Supporting Information

Suzuki et al. 10.1073/pnas.1302426110

SI Materials and Methods

Sampling Sites and Sample Collections. Samples for geochemical and microbiological analysis were collected from three different ultrabasic springs: Grotto pool spring 1 (GPS1: elevation 273 m, N: 38° 37.268′ W: 123° 8.014′), Barns spring 5 (BS5: elevation 282 m, N: 38° 37.282′, W: 123° 7.987′), and Nipple spring 1 (NS1: elevation 321 m, N: 38° 37.365′ W: 123° 8.114′), and also the main creek (i.e., Austin Creek Headwaters) (elevation 282 m, N: 38° 37.282′, W: 123° 7.987′) within The Cedars (1). Water samples for biological analysis were pumped and collected with Tygon Chemical Tubing (Masterflex) or Chem-Durance Tubing (Masterflex) equipped with PFA In-Line filter holders (Advantec). The 0.22-μm in-line filters (Millipore) were used for microbial cell collection. For microbial community analysis, ∼200 L, 50 L, and 30 L of spring waters were collected at GPS1, NS1, or BS5, respectively each year. The filtered cells were frozen immediately with dry ice at each sampling site and kept at dry-ice temperature during the transportation. A calcite "skin" at the surface of BS5 was removed before water sampling but calcite slowly but continuously precipitated at the surface. The calcite can potentially fall into the BS5 spring pool during the water sampling. Therefore, we carefully set the inlet of tube at the location where water and gas discharged into the pool and suspended the membrane filter ∼2 m above the inlet to avoid collecting calcite crystals. The creek sample was collected from the main creek adjacent to the Barnes spring complex. GPS1 samples were collected by placing a clean tube directly into the spring opening.

Geochemical Analysis. A portable pH/E_h/conductivity meter (Orion) was used for the pH, E_h, conductivity, and dissolved oxygen analysis. Metrosep A Supp 5 and Metrosep C2 columns were used for anion and cation detection, respectively. Mobile phase for anion detections was 3.2 mM Na_2CO_3, 1.0 mM $NaHCO_3$, and for cation detection was 4 mM tartaric acid, 0.75 mM dipicolinic acid. Detection limits of all anions and cations with this method are 1–2 μM.

Construction of Small-Subunit rRNA Gene Clone Libraries and Phylogenetic Analyses. Microorganisms in alkaline environments may have specialized outer membranes to adapt to their environment. Therefore, we tested physical and chemical cell disruption and an UltraClean soil DNA extraction kit (MO Bio) was chosen because we sought to obtain the highest amounts of DNA upon extraction. Cell lyses methods were also examined for the further optimization and a protocol with two rounds of vortex horizontally for 2 min and heat to 70 °C for 5 min was selected. The clone library construction was performed as described previously (2). In brief, universal bacterial (U27F and U1492R) (2), archaeal [Arc4F (3) and U1492R], and eukaryal [EukF (4) and EukR (5)] PCR primers were used to amplify the region of small-subunit (16S and 18S) rRNA genes. When a PCR product was not detected with 25 cycles, cycles were increased to 35. Amplified fragments were ligated into a pGEM-T vector (Promega) and cloned into *Escherichia coli* JM109 competent cells. PCR-amplified 16S and 18S rRNA gene fragments were recovered by PCR using primers M13F and M13R (the primers targeted the pGEM-T vector sequences flanking the insertion), then determined bacterial, archaeal, and eukaryal small-subunit rRNA sequences by using 907F (2), Arc 4F (3), and EukF (4), respectively. ABI 3730xl sequencer (Life Technologies) was used for the sequence determination. When the sequences were not determined by the primers above, alternative primers reading from the vector (M13F and M13R) were used. The sequences were aligned using CLC Genomics Workbench v5.1 (CLCbio) and assigned to operational taxonomic unit (OTU; classified as an OTU, >99% cutoff). Identification of chimeric sequences was conducted using JCVI 16S/18S small subunit analysis pipeline (5) and Greengenes' chimera check with Bellerophan (v3) (6).

