

Supporting Information

Biddle *et al.* 10.1073/pnas.0709942105

SI Text

DNA Extraction Details. Sediment was separated into multiple MoBio DNA Isolation kit bead tubes (≈ 0.5 g per tube), 300 μ l of bead solution, and 50 μ l of M1 solution were added to each tube, and samples were vortexed for 5 seconds. Samples were placed at 65°C for 20 min, after which they were homogenized in a bead beater for 1 min. Samples were then centrifuged, and the remaining extraction was performed as per manufacturer protocol. The final elution was pooled with all other samples from the same horizon and precipitated by using ammonium acetate and isopropanol. Precipitations were left at -20°C overnight and then centrifuged at $16,000 \times g$ for 15 min. Supernatant was removed, and the pellet was washed with 70% ethanol. Samples were centrifuged at $16,000 \times g$ for 2 min, the ethanol solution was removed, and the sample was air-dried for 1–3 h. Samples were resuspended in small volumes of water. The ability of extracted DNA to be amplified by PCR was checked by the amplification of intergenic transcribed spacer by using both bacterial and archaeal primer sets as described (1). The community composition of the 1-mbsf amplified and original samples was additionally checked by comparison of RsaI (New England BioLabs) digestion patterns from PCR-amplified 16S rRNA products, and these were identical.

In-Silico Tests of Mock Pyrosequence Data. In the early stages of this work, we evaluated the effectiveness of analyzing archaeal genome data with short sequence lengths. The genome sequence of the high-temperature crenarchaeon *Pyrobaculum aerophilum* was randomly separated into 100-bp fragments to serve as mock pyrosequence data. We then removed all *Pyrobaculum* sequences from the nonredundant database and used this modified database as the target for BLAST to compare searches performed either with the short fragments or with full-length genes. For 20% of the comparisons, the short sequence and the corresponding full-length gene both identified the same database entry as the BLAST best hit. For an additional 35% of the short sequences, the functional identification recovered was equally correct as the identification provided by the best hit of the full gene. For an additional 1% of the short sequences tested, identification was correct but more general, e.g., ABC transporter not sugar ABC transporter, and the remaining short sequences did not have a functional identification. This test found misleading BLAST best hits from short sequences to be exceptionally rare and thus indicated that short reads can be useful in a metagenomic analysis, a result mirrored by tests of the bacterial *Bdellovibrio bacteriovorus* HD100 and *Escherichia coli* K12 genomes (2).

Sequence Analysis Details. BLAST analyses were performed against the NCBI nonredundant amino acid and nucleotide databases. BLASTX results were used to categorize the sequences into tentative taxonomic clades (Fig. 1). The taxonomic clade of the best matching sequence (minimum expectancy = 0.001) in the nonredundant database was used as an estimate of the taxonomy of the sampled organism. Clade information was determined by using the gi number to taxonomy mapping file provided by NCBI (gi_taxid_nucl.dmp), and by moving up to the “phylum” level by using the NCBI taxonomy node hierarchy file (nodes.dmp).

To compare the different samples, a crude functional annotation was performed, placing many of the sequences into a prechosen set of functional categories (Table 2). The functional

categories correspond to specific terms in the Gene Ontology database (3) and were chosen to be similar to those used in the published Sargasso sea metagenome analysis (4). Sequences were placed into functional categories by a series of mapping steps. First, HMMSEARCH (<http://hmmer.wustl.edu/>) was used to search each entry from the Pfam fs database (<http://pfam.janelia.org/>), against all our sequences. GO terms could be assigned for sequences with matches to Pfam entries included in the pfam2go mapping file (<http://www.geneontology.org/GO.indices.shtml>). Where possible, these GO terms were then categorized as falling under one of the more general GO terms chosen for our functional annotation. This was achieved by repeatedly moving up from child to parent GO terms using the GO ontology OBO file v1.2 (dated December 13, 2006). Many sequences with a Pfam match did not fall under any of these categories and were not included in the annotation.

Ribosomal protein coding genes were identified in the metagenome by matches to Pfam entries. Twenty-four Pfam entries (Table S1) representing 24 ribosomal proteins were chosen from a list of 31 ribosomal protein genes that have previously been identified as good phylogenetic markers (5, 6). The 24 Pfam entries were chosen based on a clear and simple correspondence between the Pfam entries and the NCBI COG entries used by Ciccarelli *et al.* (5) and without prior knowledge of the distribution of matches throughout the samples.

