Supporting Information

Wang et al. 10.1073/pnas.0810418106

SI Methods

DNA Isolation, Amplification, Labeling, and Microarray Hybridization. The genomic DNA was isolated from the chimneys (1) and amplified by using 50 ng of DNA as described previously (2). The amplified DNA was fluorescently labeled with cy5 dye (GE Healthcare) by using random primers and Klenow (large fragment of DNA polymerase I; Invitrogen) according to the manufacturer's directions. The microarray hybridization was conducted at 50 °C in the presence of 50% formamide, as previously described (3). All hybridization experiments were carried out in duplicate.

Microarray Scanning and Data Processing. A ScanArray 500 microarray system (PerkinElmer) was used for scanning the microarray slides. Scanned images were processed by using ImaGene, version 5.0 (Biodiscovery), as described previously (4). The signal-to noise ratio (SNR) was calculated based on the formula: SNR = (signal intensity − background)/standard deviation of background. Spots with SNR greater or equal to 3 were regarded as positive. Data processing, such as outlier removal, normalization, and poor spot removal, was carried out as previously described (2).

Constructing 16S rRNA, *mcrA*, *cbbL*, and *cbbM* Gene Clone Libraries. The archaeal and bacterial 16S rRNA gene clone libraries were constructed by using universal primer sets Arch 27F/958R and Bac21F/1492R, respectively. RubisCO Form I *cbbL* genes were amplified with primers 595F/1387R with an 800-bp fragment, whereas Form II *cbbM* genes were amplified with cbbM663F/cbbM1033R, which yields a 400-bp fragment (5). The *mcrA* gene was amplified with primers ME1 and ME2 (6). The PCR products were purified on a 1% agarose gel, extracted with a gel-extraction kit (Omega Bio-Tek Inc.). Afterward, the purified DNA products were ligated with the pMD18-T vector (Takara) and transformed to competent cells of *Escherichia coli* DH-5α according to the manufacturer's instructions.

Quantitative PCR. Serial dilutions of positive control DNA were used as calibration standards for the quantitative real-time PCR. The positive control DNA extracts were amplified with the *cbbM* primers (328 bp) (7), with sulfide-chimney DNA from this study as the template. The resulting amplicons were purified with a gel-extraction kit (Omega Bio-Tek Inc.) as recommended by the manufacturer, and were cloned to the pMD18-T vector. After reamplification with vector-specific primers according to the manufacturer's instructions, the PCR products were purified as described above. The PCR products obtained were then quantified with the NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies). The measured DNA amount was converted to target molecule numbers per microliter, and the *cbbM* standards were adjusted to 10^{10} target molecules per microliter by serial dilution. PCR and monitoring of SYBR Green I fluorogenic probe signals were performed by using an ABI Prism 7500 (Applied Biosystems). Copy numbers of the *cbbL* gene were determined by quantitative competitive PCR (QC-PCR). The liner plasmid pmD-18T Δ150bp was used as the competitive template DNA. PCR products were separated by electrophoresis using 2.0% (wt/vol) agarose with TBE buffer (90 mM Tris, 90 mM boric acid, and 2 mM Na_2-EDTA; pH 8.0) and were stained with ethidium bromide. The copy number of the target gene was estimated by considering the band intensity and sizes of target and standard DNA and the initial concentration of *cbbL* present in chimney DNA; this concentration was subsequently adjusted to obtain a gene copy value on a wet-weight basis.

1. Zhou J, Bruns MA, Tiedje JM (1996) DNA recovery from soils of diverse composition. *Appl Environ Microbiol* 62:316–322.
2. Wu L, Liu X, Schadt CW, Zhou J (2006) Microarray-based analysis of subnanogram quantities of microbial community DNAs by using whole-community genome amplification. *Appl Environ Microbiol* 72:4931–4941.
3. Yergeau E, Kang S, He Z, Zhou J, Kowalchuk GA (2007) Functional microarray analysis of nitrogen and carbon cycling genes across an Antarctic latitudinal transect. *Isme J* 1:163–179.
4. Wu L, et al. (2004) Development and evaluation of microarray-based whole-genome hybridization for detection of microorganisms within the context of environmental applications. *Environ Sci Technol* 38:6775–6782
5. Elsaied H, Naganuma T (2001) Phylogenetic diversity of ribulose-1,5-bisphosphate carboxylase/oxygenase large-subunit genes from deep-sea microorganisms. *Appl Environ Microbiol* 67:1751–1765.
6. Hales BA, et al. (1996) Isolation and identification of methanogen-specific DNA from blanket bog peat by PCR amplification and sequence analysis. *Appl Environ Microbiol* 62:668–675.
7. Campbell BJ, Stein JL, Cary SC (2003) Evidence of chemolithoautotrophy in the bacterial community associated with Alvinella pompejana, a hydrothermal vent polychaete. *Appl Environ Microbiol* 69:5070–5078.

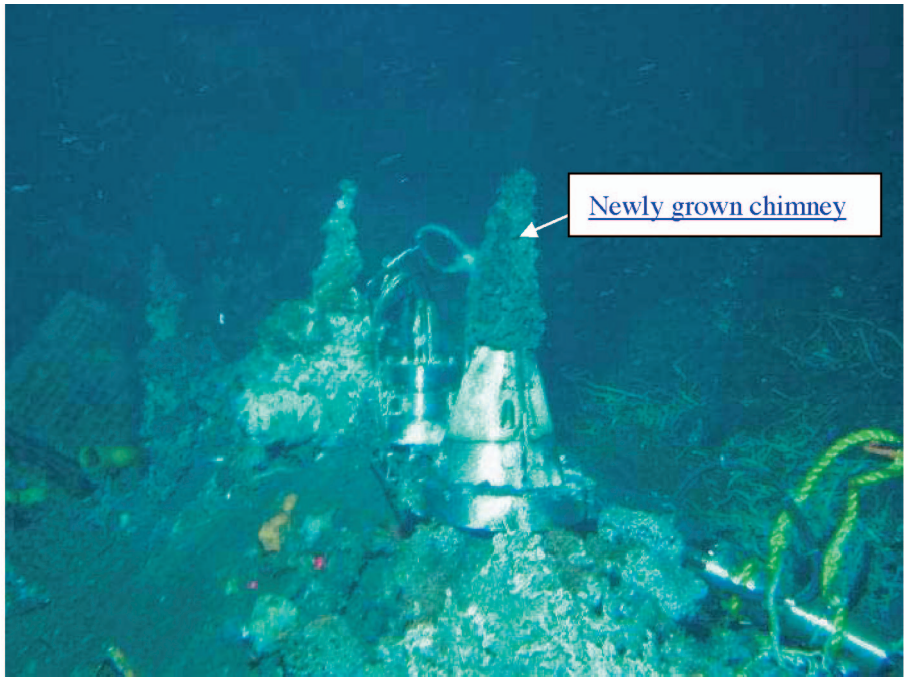


Fig. S1. A photo showing a newly grown chimney from the opening of the stainless steel cap deployed on top of a chimney venting at ≈316 °C at Main Endeavour, Juan de Fuca Ridge.

