ORIGINAL ARTICLE

Sediment microbial communities in Great Boiling Spring are controlled by temperature and distinct from water communities

Jessica K Cole[1], Joseph P Peacock[1], Jeremy A Dodsworth[1], Amanda J Williams[1], Daniel B Thompson[1], Hailiang Dong[2,3], Geng Wu[4] and Brian P Hedlund[1]

[1]School of Life Sciences, University of Nevada, Las Vegas, NV, USA; [2]Department of Geology and Environmental Earth Science, Miami University, Oxford, OH, USA; [3]State Key Laboratory of Biogeology and Environmental Geology, China University of Geosciences, Beijing, China and [4]State Key Laboratory of Biogeology and Environmental Geology, China University of Geosciences, Wuhan, China

Great Boiling Spring is a large, circumneutral, geothermal spring in the US Great Basin. Twelve samples were collected from water and four different sediment sites on four different dates. Microbial community composition and diversity were assessed by PCR amplification of a portion of the small subunit rRNA gene using a universal primer set followed by pyrosequencing of the V8 region. Analysis of 164 178 quality-filtered pyrotags clearly distinguished sediment and water microbial communities. Water communities were extremely uneven and dominated by the bacterium *Thermocrinis*. Sediment microbial communities grouped according to temperature and sampling location, with a strong, negative, linear relationship between temperature and richness at all taxonomic levels. Two sediment locations, Site A (87–80 °C) and Site B (79 °C), were predominantly composed of single phylotypes of the bacterial lineage GAL35 ($\bar{p} = 36.1\%$), *Aeropyrum* ($\bar{p} = 16.6\%$), the archaeal lineage pSL4 ($\bar{p} = 15.9\%$), the archaeal lineage NAG1 ($\bar{p} = 10.6\%$) and *Thermocrinis* ($\bar{p} = 7.6\%$). The ammonia-oxidizing archaeon 'Candidatus Nitrosocaldus' was relatively abundant in all sediment samples <82 °C ($\bar{p} = 9.51\%$), delineating the upper temperature limit for chemolithotrophic ammonia oxidation in this spring. This study underscores the distinctness of water and sediment communities in GBS and the importance of temperature in driving microbial diversity, composition and, ultimately, the functioning of biogeochemical cycles.
The ISME Journal (2013) 7, 718–729; doi:10.1038/ismej.2012.157; published online 13 December 2012
Subject Category: microbial population and community ecology
Keywords: archaea; bacteria; pyrosequencing; thermophiles; Great Basin; microbial community

Introduction

Over the last 20 years, our understanding of life at high temperature has benefitted immensely from cultivation-independent censuses of microbial communities in continental geothermal springs, particularly those in Yellowstone National Park (YNP). The majority of these studies applied low-throughput technologies, relying on PCR amplification, cloning and Sanger sequencing of the small subunit (SSU) rRNA gene as a molecular identifier (Barns *et al.*, 1994; Hugenholtz *et al.*, 1998; Reysenbach *et al.*, 1994, 2000; Blank *et al.*, 2002; Meyer-Dombard *et al.*, 2005; Spear *et al.*, 2005). In aggregate, these studies revealed that phyla that dominate most aquatic systems (for example, *Proteobacteria*, *Firmicutes*, *Actinobacteria*, and *Bacteroidetes*) are replaced with thermophilic specialists at temperatures $\geqslant 80$ °C and the surprising abundance of thermophilic bacteria, particularly *Aquificae* (Reysenbach *et al.*, 2005; Spear *et al.*, 2005). In addition, this research uncovered unexpectedly high diversity of yet-uncultivated lineages of archaea, in spite of the predominance of *Thermoprotei* among cultivated terrestrial thermophiles (Spear *et al.*, 2005). Despite their utility, these censuses were limited by the absence of spatial or temporal sampling and their relative shallowness; thus, they provided only a snapshot of the community and an incomplete understanding of factors that influence microbial community composition and diversity.

A few studies have addressed the relationship between temperature and microbial community structure by studying temperature gradients along geothermal outflows. Two studies examined S^0 and As^{5+}/Fe^{3+}-precipitating geothermal outflows from

sulfuric acid-buffered springs in YNP. Jackson *et al.*, 2001 (Jackson *et al.*, 2001) noted a transition from a community dominated by *Hydrogenobaculum* and *Desulfurella* in the 62.1 °C source and S^0 precipitation zone of Dragon Spring to a community with abundant novel archaea in the cooler As^{5+}/Fe^{3+} precipitation zone below 55 °C. A similar study of a diverted outflow from Perpetual Spring documented the presence of *Stygioglobus*, *Caldococcus*, *Caldisphaera*, *Thermocladium* and *Hydrogenobaculum* in mature biofilms in S^0-depositing zones of ~80–70 °C, yielding to possibly less diverse communities in the cooler As^{5+}/Fe^{3+}-precipitating zones (Macur *et al.*, 2004). Other studies focused on geothermal outflows in circumneutral springs in YNP. Meyer-Dombard *et al.* (2011) examined geothermal outflows from four springs in Sentinel Meadows in the Lower Geyser Basin, focusing on comparing microbial communities in similar springs with and without streamer biofilms. All springs contained abundant *Thermocrinis* at the source, with changes in diversity that were interpreted as a possible ecotone in the transitional area between hotter chemosynthetic and cooler photosynthetic systems in two of the springs.

Recent advances in DNA sequencing technologies have allowed for much deeper microbial community censuses through pyrosequencing of SSU rRNA gene fragments (pyrotags); however, to our knowledge, only one study has employed pyrotag sequencing to study the relationship between temperature and microbial community composition and diversity in terrestrial geothermal systems (Miller *et al.*, 2009). The study examined photosynthetic communities in two silica-depositing geothermal outflows, White Creek and Rabbit Creek, in the Lower Geyser Basin of YNP between 73 °C and 39 °C. Operational taxonomic unit (OTU) richness at 99% identity was negatively related to temperature and community similarity decreased exponentially with temperature difference, demonstrating that temperature was an important factor in controlling community composition.

Great Boiling Spring (GBS) is a large circumneutral spring in the U.S. Great Basin, the microbiology and geochemistry of which have been studied in some detail (Huang *et al.*, 2007; Costa *et al.*, 2009; Lefèvre *et al.*, 2010; Dodsworth *et al.*, 2011; Hedlund *et al.*, 2011; Miller-Coleman *et al.*, 2012). GBS is sourced with ~85 °C Na^+/Cl^--dominated hydrothermal fluid that accumulates in the source pool for ~1.5 days, leading to oxic conditions in the bulk water (Costa *et al.*, 2009). The geothermal source contains ammonia as the dominant form of dissolved inorganic nitrogen, which supports a highly active nitrogen cycle in which ammonia is oxidized to nitrite by a close relative of the ammonia-oxidizing archaeon 'Candidatus Nitrosocaldus yellowstonii' (Dodsworth *et al.*, 2011). In turn, oxidized nitrogen supports both dissimilatory nitrate reduction to ammonium and denitrification, which is incomplete at high temperatures, leading to high nitrous oxide flux from the spring source (Dodsworth *et al.*, 2011; Hedlund *et al.*, 2011). Cultivation-independent censuses of the microbial community in GBS that were performed using low-throughput Sanger sequencing revealed a high proportion of novel lineages of archaea and bacteria. (Costa *et al.*, 2009). This high degree of novelty was further supported by pyrosequencing of single samples of water and sediment communities, which also defined a stark contrast between the two communities (Dodsworth *et al.*, 2011).