Additionally, sequences were analyzed by BLASTN analysis against a database of 77,957 small-subunit ribosomal genes downloaded from the “greengenes” server (<http://greengenes.lbl.gov/cgi-bin/nph-index.cgi>; downloaded April 5, 2006; includes 110,979,975 nucleotides). Matches with an expectancy value $> 1 \times 10^{-6}$ were accepted. Queries that matched the database were extracted from the large sequence output and were re-analyzed by BLASTN analysis online (www.ncbi.nlm.nih.gov/BLAST/). These results were screened for percentage of identity and length of match, and those with a match length < 49 nt were removed. The closest match was recorded by gi number, taxonomy, percent of match, and length of match (Tables S2–S6).

Quantitative PCR Details. Controls were created for qPCR analysis by PCR amplifying DNA from 1 mbsf with the previously described primers. The product was then cloned by using the TOPO-TA kit (Invitrogen). Clone libraries were analyzed, and plasmid inserts were sequenced. It was determined that the archaeal primers amplified only low-temperature Crenarchaeota, as designed (7). The bacterial primers, however, amplified Archaea and Eukaryotes in addition to Bacteria. It was determined that 9% of the amplification from the bacterial primer set can be due to nonbacterial products. An archaeal clone and a bacterial clone were chosen for use as quantitative PCR controls. Each plasmid stock was measured by the PicoGreen reagent and diluted to 1 ng/ μ l. Eight more 1:10 dilutions were made. Each dilution was used for qPCR, and a standard curve was created to relate threshold cycle to the expected number of ribosomal operons present, based on the weight of plasmid added to the reaction.

Deep sediment-extracted DNAs (original samples, or non-Repli-G amplified) were diluted to concentrations of 0.01 ng and 0.001 ng/ μ l. Each sample was applied in triplicate for each dilution. Bacterial and archaeal rRNA genes in Repli-G-amplified DNAs were quantitated by the same protocol; however, a 1:1,000 dilution of the Repli-G product was measured by Pico-Green, and 1 μ l of this dilution was used as DNA input into

the qPCR. Negative controls with no DNA added were included on each plate. The following reactions were prepared: 1 μ l of DNA, 12.5 μ l of SYBR green PCR Master Mix (Applied Biosystems), 1.125 μ l of 20 μ M forward primer, 1.125 μ l of 20 μ M reverse primer, 9.25 μ l of water. Plates were loaded into an

Applied Biosystems 7300 Real Time PCR System and run with the following cycle: 10 min at 95°C; 40 cycles of 30 sec at 95°C, 1 min at 60°C, 1 min at 72°C; followed by a dissociation cycle to measure dissociation curves.

1. Biddle JF, House CH, Brenchley JE (2005) Enrichment and cultivation of microorganisms from sediment from the slope of the Peru Trench (ODP Site 1230). *Proc ODP Sci Results* 201:107.
2. Huson DH, Auch A, Qi J, Schuster SC (2007) MEGAN analysis of metagenomic data. *Genome Res* 17:377.
3. The Gene Ontology Consortium (2000) Gene ontology: Tool for the unification of biology. *Nat Genet* 25:25–29.
4. Venter JC, et al. (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 304:66–74.
5. Ciccarelli FD, et al. (2006) Toward automatic reconstruction of a highly resolved tree of life. *Science* 311:1283.
6. Von Mering C, et al. (2007) Quantitative phylogenetic assessment of microbial communities in diverse environments. *Science* 315:1126.
7. Ochsenreiter T, Selesi D, Bonch-Ozmolovskaya L, Quaiser A, Schleper, C (2003) Diversity and abundance of Crenarchaeota in terrestrial habitats studied by 16S RNA surveys and real time PCR. *Environ Microbiol* 5:787–97.