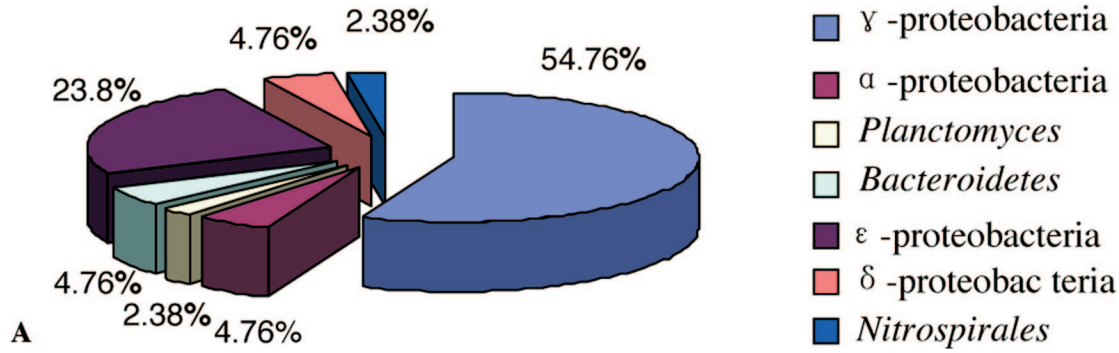
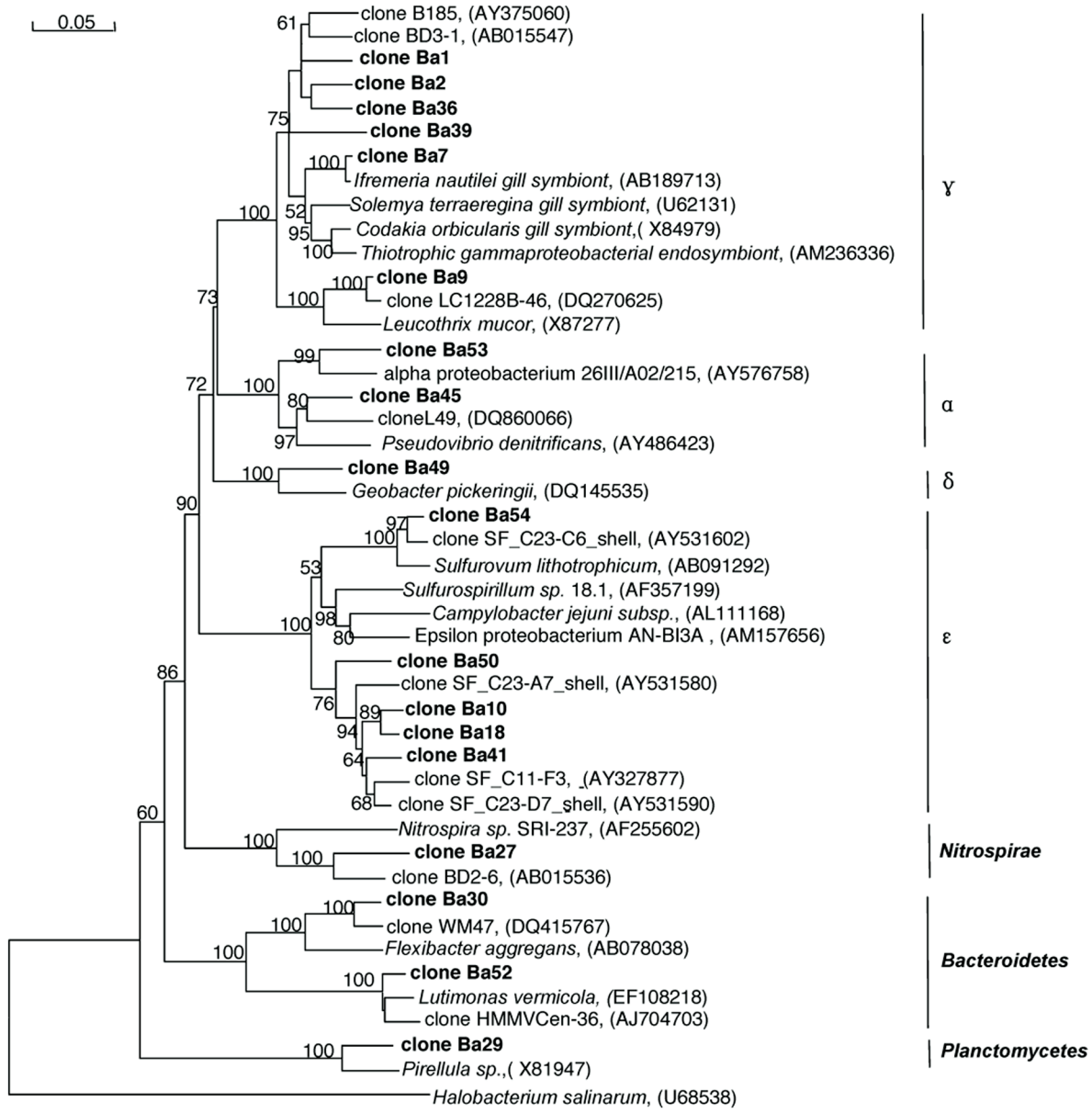
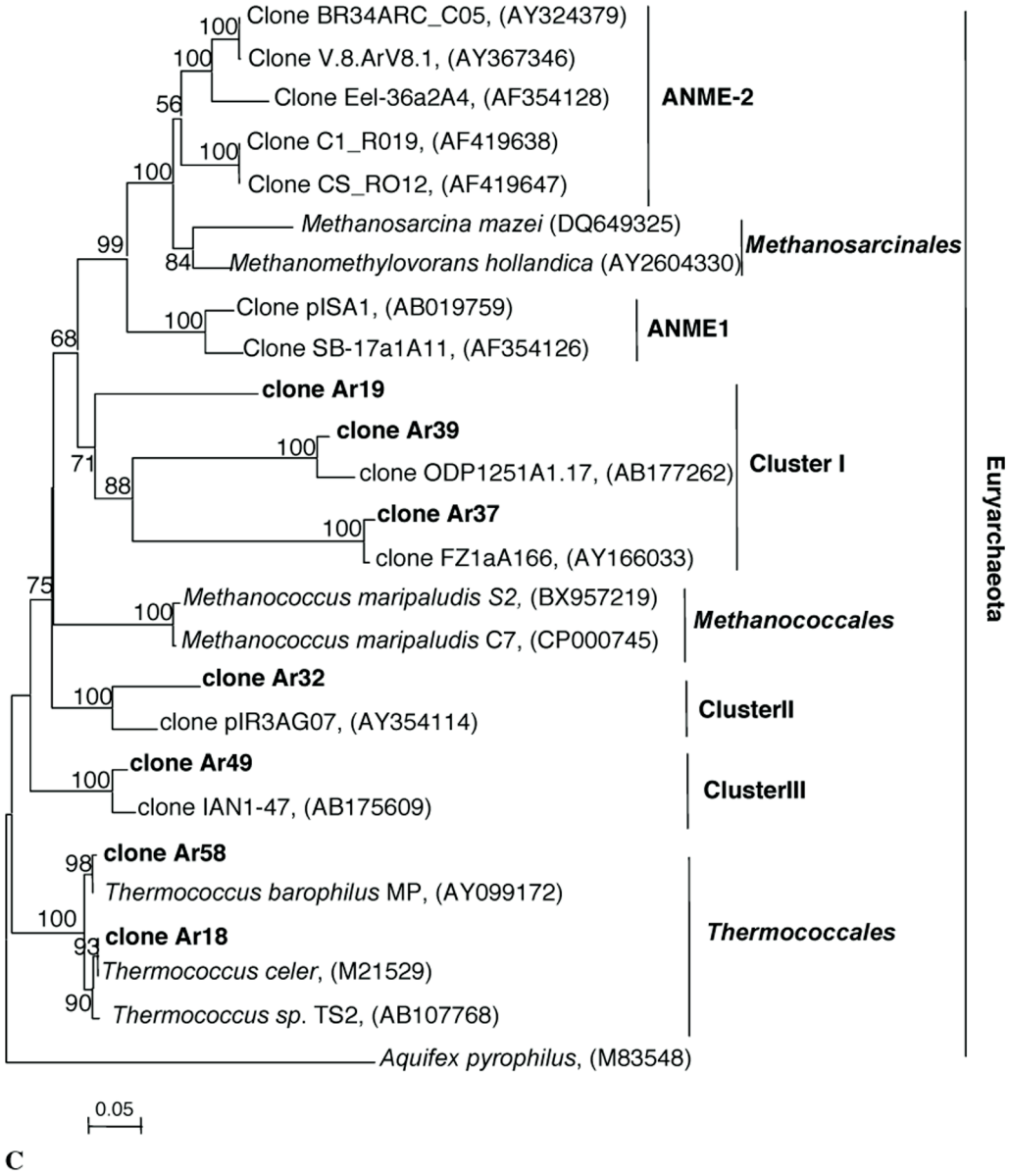