Phylogenetic affiliations were determined by the combination of Greengenes (6), BLAST (7), RDP (8), and SILVA (9). However, most of the 16S rRNA sequences from GPS1 were unassigned to taxa or assigned different taxa by programs. Therefore, the taxonomic assignment of each phylotype was conducted by the phylogenetic trees created with the closest relatives. Nearly the full-length of 16S rRNA gene was used and a maximum 200 of the closest relatives (with minimum identity 50%) were collected from SILVA as reference sequences for creating phylogenetic trees. The phylotype's taxon was determined based on where the phylotype was clustered in the phylogenetic tree. The 16S rRNA gene sequences were aligned with cmalign (v1.1rc1) from the Infernal package (10) and trimmed with TrimAl (v1.2rev59) (11) for the phylogenetic tree construction. The trees were created by using maximum-likelihood with RaxML (12). Phylogenetic trees including limited representative sequences (Fig. S5) were also created by RaxML and the robustness of furcated branches was supported by bootstrap values (100 replicates). A rarefaction analysis was conducted using the Analytic Rarefaction program. A multidimensional scale plot was conducted using JCVI 16S/18S small subunit analysis pipeline (13). Chao1 richness, Shannon's index, and Sørensen similarities among the bacterial communities were calculated using Estimate S (14).

Correlation Between Community Composition and Environmental Factors. Canonical correspondence analysis (XLSTAT) was performed for describing correlation between community composition and environmental factors (15). The fraction of deep and shallow groundwater was estimated as described previously (1). In brief, the fraction of the groundwater in each spring was calculated using a two-component mixing model. However, the geochemistry suggests that the creek and GPS spring fluids were the best proxies attainable to represent the shallow and deep groundwater, respectively. Therefore, for the purpose of indicating relative mixing between the two end members, the concentration Cl^- of the creek fluid was used for the Cl^- concentration of the shallow groundwater, and the concentration Cl^- of the GPS1 fluid was used for the Cl^- concentration of the deep groundwater. Correlation coefficients for the concentrations of Br^-, I^-, Li^+, Rb^+, Na^+, K^+, and Rb^+, and conductivity with f_{deep} were $r^2 \geq 0.98$, indicating that the changes in the concentrations of these ions were simply a result of physical mixing between the two water sources.

1. Morrill PL, et al. (2013) Geochemistry and geobiology of a present day serpentinization site in California: The Cedars. *Geochim Cosmochim Acta* 109:222–240.
2. Ishii S, et al. (2012) Functionally stable and phylogenetically diverse microbial enrichments from microbial fuel cells during wastewater treatment. *PLoS ONE* 7(2): e30495.
3. Giovannoni SJ, DeLong EF, Olsen GJ, Pace NR (1988) Phylogenetic group-specific oligodeoxynucleotide probes for identification of single microbial cells. *J Bacteriol* 170(2):720–726.
4. Medlin L, Elwood HJ, Stickel S, Sogin ML (1988) The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. *Gene* 71(2):491–499.

5. Moon-van der Staay SY, et al. (2000) Abundance and diversity of prymnesiophytes in the picoplankton community from the equatorial Pacific Ocean inferred from 18S rDNA sequences. *Limnol Oceanogr* 45(1):98–109.
6. DeSantis TZ, et al. (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* 72(7):5069–5072.
7. Karlin S, Altschul SF (1990) Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. *Proc Natl Acad Sci USA* 87(6):2264–2268.
8. Cole JR, et al. (2009) The Ribosomal Database Project: Improved alignments and new tools for rRNA analysis. *Nucleic Acids Res* 37(Database issue):D141–D145.
9. Quast C, et al. (2013) The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res* 41(Database issue, D1):D590–D596.
10. Nawrocki EP, Kolbe DL, Eddy SR (2009) Infernal 1.0: Inference of RNA alignments. *Bioinformatics* 25(10):1335–1337.
11. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T (2009) trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25(15):1972–1973.
12. Stamatakis A (2006) RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22(21):2688–2690.
13. Li K, Bihan M, Yooseph S, Methé BA (2012) Analyses of the microbial diversity across the human microbiome. *PLoS ONE* 7(6):e32118.
14. Colwell RK (2009) EstimateS: Statistical estimation of species richness and shared species from samples. Version 8.2. User's guide and application. Available at http://purl.oclc.org/estimates. Accessed December 10, 2011.
15. ter Braak CJF (1986) Canonical correspondence analysis: A new eigenvector technique for multivariate direct gradient analysis. *Ecology* 67(5):1167–1179.