Table S1. Twenty-four Pfam entries represented in the distribution of ribosomal proteins (Fig. 2)

Ribosomal protein	Pfam
Predicted GTPase	PF02824
Arginyl tRNA synthetase	PF03485
Ribosomal protein S12	PF00164
Ribosomal protein S2	PF00318
Ribosomal protein L1p/L10e family	PF00687
Ribosomal protein L3	PF00297
Ribosomal protein L5	PF00281
Ribosomal protein S11	PF00411
Ribosomal protein S9	PF00380
Seryl tRNA synthetase	PF02403
Ribosomal protein L16/L10E	PF00252
Preprotein translocase subunit SecY	PF00344
Ribosomal_S7	PF00177
Ribosomal.L11	PF00298
Ribosomal.L22	PF00237
Ribosomal.L14	PF00238
Ribosomal.S8	PF00410
Ribosomal.S5	PF00333
Ribosomal.S13	PF00416
Ribosomal.L13	PF00572
Ribosomal.S15	PF00312
Ribosomal.S17	PF00366
Ribosomal.L15	PF01305
Ribosomal_S4	PF00163

Table S2. Results from 1H1 amplified, 1 mbsf

<u>Site</u>	<u>Query</u>	<u>GenBank match</u>	<u>ODP clone matched</u>	<u>Other clone matched</u>	<u>Bacteria/Archaea</u>	<u>% Identity</u>	<u>Length of match</u>
1H1amp	204522_3625_2287	DQ301998	1H5_H12		archaea	106	99
1H1amp	046886_2167_0677	AM072598		280NF15	bacteria	100	65
1H1amp	058142_2554_3018	AY592773		Napoli-48-55	bacteria	97	108
1H1amp	058665_4017_1566	AY093464.1		MB-A2-110	bacteria	100	118
1H1amp	090305_2605_0840	AY193140.1		OD1 bacterium clone Bo16	bacteria	96	108
1H1amp	116786_2154_3291	DQ252423.1		JRCBIII16	bacteria	94	97
1H1amp	120523_2319_3368	L15624		Mycobacterium xenopi	bacteria	74	97
1H1amp	122123_3353_4031	AB114312		MBA02	bacteria	53	100
1H1amp	133149_2602_2189	AY568889		JH12_C50	bacteria	111	100
1H1amp	146006_2180_3211	AY592728		Napoli-38-61	bacteria	102	98
1H1amp	148132_2552_3229	DQ252423		JRCBIII16	bacteria	108	94
1H1amp	161847_3882_3190	AJ012756		Mycobacterium botniense	bacteria	74	95
1H1amp	165707_2708_1325	DQ252423		JRCBIII16	bacteria	99	93
1H1amp	196047_2347_3447	DQ252423		JRCBIII16	bacteria	96	94
1H1amp	202100_3061_0066	AY093458		MB-A2-103	bacteria	56	91
1H1amp	143470_3962_1875	AY191356	ODP1176A6H_5_B		bacteria	108	97
1H1amp	161111_2292_2734	AB177050	ODP1227B18.04		bacteria	105	100
1H1amp	126218_3142_1286	AB177320	ODP1251B17.11		bacteria	112	96