Fig. S2. Microbial community diversity of 16S rRNA gene analysis in 4143-1. (*A*) Proportion of different bacterial divisions within the bacterial community of 4143-1. The percentage was calculated by dividing the number of OTUs in the bacterial division with the total number of sequenced bacterial 16S rRNA gene clones. (B) Phylogenetic tree constructed from bacterial 16S rRNA gene sequences. The phylogenetic relationships among the retrieved bacterial 16S rRNA gene sequences from 4143-1 and the reference sequences from GenBank are shown. The tree was inferred by neighbor-joining analysis of 16S rRNA gene sequences with software Mega 4.0 (Center of Evolutionary Functional Genomics, Biodesign Institute, Arizona State University). Clones from this study are shown in bold. The numbers in parentheses are the GenBank accession numbers for sequences obtained from the National Center for Biotechnology Information database. Bootstrap percentages were obtained by using 1,000 replicates, and values greater than 50% are indicated at the nodes. The scale bar represents 5% of changes per nucleotide position. (*C*) Phylogenetic tree constructed from archaeal 16S rRNA gene sequences. (*D*) Rarefaction curves for bacterial and archaeal 16S rRNA sequences from 4143-1 sample.



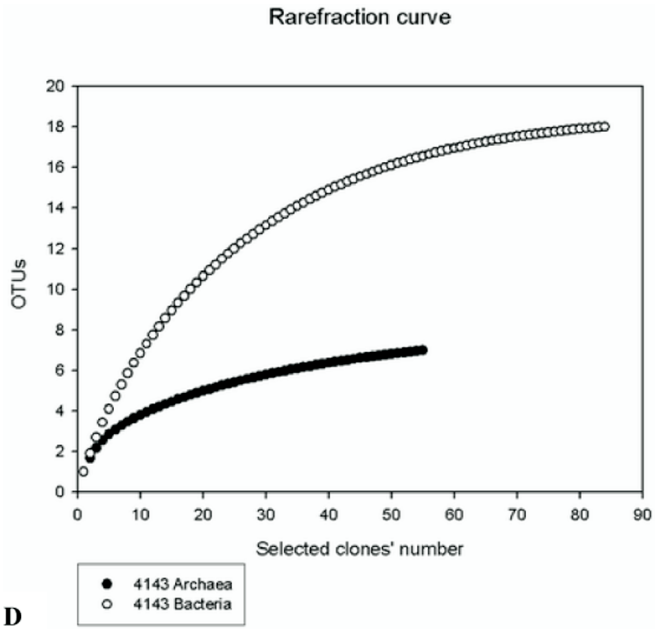
B

Fig. S2. Continued.



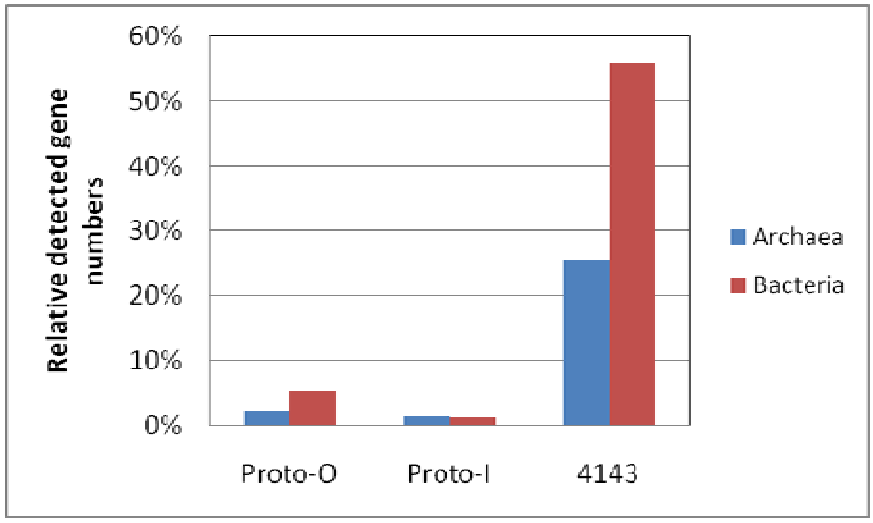
C

Fig. S2. Continued.

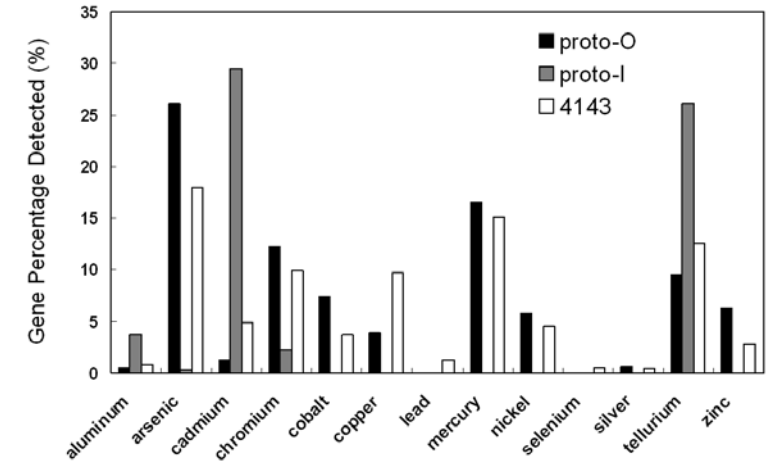


D

Fig. S2. Continued.

