steeply sloped sides as determined by bathymetric measurements are indicated. Detected, but unidentified organisms are indicated in red. Large downward arrow indicates secondary production contributions by diverse taxa in all three domains of life. Reduced iron, sulfur, and hydrogen are likely the primary modes of chemosynthetic primary production. Font size of species names indicate hypothesized relative importance in process indicated. Phototrophy by Cyanidium and Rhodovastum relatives is likely restricted to the shallow margins of lake where light penetration occurs. Gas inputs (in decreasing order) of CO₂, H₂S, H₂, and CH₄ also serve as important nutrient inputs into the system. Question mark for oxidation of nitrogen species indicates the lack of significant detection of N-oxidizing genes or species in our studies.

© 2013 John Wiley & Sons Ltd
The detection of numerous genes for reductive processes suggests that reduction of iron, sulfur, and nitrogen compounds is probably an important sink for organic carbon and a source of reduced energy. Several known iron reducers have been detected or isolated from the system (e.g. Acidimicrobium, Sulfoapillus, Alicyclobacillus tolerans and A. organivorus), and Acidimicrobium sp. have been found to be particularly abundant in clone libraries prepared from BSL water (Tables 2 and 3). Cytochrome and hydrogenase gene sequences from various Geobacter spp. were detected on the GeoChip 3.0 (not shown), suggesting that Geobacter relatives may be contributing to iron and sulfur respiration in BSL sediments. These genes also may have been horizontally transferred to thermoacidophilic anaerobic respirers that function in the system.

Given the lack of known acidophilic sulfate reducers, there were a surprisingly large number of dsrA gene probes that hybridized to BSL DNA. However, far fewer dsrA gene probes hybridized to BSL DNA than the 85 gene sequences detected in AMD habitats (Xie et al., 2011). While dissimilatory reduction of sulfate in acidic habitats has been documented (Fishbain et al., 2003; Johnson & Hallberg, 2009; Jameson et al., 2010), acidophilic sulfate reducers have proved difficult to isolate (Alazard et al., 2010). The high signal intensity associated with dsrA gene probes, and the presence of high sulfate in BSL waters, suggest that carbon limitation may restrict sulfate reduction in this habitat.

Our molecular and cultivation efforts shed some light on the biogeochemical processes but less on the players in the nitrogen cycle in BSL. Clearly, several distinct nitrogen fixation genes are abundant in the system (Tables 3 and S3, Fig. 6), which might reflect the paucity of available nitrogen (Siering et al., 2006). However, the identity of the dominant and minor diazotrophs is unclear. The GeoChip results presented here suggest that novel nifH or nif/H alleles function in nitrogen fixation under the low nutrient, low pH, and high temperature conditions in BSL. Most of the PCR primers we tested targeting nifH failed to produce amplicons (compare Table 3 to Table S2), although one primer set produced amplicons that were nearly identical to the nifH gene of the verrucomicrobial methanotroph M. fumarolicoicum, and other amplicons similar to nifH genes from Thermocrinus and Anabaena sp. Methylocacidiphilum fumarolicoicum has been prevalent in some BSL enrichment cultures. Numerous high-temperature, acidic habitats in Yellowstone National Park were shown to contain three novel lineages of nifH genes which appear to be active under the thermal, acidic conditions in these springs (Hamilton et al., 2011; Kikumoto et al., 2013); however, these nifH genes are distinct from those identified in our work. A limited number of diazotrophic enrichment/isolation attempts have proven inconclusive thus far (P. L. Siering, unpubl. data), but will be the focus of future cultivation efforts.

Furthermore, we do not understand to what extent nitrogen compounds are biotically oxidized or reduced, or what organisms are involved in these processes. Molecular and enrichment strategies have failed to detect bacterial or archaeal ammonia oxidation genes or organisms (P. Siering and J. de La Torre, unpubl. data), and the GeoChip data provide little additional insight in terms of how and by whom these ecosystem services are achieved. Similarly, PCR screens for dissimilatory nitrate reduction genes have proved unsuccessful, though the GeoChip data presented herein suggest many genes to target that have been identified from metagenomic data or other culture-independent investigations of nitrogen cycling in diverse habitats.

Summary

The majority of the lake appears to be well mixed and consistent in temperature (c. 52 °C), while the southeastern end is much warmer (c. 70–95 °C) and more variable with visibly active thermal inputs. The physical mixing within the water column allows for surprisingly similar planktonic assemblages despite a 30 °C temperature difference between the two ends of the lake (Fig. 1, sites A and D). Summer measurements of temperature, pH, chemistry, and 16S rRNA gene TRFLP patterns (data not shown) suggest a very stable community while winter measurements indicate a diverse suite of habitats and greater seasonal variability than is evident from summertime observations.