The present study uses SSU rRNA gene pyrotag sequencing to examine microbial community composition and diversity in GBS within the context of spatial and temporal sampling. The primary goals were: (i) to determine the effect of temperature on community composition, diversity and inferred function (ii) to determine whether microbial communities in sediments and bulk water are similar or distinct and (iii) to evaluate the temporal stability of the microbial community in GBS in the face of temperature fluctuations. The current study assesses both archaea and bacteria at all taxonomic levels and includes analysis of water-borne cells as well as sediment samples above, near and below the photosynthetic temperature limit at several different sampling times.

Materials and Methods

Sample collection, DNA extraction, PCR and pyrosequencing

GBS is located at N40°39′41″ W119°21′58″. Five sample locations, sites A, B, C, D and E, (Supplementary Figure S1) were identified within GBS for sediment sampling. The sediment/water interface (top ~1 cm) was sampled using sterile technique. Bulk water samples (site W) were collected using either normal filtration, with a 0.2 μm Supor filter (hydrophilic polyethersulfone, Pall Corporation, Port Washington, NY, USA) or tangential flow filtration (TFF; prep/scale filter with 30 kDa molecular weight cutoff, Millipore, Billerica, MA, USA). Sampling was performed at all sites in February, 2010, with additional samples taken at a subset of the sites in July, 2007, December, 2008, June, 2009 and July, 2010 (Table 1). Samples are identified in the text by a four number code indicating the month and year of sampling (in YYMM format) followed by a letter indicating the sample site; for example, 0706W refers to the bulk water sample taken during June 2007. Sediment and normal flow filters were placed on dry ice immediately and transported to the laboratory (<36 h) where they were stored at −80 °C until processing. Tangential flow filtration concentrates were stored on ice for <24 h, after which the cells were pelleted by centrifugation at 6000 rcf for 10 min and stored at −80 °C until processing. Temperature and pH

Table 1 Physical, chemical and mineralogical data, and sequencing statistics

Sample Name	Site[a]	Date (m-y)	Temp. (°C)[b]	pH	Predominant minerals[c]	Oxygen (μM)	NH_4^+ (μM)	NO_2^- (μM)	NO_3^- (μM)	N (initial)[d]	N (final)[e]
0706W	W	Jun-07	80	7.40	n/a	n/a	71	1.90	14.0	15 003	9781
0812W	W	Dec-08	80	7.20	n/a	28	70	0.79	1.7	35 116	17 624
0906W	W	Jun-09	82	6.85	n/a	31	93	1.04	8.8	12 266	8046
1002W	W	Feb-10	82	6.63	n/a	19	69	(bd)	7.2	21 841	13 891
0812A	A	Dec-08	80	7.20	Al, Am, Mu, Q	28	70	0.79	1.7	57 106	37 890
0906A	A	Jun-09	82	6.85	As, Mu,Q	31	93	1.04	8.8	10 511	5 706
1002A	A	Feb-10	82	6.63	An, L, Mu, Q	25	69	(bd)	7.2	27 235	17 944
1007A	A	Jul-10	87	n/a	Mu, Q	n/a	n/a	n/a	n/a	26 687	6689
1002B	B	Feb-10	79	6.81	Mu, Q	25	25	(bd)	3.0	30 813	20 720
1002C	C	Feb-10	72	7.05	Mi, Mu, Pa, Ps, Q	56	15	0.80	1.6	15 877	9768
1007C	C	Jul-10	76	n/a	Mu, Q	n/a	n/a	n/a	n/a	16 742	5907
1002E	E	Feb-10	62	7.24	Mu, O, Q	56	47	0.80	9.1	16 522	10 212
									Sums:	285 719	164 178

[a]See Figure 1 for site locations in GBS.
[b]Temperature for sediment samples recorded in water directly above sample sites.
[c]Al, albite; Am, analcime; An, andalusite; As, andesine; bd, Below detection limit ($<0.1\,\mu M$); L, labradorite; Mi, microcline; Mu, muscovite; n/a, Data not collected; O, orthoclase; Pa, paragonite; Ps, pseudomalachite; Q, quartz.
[d]Pyrotags generated by 454 sequencing.
[e]Pyrotags remaining after chimera check and other quality filters.

were measured in the bulk water using a LaMotte pH5 Series pH/temperature meter (LaMotte, Chestertown, MD, USA).

DNA was extracted from sample 0812W using the Joint Genome Institute's (JGI) cetyltrimethyl ammonium bromide protocol (Department of Energy Joint Genome Institute & California, 2010). DNA was extracted from all other samples using the FastDNA Spin Kit for Soil (MP Biomedicals, Solon, OH, USA) using modifications specified previously (Dodsworth et al., 2011). DNA was precipitated with 70% ethanol and resuspended in $0.5 \times TE$ (5 mM Tris, 0.5 mM EDTA, pH 8). DNA retrieved from the sediment sample collected at site D was insufficient for PCR and therefore site D was excluded from further analysis. DNA was shipped on dry ice to the JGI, where a portion of the SSU rRNA gene, including the V6–V8 hypervariable regions, was amplified using JGI's universal PCR primers 926F (5′-cctatccctgtgtgccttggcagtctcagAAACTYAAAKGA ATTGACGG-3′) and 1392R (5′-ccatctcatccctgcgtg tctccgactcag-<XXXXX>-ACGGGCGGTGTGTRC-3′). Primer sequences were modified by the addition of 454 A or B adapter sequences (lower case). In addition, the reverse primer included a 5 bp barcode, shown above, for multiplexing of samples during sequencing. Twenty μl PCR reactions were performed in duplicate and pooled to minimize PCR bias using 0.4 μl Advantage GC 2 Polymerase Mix (Advantage-2 GC PCR Kit, Clontech, Mountain View, CA, USA), 4 μl 5 × GC PCR buffer, 2 μl 5 M GC Melt Solution, 0.4 μl 10 mM dNTP mix (MBI Fermentas, Glen Burnie, MD, USA), 1.0 μl of each 25 nM primer and 10 ng sample DNA. The thermal cycler protocol was 95 °C for 3 min, 25 cycles of 95 °C for 30 s, 50 °C for 45 s, 68 °C for 90 s and a final 10-min extension at 68 °C. PCR amplicons were

purified using solid-phase reversible immobilization beads and quantified using a Qubit flurometer (Invitrogen, Carlsbad, CA, USA). Samples were diluted to 10 ng/μl and mixed in equal concentrations. Emulsion PCR and sequencing of the PCR amplicons were performed following the Roche 454 GS FLX Titanium technology (454 Life Sciences, Branford, CT, USA) manufacturer's instructions. Results obtained with the universal JGI primer set used for amplification indicated a lack of detection of archaea known to be present in GBS (data not shown). Comparison of the standard JGI forward primer with near full-length SSU rRNA gene sequences obtained from GBS (Costa et al., 2009) and other geothermal environments revealed that the corresponding sequence in most archaea had at least one mismatch to the forward primer (data not shown). Therefore, a revised forward primer, 926F454TitFNew (AAACTYAAAKGAATTGRCGG), was designed and used for pyrotag amplification. Choice of primer pair during PCR can heavily influence estimates of microbial richness and evenness due to varying degrees of binding efficiency between primers and the sequences of the various SSU rRNA gene primer binding sites present in the sample (Engelbrektson et al., 2010). In this case, use of the newly designed primer pair led to a dramatic increase in species richness, as compared with a dataset obtained from using the standard primer set (data not shown).