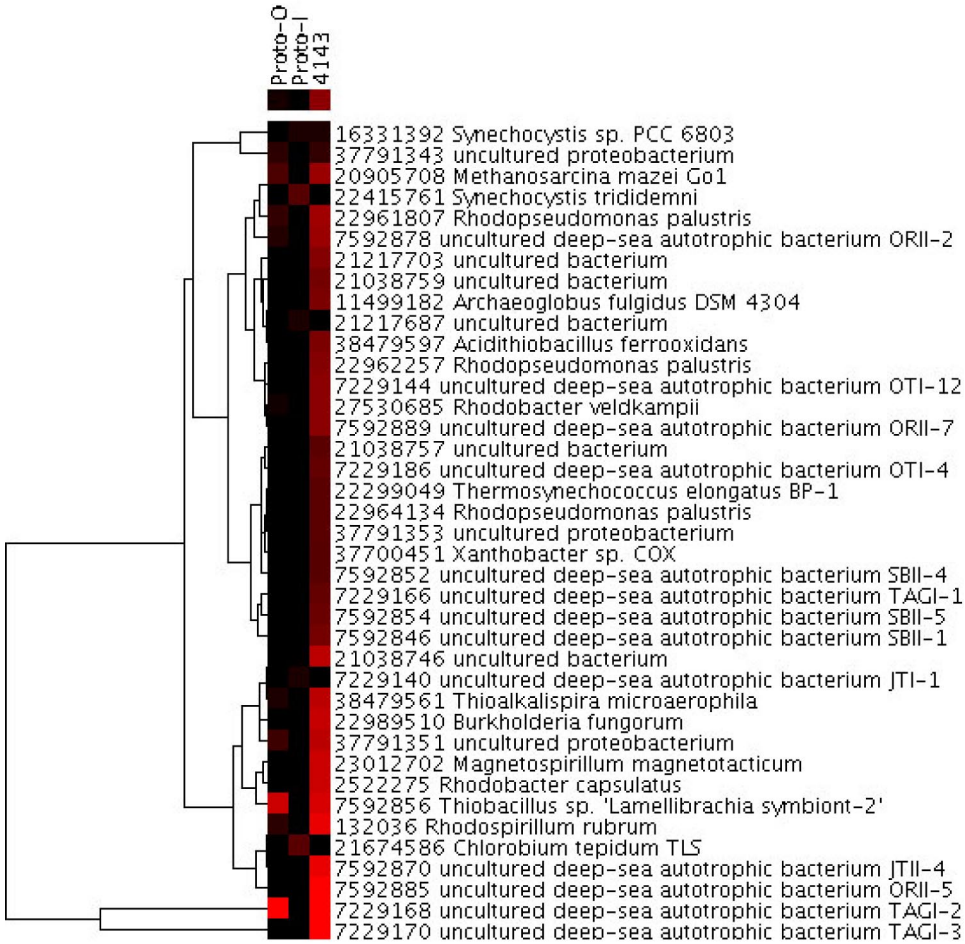


A



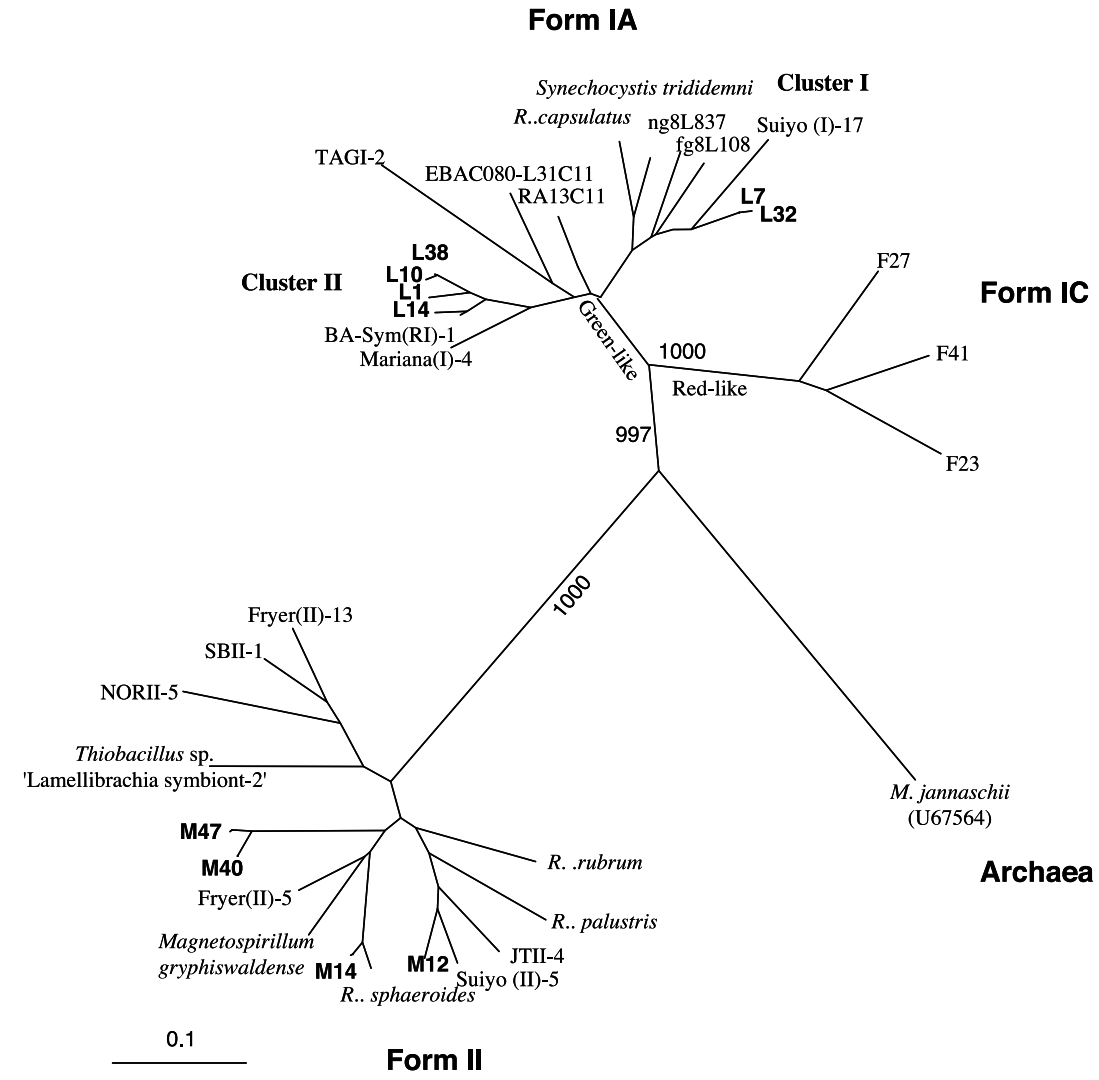
B

Fig. S3. Relative abundance of genes detected. (*A*) Relative proportions of detected archaeal or bacterial gene numbers from the chimney samples versus their respective total gene numbers on the chip. There are a total of 8,371 and 594 genes from Bacteria and Archaea on the chip, respectively. (*B*) Percentage of genes for metal resistance. The percentage of genes involved in aluminum, arsenic, cadmium, chromium, cobalt, copper, lead, mercury, nickel, selenium, silver, tellurium, and zinc were calculated by dividing the total signal intensity values for each individual metal-resistance group by the total intensity values of all metal-resistance genes detected on the array.