GeoChip analysis of DNA from BSL water suggests that the oligotrophic, low-nutrient habitat of BSL contains both autotrophic and heterotrophic C processes. Scavenging of very limited N and organic C, as well as significant N₂ fixation, is well supported by strong hybridization to probes for N and C scavenging genes in GeoChip. Although we lack data on Eh vs. depth, the rapid input of reduced gases (Janik & McLaren, 2010) and low solubility of O₂ at 52 °C suggest that reductive processes dominate. This is supported by hybridization to probes for both N and S cycle genes involved in anaerobic respiration.

Finally, several biogeochemically important processes are missing or undersampled with GeoChip 3.0. Like AMD sites, Fe redox processes are likely very important in BSL energy production. The recently recognized redox cycle of P (Hanrahan et al., 2005) is too new to have been incorporated into GeoChip, but is likely to be important in BSL, based on the importance of other reductive processes and very low phosphate concentrations. We suspect that C, N, and P might all act as limiting nutrients on microbial production in this environment. The low nutrients coupled with the high temperature and low pH of the system act to select for certain gene variants for key ecosystem
processes, and thus, BSL should be a good source for novel genes involved in nutrient cycling.

ACKNOWLEDGMENTS

Funding was provided by the National Science Foundation (nos. MCB-0702018, MCB-0702069, and MCB-0702020), HSU Howard Hughes Medical Institute undergraduate science education grant (no. 52002680), and a grant from the NASA Astrobiology Institute’s Directors Discretionary Fund grant no. NNA11AC01G. Jennifer Kyle was supported by the NASA Astrobiology Institute Postdoctoral Fellowship program. We are indebted to PSU mechanical engineering students, who designed and built the ROV as part of a capstone project in 2008 under the direction of Dr. Faryar Etesami, PSU. We thank Laura Ramos for GPS data collection, and ‘Mile Brian’ Peasley for assistance with the GIS programming and drafting. Data files relating to this project have been deposited with the Lassen Volcanic National Park Geographic Information System (LAVO GIS). We also acknowledge Billie Reeder (CSUC) for help with ROV and sediment data collection, Dave Brown & Rachel Teasdale (CSUC) for additional T data, Jose De La Torre (SFSU) for sharing unpublished results, and the NPS LVNP staff, Louise Johnson and Michael Magnuson, for their generous efforts assisting with site access and use of LVNP Science Center facilities for sample processing. The authors of this manuscript have no conflict of interest to declare with respect to publication of this manuscript.

REFERENCES


© 2013 John Wiley & Sons Ltd


SUBMITTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Data S1 Materials and methods.

Fig. S1 X-ray diffraction patterns of sediment collected from the middle of the lake (a) and near the lake edge (b). K = kaolinite, Al = alunite, S = smectite. Unmarked peaks have not been identified.

Fig. S2 ROV GPS position from 8-16-09, showing correlation between onboard GPS (Lowrance LMS-480M) and shore-based sub-meter SX Blue II GPS (Geneg). Shore-based positions available only when ROV was close to handheld receiver at N/W shore.

Fig. S3 Top: sonar soundings during 8-16-09 transect, showing strong echoes from steeply sloped sides (yellow), with weak echoes from gas bubbles (white arrows) or soft mud or sediment deposits, with hints of deep vents (red arrows). Bottom: gray trace shows raw sonar data, including upwelling gas bubble reflections. Red trace shows manual corrections for soft mud bottom.

Fig. S4 Contour maps of water temperatures at surface (left) and 1 m depth (right). Hot spots at 1 m were shallow near-shore sediments.

Fig. S5 Time-lapse images from 9-8-09, showing convection from several gas vents. Lake water has been contrast-enhanced to reveal suspended sediment. Red arrows show vent locations.

Table S1 Genes and PCR primers used to screen N cycle pathways.

Table S2 Genes and PCR primers used to screen CO₂ fixation pathways.

Table S3 Highest signal intensity hybridizations. The top 25 detected genes are shown in descending order; numbers rounded to nearest 0.1.

Video S1 a. BSL 8-3-10, b. 8-4-10: 15-s unedited images taken above lake, merged in Windows Movie Maker, with time stamp added. Site D is in foreground, site A at top.

Video S2 a. BSL 8-3-10, b. 8-4-10 contrast enhanced: images were batch-processed in Photoshop (Adobe Inc.) using the Magic Wand tool to select the lake water, and apply automatic contrast enhancement to help visualize water convection. This produces artifacts around the lake’s perimeter and shoreline.