Table S3. Results from 1H1 original, 1 mbsf

Site	Query	GenBank match	ODP clone matched	Other clone matched	Bacteria/Archaea	% Identity	Length of match
1h1ori	077076_1373_0467	DQ301985		archaeon clone 1H5_E02	archaea	100	111
1h1ori	076814_0376_1756	AY835427		archaeon clone 4E09	archaea	90	116
1h1ori	170799_1830_3529	AF419648		AT_R003	archaea	94	72
1h1ori	064689_0775_1449	AY454710		eurychaeote clone E_C09	archaea	99	106
1h1ori	168259_1770_1349	AY769057		FE2ArchBot11	archaea	98	108
1h1ori	138323_0456_3000	AB237760		HDBW-WA27	archaea	98	107
1h1ori	149434_0413_2418	AB237760		HDBW-WA27	archaea	94	89
1h1ori	095023_1688_3329	AY592522		Napoli-3A-11	archaea	100	106
1h1ori	077493_1584_0458	AY592540		Napoli-3A-34	archaea	100	108
1h1ori	087953_0998_0402	AB213094		Papm3A35	archaea	100	99
1h1ori	075993_0240_1016	AE006720		Sulfolobus solfataricus P2	archaea	100	59
1h1ori	121595_0346_2384	AM086136		c5LKS53	bacteria	92	91
1h1ori	169388_1260_2001	AF146242		CLEAR-19	bacteria	98	62
1h1ori	171961_1096_3053	AF146242		CLEAR-19	bacteria	98	62
1h1ori	173479_1291_2466	AF146242		CLEAR-19	bacteria	93	96
1h1ori	138276_0345_2977	AY551086		Clostridium sp. 17cr1	bacteria	96	63
1h1ori	133633_1460_2944	CP000027		Dehalococcoides ethenogenes	bacteria	89	59
1h1ori	065185_0714_0897	AY710904		delta proteobacterium clone SIMO-1464	bacteria	97	116
1h1ori	156932_0541_1546	AY869670		FS117-47B-02	bacteria	95	115
1h1ori	114929_1050_2098	AY704394		FS118-62B-02	bacteria	93	105
1h1ori	155688_1135_0167	AY280660		G4-23	bacteria	98	94
1h1ori	093319_0896_1132	CP000148		Geobacter metallireducens GS-15	bacteria	89	74
1h1ori	133653_0458_1898	BA000045		Gloeobacter violaceus	bacteria	96	53
1h1ori	121509_0983_1898	AF424166		MERTZ_21CM_129	bacteria	98	98
1h1ori	047644_0979_0426	AF424184		MERTZ_2CM_30	bacteria	98	116
1h1ori	133097_0967_1293	AY356377		N3-40	bacteria	98	74
1h1ori	092053_0357_1084	AB094814		OHKB2.40	bacteria	99	106
1h1ori	056187_1948_1342	AY822427		planctomycete clone C5T12	bacteria	100	51
1h1ori	068420_0095_1660	AF420341		proteobacterium isolate a2b027	bacteria	93	117
1h1ori	105759_1912_3799	AY710904		SIMO-1464	bacteria	97	114
1h1ori	054776_0866_2353	DQ131907		Thermotoga sp. G2	bacteria	96	74
1h1ori	081059_1362_2831	DQ131907		Thermotoga sp. G2	bacteria	97	74
1h1ori	053011_0214_2896	AY164454		ULIB-91	bacteria	95	90
1h1ori	037464_0360_3377	AB177336	ODP1251B3.24		bacteria	97	109
1h1ori	062268_1024_3243	AB177336	ODP1251B3.24		bacteria	97	109
1h1ori	087084_1855_1560	AB177336	ODP1251B3.24		bacteria	97	109
1h1ori	101550_1125_1873	AB177189	ODP1230B3.02		bacteria	95	112
1h1ori	140596_1767_3073	AB177192	ODP1230B3.18		bacteria	100	118
1h1ori	154688_1662_3037	AB177209	ODP1230B33.10		bacteria	97	118