Sequence processing
285 719 pyrotags from 12 samples were quality filtered, aligned and analyzed using the mothur software package v. 1.20.2 (Schloss et al., 2009). Quality filtering discarded sequences with

ambiguous base calls, homopolymers of seven nucleotides or longer, and sequences with an average quality score <27 (as recommended by (Kunin & Hugenholtz, 2010)) over a window size of 50 nt. The mothur-provided bacterial and archaeal Silva reference files were concatenated into a single file against which quality-filtered sequences were aligned using default Needleman–Wunch algorithm parameters. The alignment was manually curated and sequences that did not end at nucleoide 1386 in the *E. coli* SSU rRNA gene were discarded. Preclustering was performed at a level of 1% difference. The alignment was checked for chimeras using mothur's chimera.slayer command (modeled after ChimeraSlayer by Broad Institute, Cambridge, MA, USA; http://microbiomeutil.sourceforge.net/#A_CS) with abundant sequences in the sample set serving as the reference set. Sequences with very low abundance or those that did not align appropriately were searched against the NCBI BLAST database using blastn to further evaluate sequence quality (Zhang *et al.*, 2000). Confirmed and possible chimeric sequences, eukaryotic sequences, chloroplast sequences and low-abundance sequences with little similarity to any known organism were discarded, resulting in a curated set of 164 178 sequences, 200–230 nt in length, approximately covering bases 1156–1377 (including the V8 hypervariable region) of the *E. coli* SSU rRNA gene. SFF files containing the original unfiltered pyrosequences were submitted to the NCBI Sequence Read Archive (0812W, SRX201232; 0812W, SRX201232; 0906W, SRX201233; 1002W, SRX201234; 0812A, SRX201235; 0906A, SRX201236; 1002A, SRX201239; 1007A, SRX201241; 1002B, SRX201242; 1007C, SRX201244; 1002C, SRX201245; 1002E, SRX201246) and associated with NCBI BioSample SUB112233 and NCBI BioProject SUB112230.

The Quantitative Insights into Microbial Ecology (QIIME) software package v 1.3.0 (Caporaso *et al.*, 2010) was used to designate OTUs using the UCLUST algorithm (Edgar, 2010). Taxonomy was assigned according to the QIIME-compatible Green-Genes database-derived OTU set released on 4th Feb, 2011 (DeSantis *et al.*, 2006) using the blast algorithm, and also the QIIME-compatible pre-built Ribosomal Database Project (RDP) using the RDP classifier algorithm (Cole *et al.*, 2009). Each OTU was assigned to the most detailed lineage in the database matching $\geqslant 90\%$ of the sequences within the OTU. Where the taxonomy assignments of the 97% OTUs derived from the GreenGenes database and the RDP classifier disagreed, the representative sequences of those OTUs were submitted to NCBI blast to further assess the taxonomic assignment (Johnson *et al.*, 2008). The GAL35 group is currently one of two candidate class-level groups within the candidate phylum OP1 but was given a phylum-level designation in a previous edition of the Greengenes phylogeny (29 November, 2010). In our own phylogenetic analyses, the position of the GAL35 group is uncertain, but may be affiliated with the *Chloroflexi* (Costa *et al.*, 2009) (Dodsworth *et al.*, unpublished results). We therefore chose the conservative approach of retaining its designation as a distinct lineage rather than grouping it with other previously named phylum-level groups.

Statistical analyses

Rarefaction and collector's curves were generated for OTUs observed at the 97% OTU level using mothur. Rarefaction curves were prepared at sequencing depth of $\leqslant 915$ for archaea and $\leqslant 4410$ for bacteria. QIIME was used to calculate alpha diversity measures at OTU definitions of 80, 85, 92, 95 and 97% sequence identity, including OTUs observed, Chao1 (Chao, 1984), the Shannon Index, and Simpson's Index of Diversity. Simpson's Evenness was calculated in Microsoft Excel 2007.

QIIME was used to perform beta diversity analyses. Bray–Curtis dissimilarity (Bray and Curtis, 1957) was calculated at the 97% OTU level and principal coordinates analysis and hierarchical cluster trees were constructed from the resulting matrix. A hierarchical cluster tree at high taxonomic level (phylum level for bacteria and class level for archaea) was constructed based upon taxonomy assigned to 97% OTUs.

Analysis of Similarity (ANOSIM) and Similarity Percentage (SIMPER) calculations were conducted using PAST v. 2.11 (Hammer *et al.*, 2001). ANOSIM was used to test for significant differences between sample groups defined by the 97% OTU cluster tree. Holm's sequential Bonferroni correction was used to correct for multiple comparisons (Holm, 1979). Those groups found to be significantly different according to uncorrected *P*-values were subjected to SIMPER analysis based on Bray–Curtis dissimilarity in order to define the OTUs primarily responsible for the differences between the groups.

Pearson's χ^2 goodness-of-fit calculations and linear regressions were performed using the R Project for Statistical Computing (R Development Core Team, 2011). In regressions of diversity versus temperature, sediment community samples collected at the same location in different years were treated as independent samples because the temperatures differed between years and the time intervals between sampling were sufficiently long that communities could adjust to changed temperature conditions.

Mineralogy

Samples were dried at ambient temperature in an anaerobic chamber (Coy, Gross Lake, MI, USA) with an atmosphere of 90% N_2, 5% CO_2, and 5% H_2. Samples were then scanned with a Scintag X1 X-ray powder diffractometer (Scintag, Cupertino, CA, USA) using CuKα wavelength, a fixed slit scintillation detector, and a power of 1400 W (40 kV, 35 mA).

Scans were performed from 2–70° 2-theta stepping at 0.02° with a count time of 2 s per step. Qualitative identification of mineral phase was made by using the Jade 7 program (MDI, Livermore, CA, USA). This program utilized the International Center for Diffraction Data Powder diffraction File database (ICDD PDF-2, Sets 1–46, 1996) as a reference source.

Results and discussion

Overview of pyrotag dataset

A total of 164 178 quality-filtered SSU rRNA gene pyrotags were obtained from twelve samples collected on up to four sampling dates from bulk water (site W) and four sediment sites (sites A, B, C and E; Figure 1). Total community OTU abundance ranged from 25–252 at the 97% level and from 10–27 at the 80% level, roughly corresponding to species and phyla, respectively (Supplementary Table S1). The V8 region, which possesses less variation than the unmasked full-length SSU rRNA gene, was the only hypervariable region contained within the quality-filtered pyrotags. Use of only the V8 region likely caused an underestimation of the richness of the communities sampled in this study (Youssef et al., 2009). Collector's curves suggested that sampling did not approach completion at any sample location (Supplementary Figure S1).

Water and sediment communities are distinct

Cluster analysis calculated at the phylum/class level (Figure 2a) and the species level (Figure 3a) revealed distinct clustering of microbial communities from water and sediment samples. Furthermore, principal coordinates analysis at the species level showed that the primary axis of variation separated water versus sediment communities (Figure 3b). ANOSIM confirmed highly significant differences between water and sediment communities ($P < 0.01$, corrected $P < 0.05$; Supplementary Figure S2, Node 1). In order to isolate differences between water and sediment communities from variation in sample temperature, ANOSIM was performed comparing the microbial communities in the four water samples against sediment samples collected within the same temperature range (82–80 °C; 0812A, 0906A, 1002A). This analysis demonstrated a significant difference between the communities in the water and equivalent-temperature sediment samples ($P < 0.01$, corrected $P < 0.05$; Supplementary Figure S2, AB-W). Alpha diversity calculations with the same equivalent-temperature sample set showed that water and sediment communities had similar richness at all taxonomic levels; however, Student's t-tests indicated the water community was significantly less even at OTU levels from 95, 92 and 85%, roughly corresponding to the level of genus, family and order, respectively (Figure 4; Supplementary Tables S1–S3).