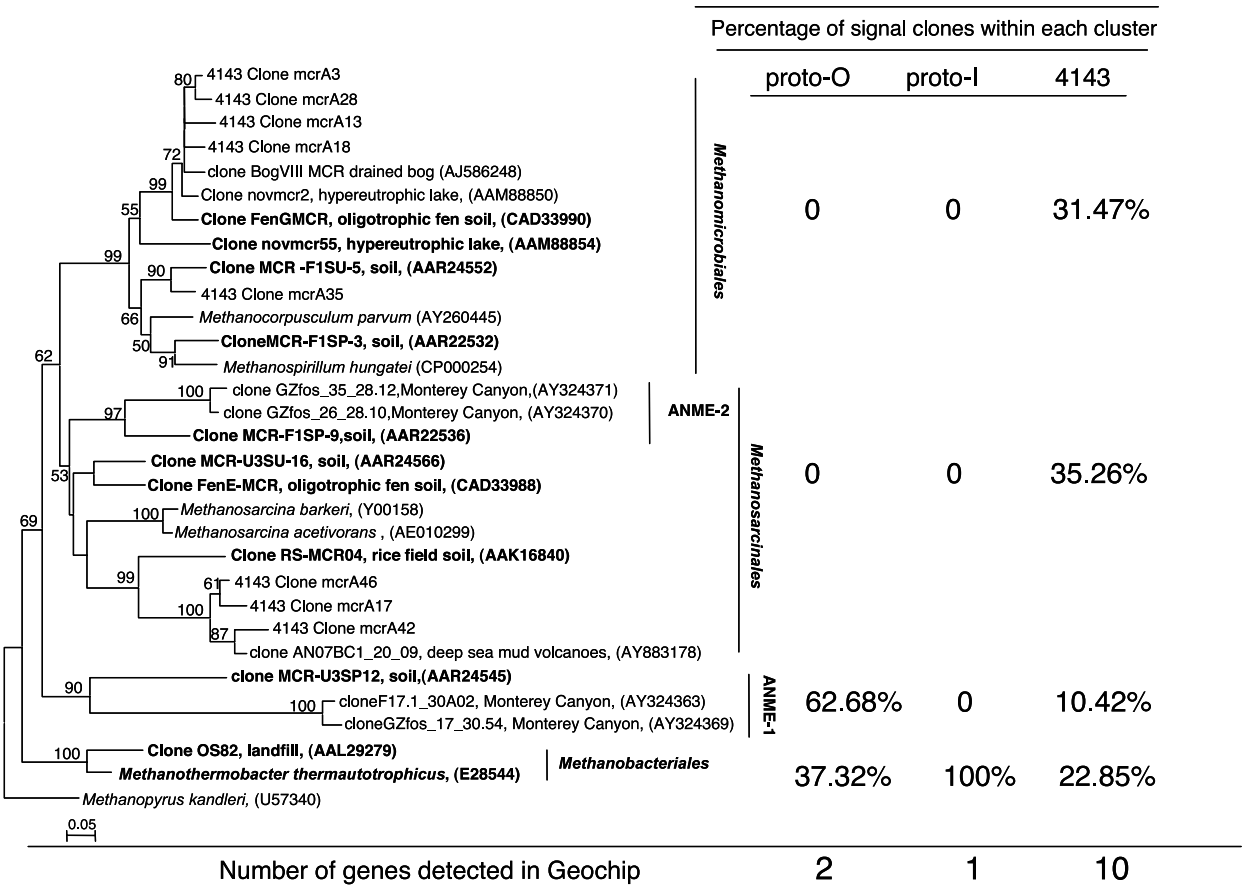
A

Fig. S4. Analysis of CO_2 fixation genes detected. (*A*) Hierarchical cluster analysis of *rbcL* genes based on hybridization signals for Proto-I, Proto-O, and 4143-1. The figure was generated by using CLUSTER (http://rana.lbl.gov/EisenSoftware.htm) and visualized with TREEVIEW (http://rana.lbl.gov/EisenSoftware.htm). Black represents no hybridization above background level, and red represents positive hybridization. The color intensity indicates differences in hybridization patterns. All genes detected in Proto-I and Proto-O are listed, but only the genes with signal intensities greater than 1 from 4143-1 were included here. (*B*) Phylogenetic tree based on the deduced amino acid sequences of RubisCO large-subunit genes. Clone sequences (OTUs) retrieved from *cbbL* and *cbbM* clone libraries are indicated in bold. (Scale bar: 0.1 substitutions per site.) See the Fig. S2*B* legend for details.



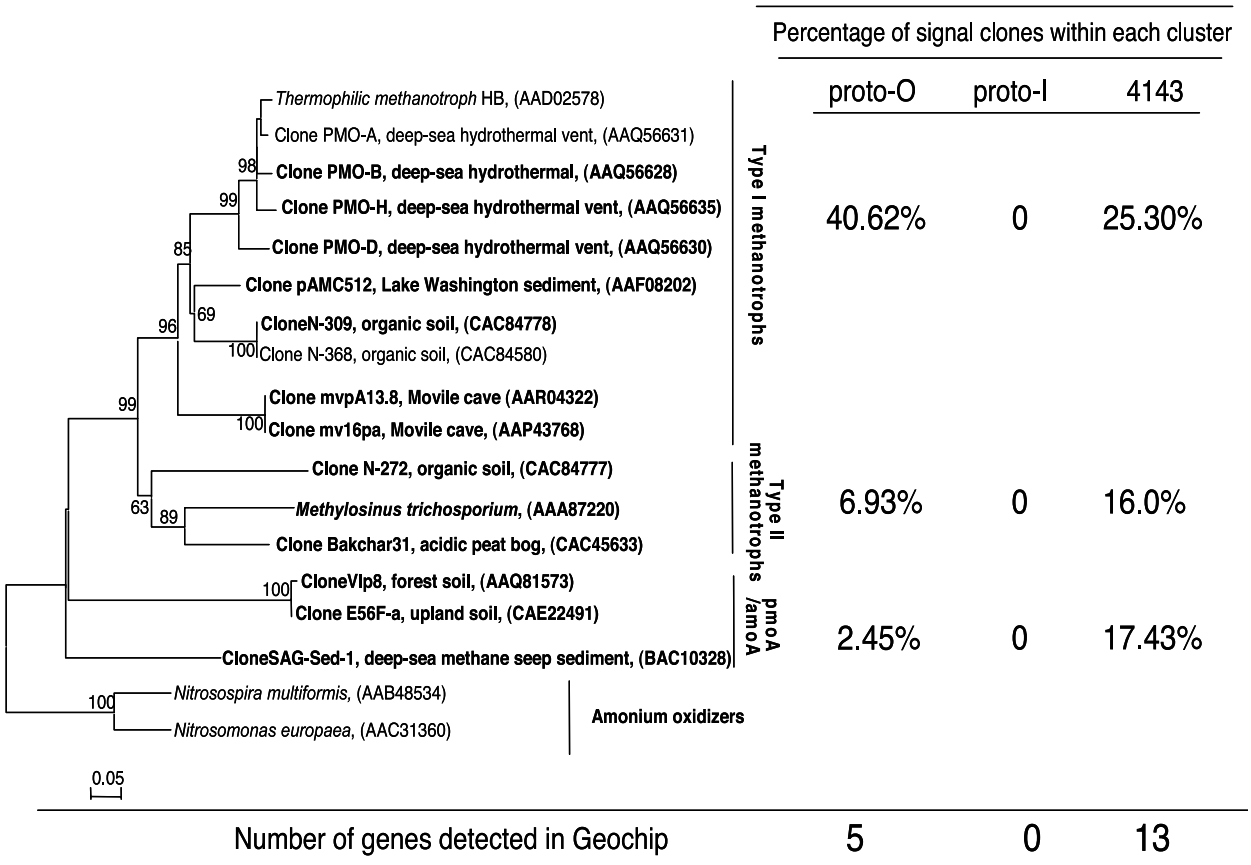
B

Fig. S4. Continued.