Table S4. Results from 3H2, 16 mbsf

Site	Query	GenBank match	ODP clone matched	Other clone matched	Bacteria/Archaea	% Identity	Length of match
12293H2	146111_0601_3521	DQ302014.1	12H3_ar23		archaea	97	119
12293H2	208315_1432_0450	DQ302014.1	12H3_ar23		archaea	98	108
12293H2	191932_0236_1699	AJ867793.1	42-AA7		archaea	93	108
12293H2	197732_0769_1132	DQ302020.1	4H3_ar12		archaea	94	90
12293H2	048400_0576_3988	DQ302027.1	4H3_ar42		archaea	100	49
12293H2	156868_1244_0710	AJ867784.1	6-AA2		archaea	97	127
12293H2	178786_1176_1578	AJ867790.1	86-AA1		archaea	91	116
12293H2	040374_1002_3496	AJ867806.1	clone 6-AA6		archaea	98	117
12293H2	034905_0118_1732	AJ867805.1	clone 6-AA7		archaea	99	118
12293H2	091113_0998_1543	AB177027.1	ODP1227A3.14		archaea	97	85
12293H2	110090_1108_3443	AB177086.1	ODP1230A10.04		archaea	93	73
12293H2	234857_1562_0989	AB177086.1	ODP1230A10.04		archaea	95	60
12293H2	146811_0607_3599	AB177095.1	ODP1230A13.29		archaea	94	117
12293H2	040025_0452_0798	DQ278137.1		crenarchaeote clone RotD-9500	archaea	100	56
12293H2	051105_1259_3846	AY454621.1		crenarchaeote clone E_E11	archaea	99	104
12293H2	056004_0904_2471	AY592006.1		clone Kazan-2A-29/BC19-2A-29	archaea	98	107
12293H2	057808_0325_0285	AB237758.1		archaeon HDBW-WA25	archaea	90	110
12293H2	065070_1407_0534	AY454603.1		crenarchaeote clone EI_H08	archaea	98	99
12293H2	099208_0862_3251	AF119129.1		archaeon CRA9-27cm	archaea	97	100
12293H2	100140_0699_1423	AY454621.1		crenarchaeote clone E_E11	archaea	97	105
12293H2	100445_1494_2832	AY454621.1		crenarchaeote clone E_E11	archaea	100	104
12293H2	107776_1800_2578	DQ188819.1		crenarchaeote clone TCren8	archaea	100	50
12293H2	112189_0803_3515	AY454621.1		crenarchaeote clone E_E11	archaea	100	104
12293H2	114748_0530_0834	AF419653.1		archaeon AT_R021	archaea	91	111
12293H2	120464_1525_1837	AB237758.1		archaeon HDBW-WA25	archaea	96	55
12293H2	136199_0874_0769	AY454621.1		crenarchaeote clone E_E11	archaea	100	104
12293H2	137937_1383_3780	AF419653.1		archaeon AT_R021	archaea	88	117
12293H2	159879_1072_2631	AY592472.1		Napoli-2A-07	archaea	97	103
12293H2	177073_0508_0461	AB237759.1		archaeon HDBW-WA26	archaea	95	73
12293H2	178686_0557_2844	AB237745.1		HDBW-WA12	archaea	96	57
12293H2	184156_1611_3135	AY592006.1		Kazan-2A-29/BC19-2A-29	archaea	98	97
12293H2	185341_1676_1221	AB007303.1		pOWA133	archaea	93	60
12293H2	224936_1300_1386	AY454621.1		crenarchaeote clone E_E11	archaea	94	58
12293H2	160004_0639_0239	AB177042.1	ODP1227B1.19		bacteria	100	78
12293H2	137884_1075_3784	AB177053.1	ODP1227B18.18		bacteria	95	101
12293H2	120370_1855_0526	AB177059.1	ODP1227B19.11		bacteria	98	65
12293H2	117892_1614_2662	AB177068.1	ODP1227B20.01		bacteria	97	93
12293H2	144138_1844_1289	AB177068.1	ODP1227B20.01		bacteria	90	86
12293H2	168556_0420_2898	AB177126.1	ODP1230B1.01		bacteria	100	96
12293H2	086365_1024_0653	AB177135.1	ODP1230B10.04		bacteria	95	102
12293H2	103271_0109_0937	AB177138.1	ODP1230B10.12		bacteria	95	121
12293H2	168615_0812_2712	AB177164.1	ODP1230B2.02		bacteria	93	114
12293H2	020135_0300_3554	AB177167.1	ODP1230B2.29		bacteria	95	87
12293H2	168048_0298_0587	AB177182.1	ODP1230B25.18		bacteria	95	109
12293H2	063739_0153_2592	AB177183.1	ODP1230B25.22		bacteria	97	113
12293H2	119759_0966_3244	AB177183.1	ODP1230B25.22		bacteria	92	55
12293H2	030048_0622_2838	AB177186.1	ODP1230B27.24		bacteria	92	96
12293H2	012210_1149_0599	AB177188.1	ODP1230B28.22		bacteria	96	78
12293H2	209665_0770_3148	AB177188.1	ODP1230B28.22		bacteria	94	90
12293H2	208772_0623_0912	AB177209.1	ODP1230B33.10		bacteria	95	116
12293H2	176457_0843_1840	AB177210.1	ODP1230B33.19		bacteria	91	119
12293H2	181593_0530_0655	AB177210.1	ODP1230B33.19		bacteria	93	111
12293H2	091735_0475_0249	AF407406.1		bacterium clone RA13C7	bacteria	98	68
12293H2	097415_0391_2765	AY484713.1		bacterium clone 70116	bacteria	100	59
12293H2	203559_0520_4008	AJ306742.1		SHA-20	bacteria	93	49
12293H2	223947_1350_3133	AF392750.1		clone LBF4	bacteria	81	126
12293H2	234645_1132_2114	AY093456.1		clone MB-A2-101	bacteria	90	108

Site, the sample being analyzed; Query, the code for the sequence read; GenBank match, most homologous sequence in the database; ODP clone matched, name of clone from ODP sediment study if it is most homologous sequence matched; Other clone matched, name of clone if not from ODP study; Bacteria/Archaea, domain of match; % Identity, percentage of exactly matching nucleotides; Length of match, number of nucleotides matched per sequence read.