SIMPER analysis was performed to identify species that contributed to the dissimilarity between water and sediment samples (Table 2). Two OTUs, a Thermocrinis sp. ($\bar{p} = 89.47\%$) and Pyrobaculum caldifontis ($\bar{p} = 8.63\%$), dominated water communities, yet each was present at low relative abundance in sediment communities. The high relative abundance of Thermocrinis and Pyrobaculum in the water column could result from these populations shedding from sediment communities; however, the lower evenness and higher abundance of cells in the GBS water column compared with a nearby spring with a much shorter water residence time
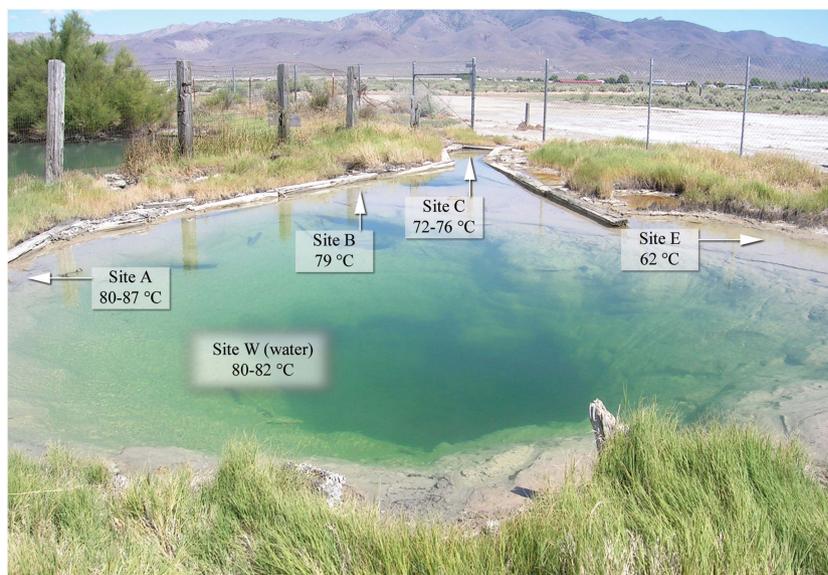

Figure 1 Photograph of GBS with sampling sites identified. Temperature range displayed at each site represents the temperatures recorded during sample collection. Temperature of sediment samples measured in water above each sample site.
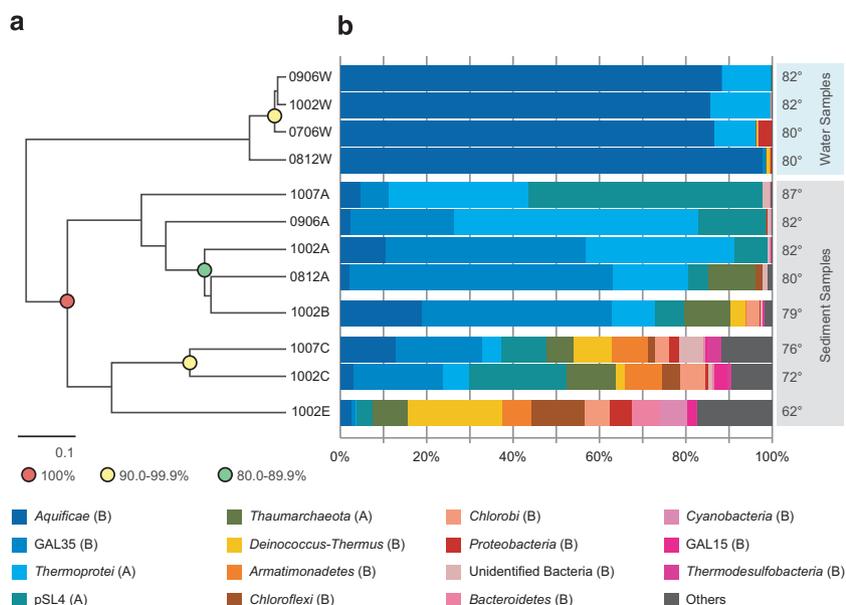
Figure 2 Visual representation of the similarity of samples based on their community composition, considered at the phylum level for bacteria and the class level for archaea. (**a**) Cluster tree calculated using Bray–Curtis dissimilarity, with jackknife values ⩾80.0% displayed. (**b**) Bar chart displaying the community composition of each sample, including the 15 most abundant taxa in all samples with remaining taxa included as 'Others'. '(A)' or '(B)' after taxon name in legend designates an archaeal or bacterial taxon, respectively.
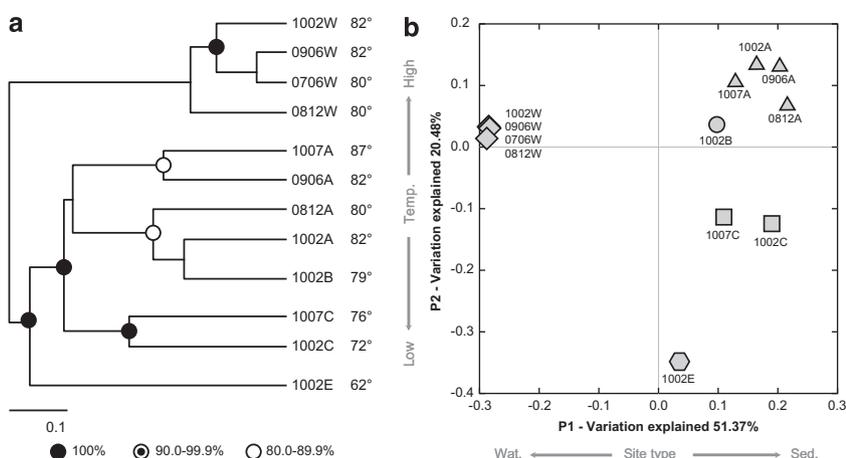

Figure 3 Visual representation of the similarity of samples based on their community composition, considered at the 97% OTU level. (**a**) Cluster tree calculated using Bray–Curtis dissimilarity, with jackknife values ⩾80.0% displayed. (**b**) Principal coordinates analysis constructed with Bray–Curtis dissimilarity. Principal coordinate 1, P1 and principal coordinate 2, P2, plotted against each other, with a total of 71.85% of variation explained.

(Dodsworth *et al.,* 2011) suggest that these two genera live and reproduce in the water column. Other taxa that contributed to dissimilarity of water and sediment communities were more abundant in sediments, including an *Aeropyrum* sp. and an OTU related to the novel archaeal lineage NAG1 ((Kozubal *et al.,* 2012); Supplementary Table S4).

The distinctness of water and sediment communities is well known in most natural aquatic systems and has been identified in a large meta-analysis as a primary driver delineating microbial community composition (Lozupone and Knight, 2007). These differences segregate biogeochemical functions in

natural aquatic systems. For example, in marine systems, the photic zone of the water column is characterized by primary productivity and highly dynamic recycling of both elements and energy (that is, the 'microbial loop'), whereas the deep sea water column and sediments are essentially heterotrophic systems that are responsible for mineralization of sinking material from above (reviewed in (Fenchel, King and Blackburn, 1998)). While biologists studying thermophilic viruses focus entirely on water-borne material, the majority of research on the biology of geothermal systems focuses entirely on sediments or mineral precipitates and potentially
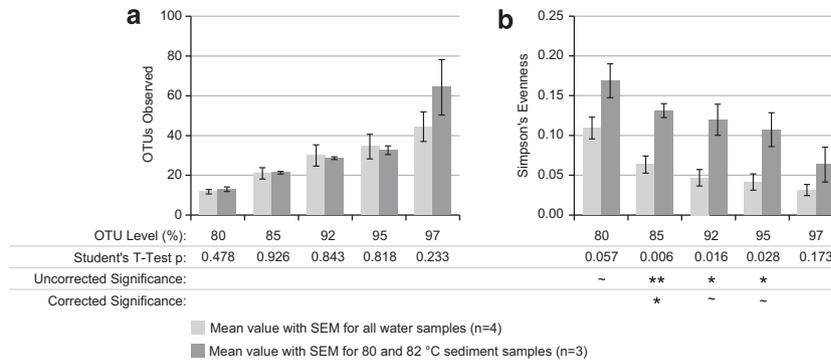
Figure 4 Comparison of alpha diversity measures of complete communities (both archaea and bacteria) in water samples to sediment samples of comparable temperature (82–80 °C). Calculations were performed with OTUs determined at five sequence identity levels and errors bars indicate s.e.m. Student's *t*-tests were performed for each pair of comparisons, water vs sediment, at each OTU level to determine significance. ***, significant at $\alpha = 0.001$; **, significant at $\alpha = 0.01$; *, significant at $\alpha = 0.05$; ~, significant at $\alpha = 0.1$. (**a**) Richness measured in the water samples compared with that of sediment samples of comparable temperature. (**b**) Simpson's evenness measured in the water samples compared with that of the sediment samples.