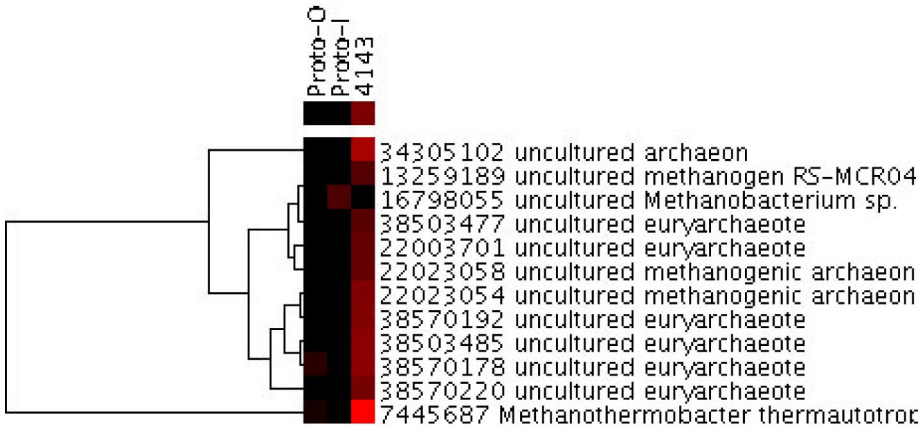
A

Fig. S5. Analysis of *mcrA* and *pmoA* genes detected. (*A*) Phylogenetic tree based on *mcrA* gene sequences. The clones retrieved from 4143-1 through the *mcrA* clone library are designated ''4143-1 Clone *mcrA*,'' followed by the clone number. (*B*) Phylogenetic tree based on *pmoA* gene sequences. Sequences collected from GeoChip hybridization are shown in bold. The tree was inferred by the neighbor-joining method with Mega 4.0. See the Fig. S2*B* legend for details. (*C*) Hierarchical cluster analysis of *mcrA* genes. (*D*) Hierarchical cluster analysis of *pmoA* genes. See Fig. S4*A* legend for details

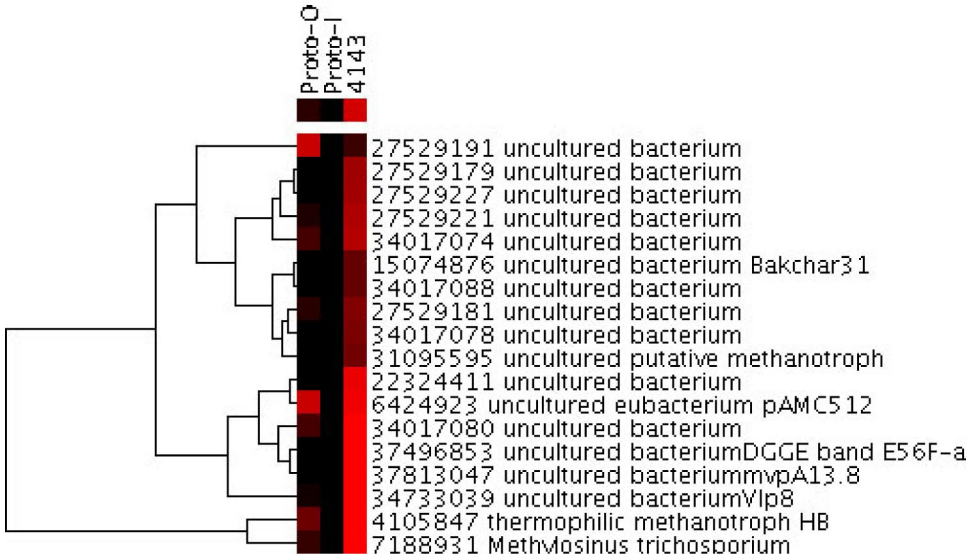


B

Fig. S5. Continued.



C



D

Fig. S5. Continued.

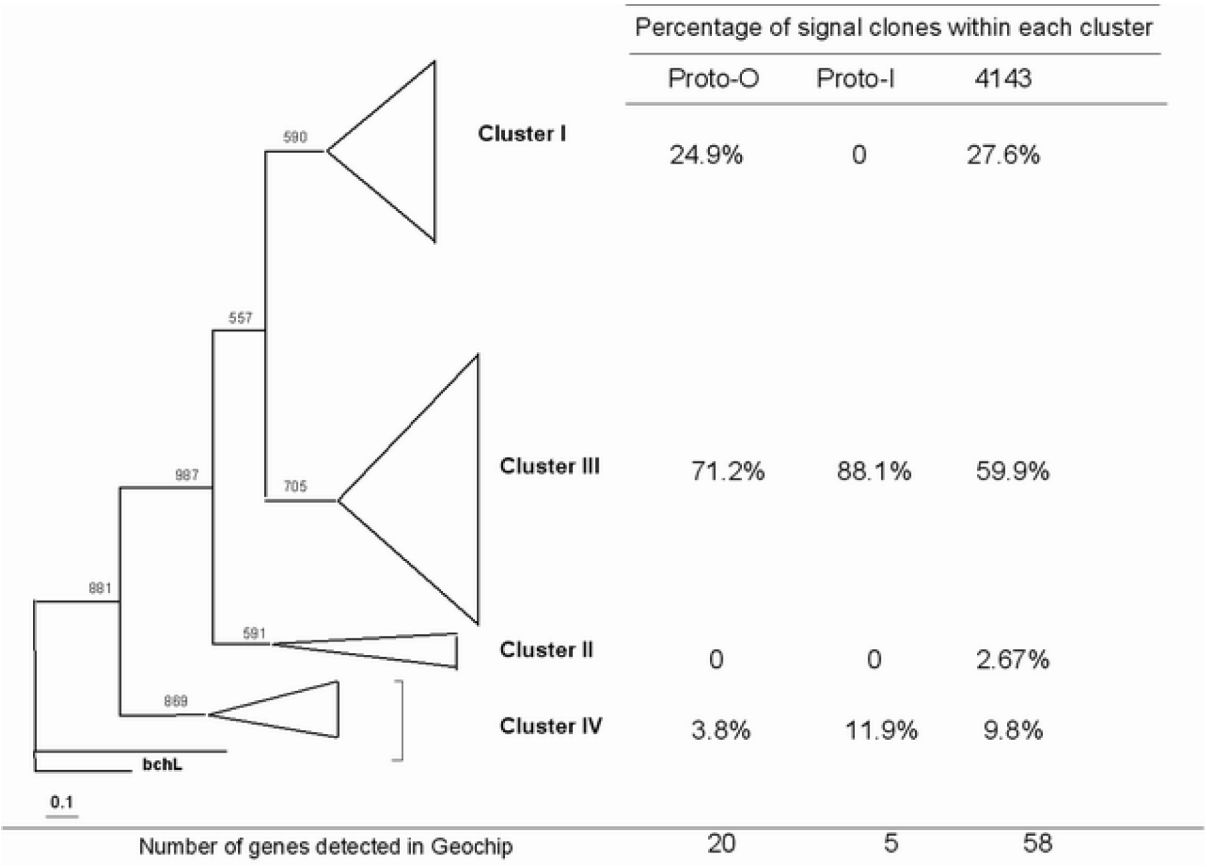


Fig. S6. Maximum-likelihood phylogenetic tree of *nifH* sequences obtained from GeoChip showing the phylogenetic relationship among the 4 *nifH* clusters. The depth and width of each wedge are proportional to the branch lengths and number of *nifH* sequences within each cluster, respectively. Cluster I contains α-, β-, and γ-proteobacterial *nifH* sequences; cluster III contains anaerobic bacterial *nifH* sequences; cluster II contains methanogen and alternative *nifH* sequences; and cluster IV contains divergent archaeal *nifH* sequences. Light-independent protochlorophyllide reductase subunit L gene (*bchL*) from *Chlorobium tepidum* was used as an outgroup. The percentage of signal clones detected in each sample was listed near the phylogenetic tree. The total number of genes detected in GeoChip was summarized.