Table S5. Results from 4H5, 32 mbsf

Site	Query	GenBank match	ODP clone matched	Other clone matched	Bacteria/Archaea	% Identity	Length of match
12294H5	094755_3915_0560	DQ302017.1	12H3_ar19		archaea	99	105
12294H5	107751_3219_1740	DQ302017.1	12H3_ar19		archaea	99	125
12294H5	007653_3345_3356	DQ301980.1	1H5_C08		archaea	99	130
12294H5	016005_3658_0254	AJ867795.1	30-AC4		archaea	100	115
12294H5	221502_3766_3423	AJ867795.1	30-AC4		archaea	96	112
12294H5	162700_3039_3139	AJ867792.1	42-AB6		archaea	100	116
12294H5	063435_3672_2350	DQ302021.1	4H3_ar13		archaea	100	68
12294H5	213230_3978_2137	AJ867806.1	6-AA6		archaea	94	71
12294H5	060544_3797_1437	AJ867803.1	6-AB7		archaea	93	99
12294H5	062883_2700_2243	AJ867803.1	6-AB7		archaea	100	96
12294H5	075592_2559_0847	AJ867803.1	6-AB7		archaea	98	104
12294H5	086497_4046_1916	AJ867803.1	6-AB7		archaea	100	72
12294H5	181725_2216_1211	AJ867803.1	6-AB7		archaea	97	111
12294H5	046601_2558_1657	AJ867790.1	86-AA1		archaea	98	80
12294H5	052645_2526_2001	AJ867790.1	86-AA1		archaea	100	80
12294H5	215786_3669_1740	AJ867789.1	86-AA3		archaea	90	111
12294H5	180503_2556_3529	AJ867778.1	86-AC6		archaea	96	118
12294H5	238329_2145_2503	AB177016.1	ODP1227A19.02		archaea	99	120
12294H5	066542_3665_2356	AB177018.1	ODP1227A19.04		archaea	95	106
12294H5	169644_3203_0347	AB177027.1	ODP1227A3.14		archaea	93	72
12294H5	215640_2364_1516	AB177028.1	ODP1227A3.21		archaea	98	100
12294H5	039811_2903_0409	AY454621.1		clone E_E11	archaea	98	124
12294H5	042870_3694_4002	AY861963.1		clone OPPD030	archaea	100	98
12294H5	054895_3806_0803	AY454621.1		clone E_E11	archaea	98	98
12294H5	062389_2193_2025	M32222.1		M.fervidus	archaea	90	77
12294H5	064965_3436_2894	X14835.1		Thermofilum pendens	archaea	100	49
12294H5	081187_3547_3528	AY454616.1		clone E_E03	archaea	97	108
12294H5	104663_3684_0321	AY454616.1		clone E_E03	archaea	98	76
12294H5	108846_3080_2623	X14835.1		Thermofilum pendens	archaea	97	49
12294H5	118142_2883_1319	DQ128270.1		CCSD_DF2450_A14	archaea	93	61
12294H5	143190_2715_0980	AB113634.1		HAuD-LA42	archaea	98	98
12294H5	168019_2555_2396	AB109878.1		pMLA-1	archaea	96	81
12294H5	203168_2960_1142	AY591992.1		Kazan-2A-14/BC19-2A-14	archaea	96	97
12294H5	203425_2539_2752	AY592006.1		Kazan-2A-29/BC19-2A-29	archaea	90	66
12294H5	222809_2962_1545	DQ188818.1		TCren7	archaea	95	89
12294H5	255947_4011_1844	AY861925.1		OPPC060	archaea	98	55
12294H5	154338_3997_3668	AJ867595.1	86-B19		bacteria	98	110
12294H5	007673_3338_2071	AB177059.1	ODP1227B19.11		bacteria	92	110
12294H5	094273_3419_0567	AB177073.1	ODP1227B3.08		bacteria	94	109
12294H5	093003_3614_0162	AB177080.1	ODP1227B6.09		bacteria	98	99
12294H5	240198_2805_1419	AB177132.1	ODP1230B1.19		bacteria	92	54
12294H5	019831_2531_1827	AB177135.1	ODP1230B10.04		bacteria	90	104
12294H5	063813_2726_2238	AB177135.1	ODP1230B10.04		bacteria	96	96
12294H5	189045_2894_1725	AB177135.1	ODP1230B10.04		bacteria	91	93
12294H5	193671_3659_0753	AB177135.1	ODP1230B10.04		bacteria	95	95
12294H5	203889_3968_1197	AB177135.1	ODP1230B10.04		bacteria	95	99
12294H5	248027_2559_3623	AB177135.1	ODP1230B10.04		bacteria	99	110
12294H5	266070_2625_1900	AB177135.1	ODP1230B10.04		bacteria	88	90
12294H5	016464_2577_2188	AB177167.1	ODP1230B2.29		bacteria	96	99
12294H5	066541_2875_2673	AB177167.1	ODP1230B2.29		bacteria	97	102
12294H5	160864_3054_1830	AB177167.1	ODP1230B2.29		bacteria	92	114
12294H5	172707_2970_0084	AB177167.1	ODP1230B2.29		bacteria	96	99
12294H5	134452_3566_3983	AB177188.1	ODP1230B28.22		bacteria	94	95
12294H5	045511_2144_2917	AB177189.1	ODP1230B3.02		bacteria	97	79
12294H5	123702_2205_3476	AB177209.1	ODP1230B33.10		bacteria	99	111
12294H5	027105_2403_3928	AB177210.1	ODP1230B33.19		bacteria	95	96
12294H5	039095_2177_0873	AB177210.1	ODP1230B33.19		bacteria	95	92
12294H5	051441_2388_2941	AB177210.1	ODP1230B33.19		bacteria	98	65
12294H5	071426_3196_3018	AB177217.1	ODP1230B4.21		bacteria	100	54
12294H5	197024_2966_3239	AB177243.1	ODP1244B22.13		bacteria	92	91
12294H5	124613_3921_0733	DQ128649.1		HSB NT21_E02	bacteria	100	49
12294H5	153880_3440_1315	CP000232.1		Moorella thermoacetica	bacteria	98	50
12294H5	182552_3126_1082	AY616128.1		Vibrio cholerae	bacteria	93	49
12294H5	191470_2096_0180	CP000027.1		Dehalococcoides ethenogenes	bacteria	93	91
12294H5	215908_3980_1282	AJ582199.1		Gitt-GS-113	bacteria	88	92
12294H5	224513_2743_3163	AJ880487.1		340N52	bacteria	93	63