Table 2 SIMPER analysis results displaying top 10 97% OTUs responsible for dissimilarity between water and sediment samples[a]

Taxon[b]	Genus	Contrib (%)[c]	Mean abund[d] Wat (%)	Mean abund Sed (%)
Aquificae (B)	Thermocrinis	45.31	89.47	6.73
GAL35 (B)	Unidentified GAL35	14.77	0.20	27.18
Thermoprotei (A)	Aeropyrum	6.14	0.07	11.27
pSL4 (A)	Unidentified pSL4	5.88	0.04	10.77
Thermoprotei (A)	Pyrobaculum	4.19	8.63	1.87
Thermoprotei (A)	Unidentified NAG1	3.64	0.02	6.68
Thaumarchaeota (A)	Candidatus Nitrosocaldus	2.78	0.02	5.09
pSL4 (A)	Unidentified pSL4	2.00	0.00	3.64
Deinococcus-Thermus (B)	Thermus	1.90	0.28	3.43
Armatimonadetes (B)	Unidentified OS-L	1.15	0.01	2.10

[a]Pooled water samples compared to pooled sediment samples.
[b]Phylum-level taxa for bacteria and class-level taxa for archaea.
[c]Contribution of OTU to overall dissimilarity between groups.
[d]Average abundance of OTU in each group.

overlooks an important community of organisms in the water column.

Temperature controls richness and composition of sediment communities at all taxonomic levels
The effects of temperature on microbial communities were explored using two different approaches: first, by determining whether microbial composition clusters were concordant with differences in temperature, and, second, by determining whether quantitative relationships exist between measures of OTU diversity and temperature (Supplementary Table S1). Cluster analysis calculated at the phylum/class level (Figure 2a) and the species level (Figure 3a) showed that sediment samples clustered according to temperature and principal coordinates analysis at the species level showed that temperature corresponded with the secondary axis of variation (Figure 3b). ANOSIM confirmed a significant difference between sediment communities that were divided into three temperature groups,

87–79 °C (sites A and B), 76–72 °C (site C) and 62 °C (site E) ($P < 0.05$, corrected $P < 0.05$; Supplementary Figure S2b).

Within the sediment samples, both Simpson's and Shannon Diversity Index exhibited an inverse relationship between microbial diversity and temperature (Supplementary Tables S1–S3). For example, both Simpson's diversity values and Shannon Index for 97% OTUs had negative, linear relationships with temperature ($R^2 = 0.60$ and $R^2 = 0.76$, respectively).

To disentangle the components of alpha diversity, the relationships between temperature and evenness and richness were assessed separately. Using the combined archaea and bacteria dataset, richness at every OTU level showed a highly significant, negative, linear relationship with temperature both before and after correction for multiple comparisons ($R^2 \geqslant 0.88$, $P < 0.001$, corrected $P < 0.01$; Figure 5a). The estimated slopes (OTUs per °C ± s.e.) ranged from -8.57 ± 1.37 for 97% OTUs to -0.73 ± 0.09 for 80% OTU richness (Figure 5a). When assessed
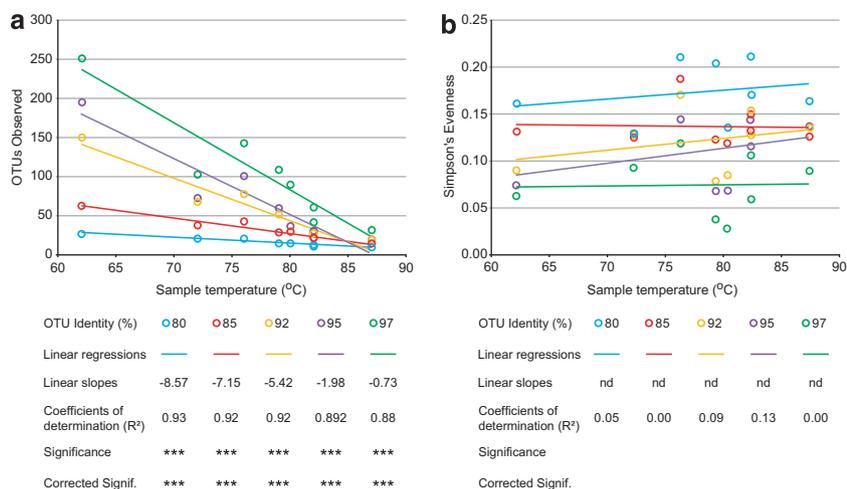
Figure 5 Alpha diversity measures of complete communities (combined archaea and bacteria) for all sediment samples, plotted against sample temperature, with regression lines for each OTU level. ***, significant at $\alpha = 0.001$; **, significant at $\alpha = 0.01$; *, significant at $\alpha = 0.05$; $\sim$, significant at $\alpha = 0.1$. (**a**) Richness vs sample temperature. (**b**) Simpson's evenness vs sample temperature. nd, slope not significantly different from zero.

separately, bacterial richness showed a significant, negative, linear relationship with temperature at all OTU definitions ($R^2 = 0.88$, $P < 0.001$, corrected $P < 0.001$; Supplementary Figure S3b), whereas archaeal richness only showed a significant relationship with temperature at the 80% ($R^2 = 0.89$ $P < 0.001$, corrected $P < 0.01$) and 85% OTU levels ($R^2 = 0.64$, $P < 0.05$, corrected $P < 0.1$) (Supplementary Figure S3a). The Chao1 richness estimator showed a similar negative relationship between estimated richness and temperature; however, the relationship was weaker than that observed for OTUs (Supplementary Table S4).

Many factors typically covary with temperature in geothermal systems, due to processes such as degassing, mineral precipitation, evaporation, autotrophy and oxidation (Nordstrom et al., 2005; Fouke, 2011). In GBS, cooler temperature covaried negatively with dissolved oxygen concentration ($r = -0.88$) (Supplementary Figure S4). Alpha diversity measurements also covaried to some degree with pH and oxygen (Supplementary Table S4). However, multiple regressions of combined OTU richness against the chemical variables showed that temperature accounted for more variation in richness than any chemical variable. In multiple regressions of 80, 85 and 95% OTU levels, richness against physicochemical variables, there were highly significant, negative relationships between richness and temperature after adjusting for other environmental variables ($R^2 = 0.99$, $P = 0.006$; $R^2 = 0.99$, $P = 0.006$; $R^2 = 0.97$, $P = 0.009$; respectively). Although other unmeasured physicochemical variables cannot be ruled out as factors that influence microbial community composition, these results suggest temperature is the major factor driving community composition and structure. The homogeneous mineralogy of GBS sediments

(Table 1) eliminated sediment heterogeneity as a factor.

Statistical analyses performed with the combined dataset (Figure 5b) or the bacteria-only dataset (Supplementary Figure S4d) revealed no significant relationships between temperature and evenness at any OTU definition. However, a significant, positive, linear relationship between evenness and temperature was evident in the archaeal dataset at OTU levels 80, 85, 92 and 95% before correction and 80 and 95% after correction (Supplementary Figure S4c).