Table S1. Diversity statistics and recovery of microbial phyla from The Cedars springs

	GPS1			NS1	BS5					Creek
					2009 May			2009 September	2011 October	
Statistics	2009 May	2009 September	2011 October	2010 April	Bottom	Surface	Rock	Bottom	Bottom	2010 April
Bacteria										
No. clones	183	139	122	109	160	69	75	176	128	68
No. OTU (99% cutoff)	14	15	15	13	23	18	7	16	30	56
Chao1 richness	15 ± 3.7	18.5 ± 7.2	15.1 ± 1.8	19.3 ± 7.6	31 ± 10.7	28.7 ± 10.2	10 ± 4.4	56.5 ± 49.1	57.2 ± 23.2	489.5 ± 279
Shannon's index	1.99	1.83	2.03	1.57	2.61	2.29	1.27	1.44	2.74	3.87
Phylum/class abundance (%)*										
Chloroflexi	33	50	37	10	17	4	—	3	16	—
Firmicutes	27	21	35	28	28	10	18	7	19	3
Alphaproteobacteria	—	—	—	—	—	4	—	1	5	9
Betaproteobacteria	—	2	—	57	23	42	59	75	28	21
Gammaproteobacteria	1	—	—	—	1	17	21	11	9	10
Deltaproteobacteria	—	—	—	1	5	4	1	1	3	12
Bacteroidetes	—	1	1	1	1	1	—	1	4	6
Cyanobacteria	—	—	—	—	4	12	—	1	2	21
Actinobacteria	—	—	—	1	1	—	—	—	—	—
Nitrospirae	—	—	1	—	1	—	—	—	2	—
Planctomycetes	—	—	—	—	—	—	—	—	—	3
Verrucomicrobia	—	—	—	—	—	—	—	—	—	1
Cand. Div. OD1	23	5	16	—	13	1	—	1	2	—
Cand. Div. NPL_UPA2	4	4	—	2	—	—	—	—	—	1
Cand. Div. OP11	—	—	—	—	—	—	—	—	—	4
Cand. Div. OP3	—	—	—	—	—	—	—	—	—	4
Cand. Div. TM7	—	—	—	—	—	—	—	—	—	1
Unclassified	12	18	11	—	7	3	—	—	6	—
Miscellaneous	—	—	—	—	—	—	—	—	2	3
Archaea										
No. clones†	42	69	45	40	45	33	0	46	40	0
No. OTU (99% cutoff)	1	1	1	1	1	1	—	1	1	—
Phylum/class abundance (%)										
Euryarchaeota	100	100	100	100	100	100	—	100	100	—
Eukarya										
No. clones†	0	0	0	0	0	25	0	0	0	NT
No. OTU (99% cutoff)	—	—	—	—	—	21	—	—	—	
Phylum/class abundance (%)										
Amoebozoa	—	—	—	—	—	20	—	—	—	
Cryptophyta	—	—	—	—	—	16	—	—	—	
Fungi	—	—	—	—	—	16	—	—	—	
Heterolobosea	—	—	—	—	—	4	—	—	—	
Metazoa	—	—	—	—	—	28	—	—	—	
Rhizaria	—	—	—	—	—	4	—	—	—	
Stramenopiles	—	—	—	—	—	8	—	—	—	
Unclassified	—	—	—	—	—	4	—	—	—	

—, not detected; NT, not tested.
*Phylum or class taxonomic classifications were based on SILVA classifications.
†0 indicates that small subunit RNA was not amplified by PCR with domain-specific primers.

Fig. S1. Sampled springs GPS1 (*A* and *B*), NS1 (*C* and *D*), BS5 (*E* and *F*), and creek (*F*) in The Cedars area. Open arrowheads indicate outlets of groundwater at GPS1 (*A*) and NS1 (*C*). Dashed arrows indicate the locations where bubbles come from the subsurface of BS5 (*E*). Filled arrow shows the sampling location of creek water (*F*). Panoramic pictures of the sampling sites (*B*, *D*, and *F*). The exact locations for the sample collections are indicated by the white squares in *B*, *D*, and *F*.

Fig. S1

Fig. S2. Rarefaction curves for the different phylotypes obtained from bacterial 16S rRNA gene clone libraries.

Fig. S2

Fig. S3. Sørensen's similarity of The Cedars bacterial communities.

Fig. S3

Fig. S4. Multidimensional scale plot comparing bacterial communities based on phylotype level taxonomy.

Fig. S4

Fig. S5. (*A*–*H*) Phylogenetic relationships among bacterial and archaeal 16S rRNA phylotypes recovered from The Cedars springs. Tree topologies are supported by bootstrap values for 100 replicates and the values greater than 70 were indicated. Red font denotes The Cedars phylotypes, purple font denotes phylotypes from other serpentinization sites, pink font denotes phylotypes from alkaline environments except for serpentinization sites, and green font denotes cultured (isolated) strains from alkaline environments. ANME, anaerobic methanotroph.

Fig. S5