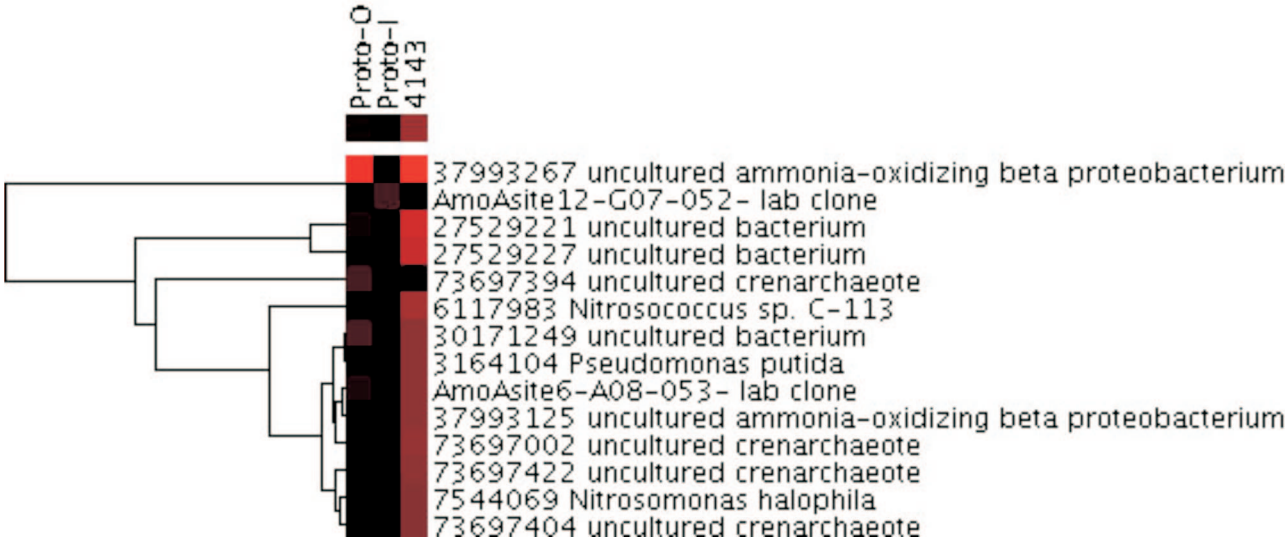


Fig. S7A. Hierarchical cluster analysis of *amoA* genes. See Fig. S4*A* legend for details.

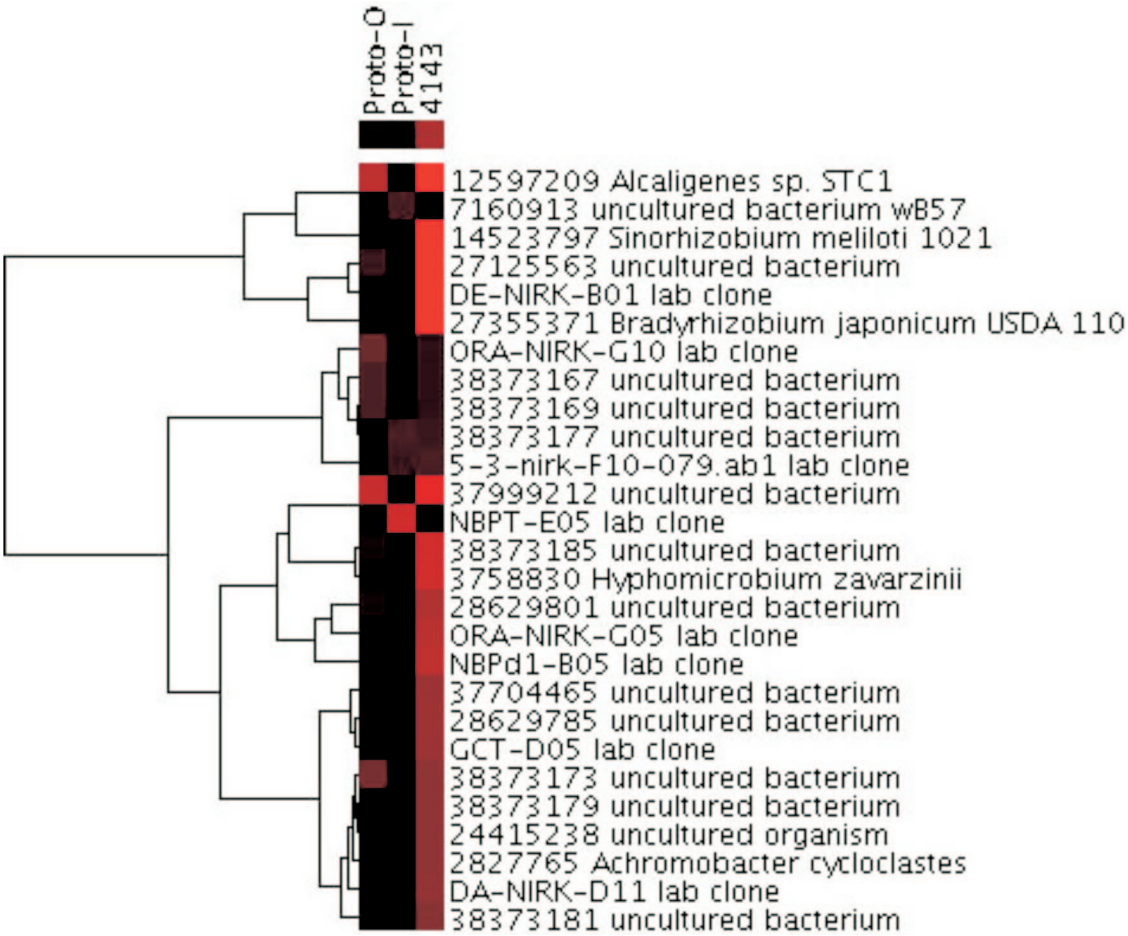


Fig. S7B. Hierarchical cluster analysis of *nirK* genes. See Fig. S4*A* legend for details.