To further investigate the relationship between temperature and diversity, we examined the community similarity as a function of pairwise temperature difference between samples. This revealed an exponential decay in community similarity as the temperature difference between samples increased (Figure 6). Our results were similar to those observed by Miller et al., 2009 in the photosynthetic zones of two alkaline geothermal outflows in YNP, ranging from 74–39 °C. However, we observed a much steeper decline in species richness as temperature increased within the temperature range we had studied, 87–62 °C, suggestive of a much stronger selective force at high temperature. To our knowledge, diversity decay dynamics along a temperature gradient had not rigorously been addressed at temperatures above the photosynthetic limit.

SIMPER was used to identify taxa that contributed most to differences between sediment communities in the three temperature groups confirmed to be significant by ANOSIM, 87–79 °C (sites A and B), 76–72 °C (site C) and 62 °C (site E) (Table 3). The highest temperature sites, A (87–80 °C) and B (79 °C), were predominantly composed of single phylotypes of the bacterial lineage GAL35 ($\bar{p} = 36.1\%$), *Aeropyrum* ($\bar{p} = 16.6\%$), the archaeal
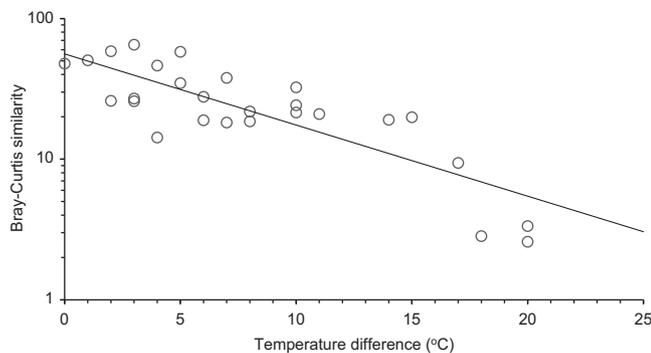
Figure 6 Exponential decay of community similarity as the difference between sample temperatures increases. The linear regression between log-transformed Bray–Curtis similarity and the difference between sample temperatures was highly significant ($y = 56.021$ e$^{-0.116x}$, $R^2 = 0.73$, $P < 0.001$).

Table 3 SIMPER analysis results displaying top ten 97% OTUs responsible for dissimilarity between high temperature (A,B) vs medium temperature (C) vs low temperature (E) sediment[a]

Taxon[b]	Genus	Contrib[c] (%)	Mean abund A,B (%)[d]	Mean abund C (%)	Mean abund E (%)
GAL35 (B)	Unidentified GAL35	17.15	36.16	17.76	0.95
Thermoprotei (A)	Aeropyrum	8.85	16.63	3.51	0.00
pSL4 (A)	Unidentified pSL4	8.38	15.91	3.27	0.01
Deinococcus-thermus (B)	Thermus	6.66	0.08	3.04	21.06
Thermoprotei (A)	Unidentified NAG1	6.20	10.64	0.08	0.00
pSL4 (A)	Unidentified pSL4	5.40	0.57	11.30	3.80
Thaumarchaeota (A)	Candidatus Nitrosocaldus	3.95	4.35	8.89	1.22
Armatimonadetes (B)	Unidentified OS-L	3.82	0.02	8.02	0.68
Aquificae (B)	Thermocrinis	3.66	7.57	6.84	2.25
Chlorobi (B)	Unidentified Chlorobi	2.21	0.65	4.49	3.88

[a]Sites A and B sediment samples compared to Site C sediment samples compared to Site E sediment sample.
[b]Phylum-level taxa for bacteria and class-level taxa for archaea.
[c]Contribution of OTU to overall dissimilarity between groups.
[d]Average abundance of OTU in each group.

lineage pSL4 ($\bar{p} = 15.9\%$), the archaeal lineage NAG1 ($\bar{p} = 10.6\%$) (Supplementary Figure S5c) and *Thermocrinis* ($\bar{p} = 7.6\%$). These five species were present at lower temperature sites C (76–72 °C) and E (62 °C), but were significantly less abundant ($\chi^2 = 77909$, $P < 0.0001$; Figure 2b, Supplementary Figure S6). Bacteria made up a significantly higher percentage of the communities at cooler temperature sites C and E (73.7% versus 59.9%; $\chi^2 = 2076$, $P < 0.0001$; Figure 2b). Species that were particularly abundant at site C (76–72 °C) included a phylotype in the *Armatimonadetes* ($\bar{p} = 15.9\%$), a distinct phylotype in the archaeal lineage pSL4 ($\bar{p} = 15.9\%$) and '*Ca.* N. yellowstonii' ($\bar{p} = 8.89\%$). Although most phylum- and class-level OTUs were represented by a single abundant species in the pyrotag dataset ($>1\%$), *Armatimonadetes* and pSL4 were exceptional in that temperature specialization was apparent (Supplementary Figure S5a Supplementary Figure S5b). The recently discovered phylum *Armatimonadetes* (Stott *et al.*, 2008) was represented primarily by three 97% OTUs, one of which was abundant at site C ($\bar{p} = 15.9\%$), and the

others which were only present at site E. pSL4 was also represented primarily by three OTUs with one dominant at high temperature sites (Supplementary Figure S6b, blue) and another dominant at lower temperature sites (Supplementary Figure S6b, purple).

Inferences on upper temperature limits of biological functions

The distribution and abundance of pyrotags can offer clues to the spatial distribution of biogeochemical functions within the spring. Two species-level lineages were identified as '*Ca.* N. yellowstonii'. The more abundant of these OTUs comprised 5–15% of all pyrotags at sites A, B and C when sampled between 80 and 72 °C and shared 99% identity with '*Ca.* N. yellowstonii' (EU239960) over the full length of the pyrotag. The less abundant OTU represented $\sim 6\%$ of pyrotags at site E (62 °C) and was 97% identical to '*Ca.* N. yellowstonii' (EU239960) over the full length of the pyrotag. Previous work has documented relatively high rates of ammonia

oxidation in GBS sediment (82 and 81 °C) and extremely high abundance of 'Ca. N. yellowstonii' (amoA $3.5–3.9 \times 10^8$ and SSU rRNA gene $6.4–9.0 \times 10^8$ copies per g of sediment) (Dodsworth et al., 2011). These results delineate an upper temperature limit of ~ 82 °C for ammonia oxidation by 'Ca. N. yellowstonii' in GBS, yet 'Ca. N. yellowstonii' has heretofore only been cultivated at temperatures up to 74 °C in the laboratory (de la Torre et al., 2008). The cause of the ~ 8 °C difference in upper temperature for laboratory growth and abundance and activity in situ remains to be elucidated.

In GBS, ammonia oxidation limits denitrification, which seems to be heterotrophic at 81 and 79 °C and is carried out at least in part by Thermus thermophilus strains that produce nitrous oxide as the terminal denitrification product (Dodsworth et al., 2011; Hedlund et al., 2011). T. thermophilus was most abundant in pyrotags at sites B (79 °C) and C (76 °C) ($\bar{p} = 2.2\%$), but was replaced by other Thermus species, primarily T. scotoductus, at lower temperatures ($\bar{p} = 9.1\%$) (Supplementary Figure S7). However, the low temperature sites likely hosted other denitrifiers as well. The temperature specialization of T. thermophilus at relatively high temperature sites in GBS, in combination with the low species richness at those sites, may drive the observed high nitrous oxide flux at high temperature in GBS sediment (Hedlund et al., 2011).

Site E (62 °C) was well below the known upper temperature limit for photosynthesis (Brock, 1967; Cox et al., 2011) and contained OTUs belonging to known photosynthetic genera. Synechococcus and Chloroflexus were each represented by at least one 97% OTU of $\sim 3.57\%$ relative abundance at site E. However, neither organism was detected at site C (76 and 72 °C), possibly due to temperature fluctuations that were common in GBS (data not shown). An OTU identified as a member of an unknown genus of Rhodospirillaceae was present at site E ($\bar{p} = 1.41\%$); however, phototrophy is paraphyletic within that family.