Table S6. Results from 7H1, 50 mbsf

Site	Query	GenBank match	ODP clone matched	Other clone matched	Bacteria/Archaea	Identity, %	Length of match
7H1	024283_0132_2107	AB109884.1		Uncultured archaeon pMLA-7	Archaea	100	126
7H1	039753_0132_1870	AJ496176.1		Uncultured crenarchaeote	Archaea	95	66
7H1	067320_0890_1464	AY454609.1		Uncultured crenarchaeote	Archaea	97	94
7H1	076398_1185_1366	EF104112.1		archaeon clone SCS-QB5-A129	Archaea	100	111
7H1	079491_0904_1623	DQ841220.1		Uncultured archaeon clone MOB4-14	Archaea	90	114
7H1	082105_1483_3818	DQ841220.1		Uncultured archaeon clone MOB4-14	Archaea	100	109
7H1	086173_0943_1801	AB109884.1		Uncultured archaeon pMLA-7	Archaea	97	87
7H1	089179_0054_1291	AY592472.1		Uncultured archaeon clone Napoli-2A-07	Archaea	98	100
7H1	101248_0402_1161	AB213061.1		Uncultured archaeon Fapm1aA37	Archaea	93	113
7H1	116577_1972_3184	EF208719.1		Uncultured archaeon clone CI75 cm.2.64	Archaea	98	53
7H1	116704_0638_0612	DQ522920.1		Uncultured archaeon clone SBAK-mid-18	Archaea	97	76
7H1	117396_1579_0488	AY592472.1		Uncultured archaeon clone Napoli-2A-07	Archaea	98	109
7H1	120870_0897_3807	AY592472.1		Uncultured archaeon clone Napoli-2A-07	Archaea	94	103
7H1	125162_1528_2068	DQ841220.1		Uncultured archaeon clone MOB4-14	Archaea	100	108
7H1	133507_1661_2036	AB109884.1		Uncultured archaeon pMLA-7	Archaea	99	115
7H1	020163_1322_3039	AY312056.1		Paracoccus sp. OI18	Bacteria	91	97
7H1	084986_1303_1320	DQ223084.1		Uncultured bacterium clone W2	Bacteria	98	98

Other Supporting Information Files

[Dataset S1 \(XLS\)](#)

[Dataset S2 \(XLS\)](#)