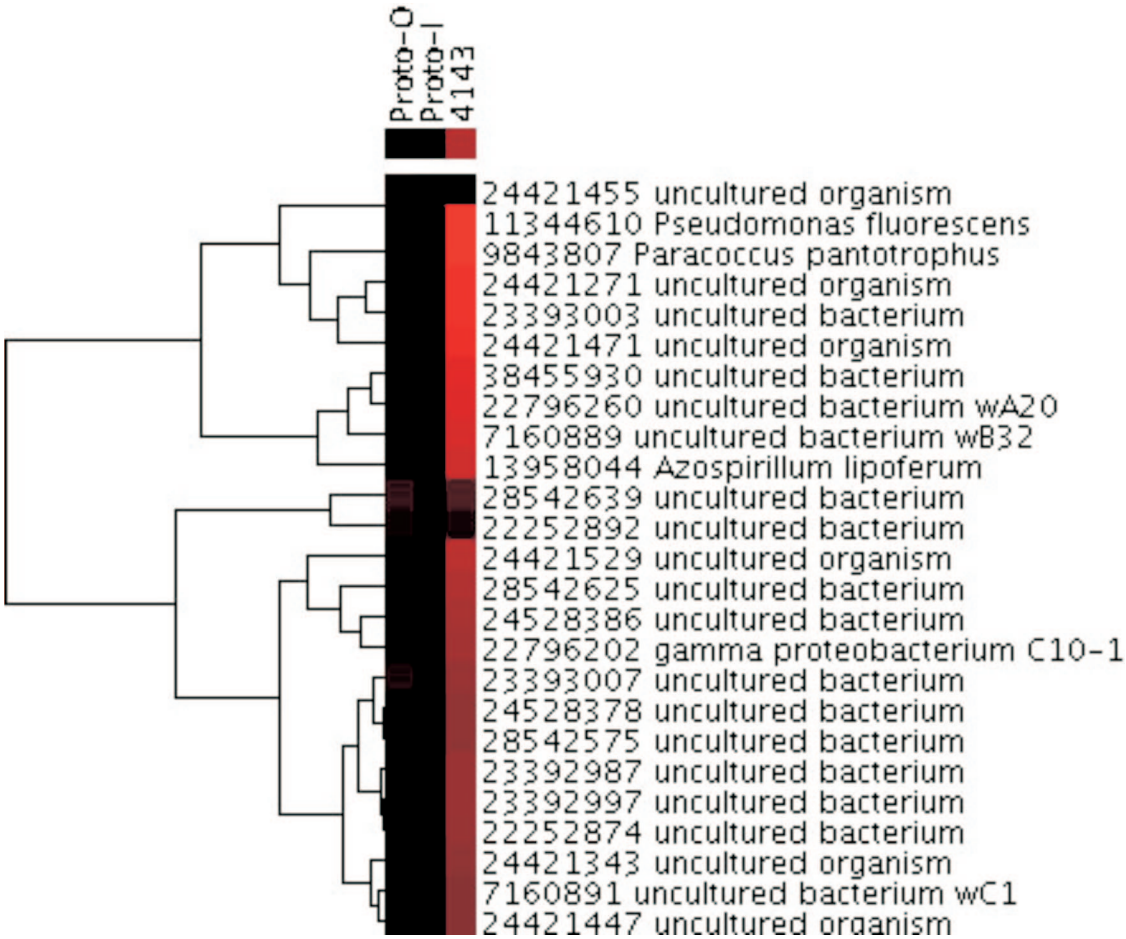


Fig. S7C. Hierarchical cluster analysis of *nirS* genes. See Fig. S4*A* legend for details.

Table S1. Distribution of functional genes detected for major metabolic processes in chimney samples from Juan de Fuca Ridge

	Sample, no. (%)		
Gene category	Proto-I	Proto-O	4143-1
Organic contaminant degradation	53 (46.9)	217 (43.1)	2,008 (37.1)
Carbon degradation	11 (9.7)	56 (11.1)	563 (10.4)
Carbon fixation	6 (5.3)	14 (2.8)	190 (3.5)
Nitrogen fixation	5 (4.4)	22 (4.4)	233 (4.3)
Nitrate reduction	8 (7.1)	30 (6)	534 (9.9)
Nitrification	11 (9.7)	23 (4.6)	391 (7.2)
Methanogenesis	1 (0.9)	2 (0.4)	62 (1.1)
Anaerobic oxidation of methane	0	1 (0.2)	3 (0.1)
Aerobic oxidation of methane	1 (0.9)	14 (2.8)	86 (1.6)
Sulfate reduction	3 (2.7)	28 (5.6)	396 (7.3)
Metal resistance	12 (10.6)	92 (18.3)	910 (16.8)
Total genes	113	504	5,414

Table S2. *rbcL* genes detected by GeoChip

Probe reference no.	Gene	Reference sequence/organism	Signal intensity
Proto-O			
BAC54026	rbcL	Rhodobacter veldkampii	0.2042
AAR21097	rbcL	Thioalkalispira microaerophila	0.357
BAA94444	rbcL	Uncultured deep-sea autotrophic bacterium ORII-2	0.4093
P04718	rbcL	Rhodospirillum rubrum	0.5155
AAR03652	rbcL	Uncultured proteobacterium	0.5576
ZP_00009413	rbcL	Rhodopseudomonas palustris	0.5996
AAR03656	rbcL	Uncultured proteobacterium	0.7224
AAM30945	rbcL	Methanosarcina mazei Go1	0.9123
BAA94433	rbcL	Thiobacillus sp. Lamellibrachia symbiont-2	2.4171
BAA92470	rbcL	Uncultured deep-sea autotrophic bacterium TAGI-2	4.1006
Proto-I			
AAM34461	rbcL	Uncultured deep-sea autotrophic bacterium JTI-1	0.2978
NP_442120	rbcL	Synechocystis sp. PCC 6803	0.3369
NP_662651	rbcL	Uncultured bacterium	0.3664
BAC10575	rbcL	Synechocystis trididemni	1.0053
AAR37722	rbcL	Chlorobium tepidum TLS	1.0612
4143			
BAA92470	rbcL	Uncultured deep-sea autotrophic bacterium TAGI-2	5.8437
BAA92471	rbcL	Uncultured deep-sea autotrophic bacterium TAGI-3	4.0193
BAA94447	rbcL	Uncultured deep-sea autotrophic bacterium ORII-5	2.9827
P04718	rbcL	Rhodospirillum rubrum	2.7958
BAA94440	rbcL	Uncultured deep-sea autotrophic bacterium JTII-4	2.7701
BAA94433	rbcL	Thiobacillus sp. Lamellibrachia symbiont-2	2.5599
ZP_00052722	rbcL	Magnetospirillum magnetotacticum	2.4597
AAC37141	rbcL	Rhodobacter capsulatus	2.4199
ZP_00034564	rbcL	Burkholderia fungorum	2.3454
AAR03656	rbcL	Uncultured proteobacterium	2.2593
AAR21097	rbcL	Thioalkalispira microaerophila	2.216
AAM26289	rbcL	Uncultured bacterium	2.1916
ZP_00009413	rbcL	Rhodopseudomonas palustris	1.9111
AAM30945	rbcL	Methanosarcina mazei Go1	1.8641
BAA94444	rbcL	Uncultured deep-sea autotrophic bacterium ORII-2	1.8121
BAA94449	rbcL	Uncultured deep-sea autotrophic bacterium ORII-7	1.6565
BAC54026	rbcL	Rhodobacter veldkampii	1.6529
BAA92481	rbcL	Uncultured deep-sea autotrophic bacterium OTI-12	1.6344
ZP_00009863	rbcL	Rhodopseudomonas palustris	1.611
AAM34469	rbcL	Uncultured bacterium	1.5313
NP_070416	rbcL	Archaeoglobus fulgidus DSM 4304	1.4807
AAR21099	rbcL	Acidithiobacillus ferrooxidans	1.4524
AAM26291	rbcL	Uncultured bacterium	1.4081
BAA94428	rbcL	Uncultured deep-sea autotrophic bacterium SBII-1	1.3347
BAA94432	rbcL	Uncultured deep-sea autotrophic bacterium SBII-5	1.2621
BAA92469	rbcL	Uncultured deep-sea autotrophic bacterium TAGI-1	1.2023
BAA92490	rbcL	Uncultured deep-sea autotrophic bacterium OTI-4	1.1499
AAM26292	rbcL	Uncultured bacterium	1.1211
AAR03657	rbcL	Uncultured proteobacterium	1.0893
ZP_00011740	rbcL	Rhodopseudomonas palustris	1.0766
NP_682296	rbcL	Thermosynechococcus elongatus BP-1	1.0719
AAR00245	rbcL	Xanthobacter sp. COX	1.0493
BAA94431	rbcL	Uncultured deep-sea autotrophic bacterium SBII-4	1.0112