Abundance of novel lineages
Novel phylum- and class-level groups had highest relative abundance in high temperature sediments (sites A,B $\bar{p} = 55.64\%$), decreasing at site C ($\bar{p} = 48.90\%$), and lowest at site E ($\bar{p} = 12.41\%$) (Supplementary Figure S8). The novel bacterial lineage GAL35 and the novel archaeal lineage pSL4 dominated this trend. GAL35 has been detected in geothermal springs in YNP, Chile, Bulgaria and China, and it has never been detected in non-thermal environments (DeSantis et al., 2006). In GBS, GAL35 had highest relative abundance at site A and site B when they were sampled at 80 and 79 °C ($\bar{p} = 52.41\%$), but it was relatively abundant in all pyrotag datasets except site E. pSL4 is affiliated with the recently proposed phylum 'Candidatus Aigarchaeota', an uncultivated group defined by environmental 16S rRNA gene sequences found in a variety of thermal environments and dominant in some anaerobic hydrothermal sediments and mats (Marteinsson et al., 2001; Nunoura et al., 2005; de la Torre et al., 2008; Nunoura et al., 2010; Nunoura et al., 2011). A composite genome of a representative of 'Candidatus Aigarchaeota', 'Candidatus Caldiarchaeum subterraneum', was found to contain an uptake hydrogenase and aerobic carbon monoxide dehydrogenase, suggesting this organism could exhibit a chemolithotrophic metabolism (Nunoura et al., 2005). In GBS, pSL4 was most abundant at site A at 87 °C ($\bar{p} = 54.13\%$) and was present in at least 3.85% relative abundance in all pyrotag datasets.

Novel genus-level groups had highest relative abundance in high temperature sediments (site A,B $\bar{p} = 70.66\%$), decreasing slightly at site C ($\bar{p} = 67.58\%$), and lowest at site E ($\bar{p} = 35.32\%$) (Supplementary Figure S9). The novel archaeal lineage NAG1 and the novel bacterial lineage OS-L were the most abundant lineages unidentified at the genus level. NAG1 sequences have been obtained from One Hundred Springs Plain Spring in YNP (Kozubal et al., 2012). In GBS, NAG1 had highest relative abundance in site A at 82 °C ($\bar{p} = 20.47\%$) and was not detectable at temperatures < 79 °C. The novel bacterial lineage OS-L is a member of the recently proposed phylum Armatimonadetes, formerly known as OP10, cultivated members of which include aerobic heterotrophs (Stott et al., 2008; Lee et al., 2011; Tamaki et al., 2011). In GBS, OS-L was most prevalent at samples taken between 76 and 72 °C ($\bar{p} = 8.50\%$) and was undetectable at temperatures > 76 °C.

Conclusions

The current study underscores a need to consider water and sediment communities of geothermal springs separately, particularly in long residence time springs, which permit growth of a specific water-borne microbial community. In GBS, the segregation of nitrogen-cycle functions between sediment and water communities has been clearly documented (Dodsworth et al., 2011; Hedlund et al., 2011) and other functions are likely segregated as well. Although a number of studies have examined microbial community composition and structure along geothermal outflows, the relative depth of the pyrotag approach employed here supported a strong statistical framework and incisive analyses. We observed a strong, negative, linear relationship between temperature and sediment microbial community richness at all taxonomic levels and illustrated the strong selective force of temperature within the range that was sampled. Temperature also affects ecosystem functioning. In this study, we documented the upper temperature limit of chemolithotrophic ammonia oxidation at ~ 82 °C in situ.

728

In addition, we identified the thermal niche of *Thermus thermophilus*, which is likely responsible for high nitrous oxide flux from high temperature GBS sediments (Dodsworth *et al.*, 2011; Hedlund *et al.*, 2011). This study also revealed a high abundance of novel lineages in GBS, particularly in the highest temperature sediments. Five species-level OTUs comprised >82.17% of all pyrotags in sediment samples ≥82 °C, and three of these OTUs could not be identified at the phylum or class level (GAL35, pSL4 and NAG1). This high degree of novelty underscores the need to uncover the roles of these organisms in order to understand the biology of this ecosystem.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

References

Barns SM, Fundyga RE, Jeffries MW, Pace NR. (1994). Remarkable archaeal diversity detected in a Yellowstone National Park hot spring environment. *Proc Natl Acad Sci* **91**: 1609–1613.

Blank CE, Cady SL, Pace NR. (2002). Microbial composition of near-boiling silica-depositing thermal springs throughout Yellowstone National Park. *Appl Environ Microbiol* **68**: 5123–5135.

Bray JR, Curtis JT. (1957). An ordination of the upland forest communities of Southern Wisconsin. *Ecol Monogr* **27**: 325–349.

Brock TD. (1967). Micro-organisms adapted to high temperatures. *Nature* **214**: 882–885.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK *et al.* (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**: 335–336.

Chao A. (1984). Nonparametric estimation of the number of classes in a population. *Scand J Stat* **11**: 265–270.

Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ *et al.* (2009). The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res* **37**: D141–D145.

Costa K, Navarro J, Shock EL, Zhang C, Soukup D, Hedlund B. (2009). Microbiology and geochemistry of great boiling and mud hot springs in the United States Great Basin. *Extremophiles* **13**: 447–459.

Cox A, Shock EL, Havig JR. (2011). The transition to microbial photosynthesis in hot spring ecosystems. *Chem Geol* **280**: 344–351.

Department of Energy Joint Genome Institute, California TRof the U of (2010). DNA isolation bacterial CTAB protocol. http://my.jgi.doe.gov/general/.

de la Torre JR, Walker CB, Ingalls AE, Könneke M, Stahl DA. (2008). Cultivation of a thermophilic ammonia oxidizing archaeon synthesizing crenarchaeol. *Environ Microbiol* **10**: 810–818.

DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K *et al.* (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* **72**: 5069–5072.

Dodsworth JA, Hungate BA, Hedlund BP. (2011). Ammonia oxidation, denitrification and dissimilatory nitrate reduction to ammonium in two US Great Basin hot springs with abundant ammonia-oxidizing archaea. *Environ Microbiol* **13**: 2371–2386.

Edgar RC. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460–2461.

Engelbrektson A, Kunin V, Wrighton KC, Zvenigorodsky N, Chen F, Ochman H *et al.* (2010). Experimental factors affecting PCR-based estimates of microbial species richness and evenness. *ISME J* **4**: 642–647.

Fenchel T, King GM, Blackburn TH. (1998). *Bacterial Biogeochemistry: The Ecophysiology of Mineral Cycling*, 2nd edn. Elsevier Academic Press: San Diego, USA.

Fouke BW. (2011). Hot-spring systems geobiology: abiotic and biotic influences on travertine formation at Mammoth Hot Springs, Yellowstone National Park, USA. *Sedimentology* **58**: 170–219.

Hammer O, Harper DAT, Ryan PD. (2001). PAST: Paleontological Statistics software package for education and data analysis. *Paleontologia Electronica* **4**: 9.

Hedlund BP, McDonald AI, Lam J, Dodsworth JA, Brown JR, Hungate BP. (2011). Potential role of *Thermus thermophilus* and *T. oshimai* in high rates of nitrous oxide (N_2O) production in ~80 °C hot springs in the US Great Basin. *Geobiology* **9**: 471–480.

Holm S. (1979). A simple sequentially rejective multiple test procedure. *Scand J Stat* **6**: 65–70.

Huang Z, Hedlund BP, Wiegel J, Zhou J, Zhang CL. (2007). Molecular phylogeny of uncultivated *Crenarchaeota* in Great Basin hot springs of moderately elevated temperature. *Geomicrobiol J* **24**: 535–542.

Hugenholtz P, Pitulle C, Hershberger KL, Pace NR. (1998). Novel division level bacterial diversity in a Yellowstone hot spring. *J Bacteriol* **180**: 366–376.

Jackson CR, Langner Heiko W, Donahoe-Christiansen J, Inskeep WP, McDermott TR. (2001). Molecular analysis of microbial community structure in an arsenite-oxidizing acidic thermal spring. *Environ Microbiol* **3**: 532–542.

Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. (2008). NCBI BLAST: a better web interface. *Nucleic Acids Res* **36**: W5–W9.

Kozubal MA, Macur RE, Jay ZJ, Beam JP, Malfatti SA, Tringe SG *et al.* (2012). Microbial iron cycling in acidic geothermal springs of Yellowstone National Park: integrating molecular surveys, geochemical processes, and isolation of novel Fe-active microorganisms. *Front Microbiol* **3**: 109.

Kunin V, Hugenholtz P. (2010). PyroTagger: A fast, accurate pipeline for analysis of rRNA amplicon pyrosequence data. *The Open Journal* 1–8.

Lee KC-Y, Dunfield PF, Morgan XC, Crowe MA, Houghton KM, Vyssotski M *et al.* (2011). *Chthonomonas calidirosea* gen. nov., sp. nov., an aerobic, pigmented, thermophilic micro-organism of a novel bacterial class, *Chthonomonadetes* classis nov., of the newly described phylum *Armatimonadetes* originally designated candidate division OP10. *Int J Syst Evol Microbiol* **61**: 2482–2490.

Lefèvre CT, Abreu F, Schmidt ML, Lins U, Frankel RB, Hedlund Brian P *et al.* (2010). Moderately thermophilic magnetotactic bacteria from hot springs in Nevada. Appl. Environ. *Microbiol* **76**: 3740–3743.

Lozupone CA, Knight R. (2007). Global patterns in bacterial diversity. *Proc Natl Acad Sci USA* 11436–11440.

Macur RE, Langner HW, Kocar BD, Inskeep WP. (2004). Linking geochemical processes with microbial community analysis: successional dynamics in an arsenic-rich, acid-sulphate-chloride geothermal spring. *Geobiology* **2**: 163–177.

Marteinsson VT, Hauksdóttir S, Hobel CFV, Kristmannsdóttir H, Hreggvidsson GO, Kristjánsson JK. (2001). Phylogenetic diversity analysis of subterranean hot springs in Iceland. *Appl Environ Microbiol* **67**: 4242–4248.

Meyer-Dombard DR, Shock EL, Amend JP. (2005). Archaeal and bacterial communities in geochemically diverse hot springs of Yellowstone National Park, USA. *Geobiology* **3**: 211–227.

Meyer-Dombard DR, Swingley W, Raymond J, Havig J, Shock EL, Summons RE. (2011). Hydrothermal ecotones and streamer biofilm communities in the Lower Geyser Basin, Yellowstone National Park. *Environ Microbiol* **13**: 2216–2231.

Miller-Coleman RL, Dodsworth JA, Ross CA, Shock EL, Williams AE, Hartnett HE *et al.* (2012). Korarchaeota diversity, biogeography, and abundance in Yellowstone and Great Basin hot springs and ecological niche modeling based on machine learning. *PLOS ONE* **7**: e35964.

Miller SR, Strong AL, Jones KL, Ungerer MC. (2009). Barcoded pyrosequencing reveals shared bacterial community properties along the temperature gradients of two alkaline hot springs in Yellowstone National Park. *Appl Environ Microbiol* **75**: 4565–4572.

Nordstrom DK, Ball JW, McCleskey RB. (2005). Ground water to surface water: chemistry of thermal outflows in Yellowstone National Park. In *Geothermal Biology and Geochemistry in Yellowstone National Park* Inskeep WP, Varley J (eds). Thermal Biology Institute, Montana State University: Canada.

Nunoura T, Hirayama H, Takami H, Oida H, Nishi S, Shimamura S *et al.* (2005). Genetic and functional properties of uncultivated thermophilic crenarchaeotes from a subsurface gold mine as revealed by analysis of genome fragments. *Environ Microbiol* **7**: 1967–1984.

Nunoura T, Oida H, Nakaseama M, Kosaka A, Ohkubo SB, Kikuchi T *et al.* (2010). Archaeal diversity and distribution along thermal and geochemical gradients in hydrothermal sediments at the Yonaguni Knoll IV Hydrothermal Field in the Southern Okinawa Trough. *Appl Environ Microbiol* **76**: 1198–1211.

Nunoura T, Takaki Y, Kakuta J, Nishi S, Sugahara J, Kazama H *et al.* (2011). Insights into the evolution of *Archaea* and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. *Nucleic Acids Res* **39**: 3204–3223.

R Development Core Team (2011). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing: Vienna. http://www.R-project.org.

Reysenbach A-L, Banta A, Civello S, Daly J, Mitchell K, Lalonde S *et al.* (2005). *Aquificales* in Yellowstone National Park. In Geothermal Biology and Geochemistry in Yellowstone National Park, Inskeep WP, McDermott TR (eds). Montana State University Publications: Bozeman, MT, internal-pdf://Reysenbach 2005 Aquificales in YNP Chapter-0789304326/Reysenbach 2005 Aquificales in YNP Chapter.pdf.

Reysenbach AL, Ehringer M, Hershberger K. (2000). Microbial diversity at 83°C in Calcite Springs, Yellowstone National Park: another environment where the Aquificales and 'Korarchaeota' coexist. *Extremophiles* **4**: 61–67.

Reysenbach AL, Wickham GS, Pace NR. (1994). Phylogenetic analysis of the hyperthermophilic pink filament community in Octopus Spring, Yellowstone National Park. *Appl Environ Microbiol* **60**: 2113–2119.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB *et al.* (2009). Introducing mothur: opensource, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75**: 7537–7541.

Spear JR, Walker JJ, McCollom TM, Pace NR. (2005). Hydrogen and bioenergetics in the Yellowstone geothermal ecosystem. *Proc Natl Acad Sci USA* **102**: 2555–2560.

Stott MB, Crowe MA, Mountain BW, Smirnova AV, Hou S, Alam M *et al.* (2008). Isolation of novel bacteria, including a candidate division, from geothermal soils in New Zealand. *Environ Microbiol* **10**: 2030–2041.

Tamaki H, Tanaka Y, Matsuzawa H, Muramatsu M, Meng X-Y, Hanada S *et al.* (2011). *Armatimonas rosea* gen. nov., sp. nov., of a novel bacterial phylum, *Armatimonadetes* phyl. nov., formally called the candidate phylum OP10. *Int J Syst Evol Microbiol* **61**: 1442–1447.

Youssef N, Sheik CS, Krumholz LR, Najar FZ, Roe BA, Elshahed MS. (2009). Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental Surveys. *Appl Environ Microbiol* **75**: 5227–5236.

Zhang CL, Hedlund BP, Meng J. (2011). Diversity of archaea in terrestrial hot springs and role in ammonia oxidation. In Handbook of Molecular Microbial Ecology II, De Bruijn, FJ De (ed). John Wiley & Sons, Inc., pp 381–394. http://onlinelibrary.wiley.com/doi/10.1002/9781118010549.ch37/summary (accessed July 5, 2012).

Zhang Z, Schwartz S, Wagner L, Miller W. (2000). A greedy algorithm for aligning DNA sequences. *J Comput Biol* **7**: 203–214.

Supplementary Information accompanies the paper on The ISME Journal website (http://www.nature.com/ismej)