Applied and Environmental Microbiology, Feb. 2011, p. 991–999 Vol. 77, No. 3
0099-2240/11/$12.00 doi:10.1128/AEM.01798-10

GeoChip-Based Analysis of the Functional Gene Diversity and Metabolic Potential of Microbial Communities in Acid Mine Drainage▽†

Jianping Xie,[1,2,3] Zhili He,[2] Xinxing Liu,[1,3] Xueduan Liu,[1,3] Joy D. Van Nostrand,[2] Ye Deng,[2] Liyou Wu,[2] Jizhong Zhou,[2] and Guanzhou Qiu[1,3]*

School of Mineral Processing and Bioengineering, Central South University, Changsha, Hunan 410083, China[1]; Institute for Environmental Genomics and Department of Botany and Microbiology, University of Oklahoma, Norman, Oklahoma 73072[2]; and Key Laboratory of Biometallurgy, Ministry of Education, Changsha, Hunan 410083, China[3]

Received 28 July 2010/Accepted 9 November 2010

Acid mine drainage (AMD) is an extreme environment, usually with low pH and high concentrations of metals. Although the phylogenetic diversity of AMD microbial communities has been examined extensively, little is known about their functional gene diversity and metabolic potential. In this study, a comprehensive functional gene array (GeoChip 2.0) was used to analyze the functional diversity, composition, structure, and metabolic potential of AMD microbial communities from three copper mines in China. GeoChip data indicated that these microbial communities were functionally diverse as measured by the number of genes detected, gene overlapping, unique genes, and various diversity indices. Almost all key functional gene categories targeted by GeoChip 2.0 were detected in the AMD microbial communities, including carbon fixation, carbon degradation, methane generation, nitrogen fixation, nitrification, denitrification, ammonification, nitrogen reduction, sulfur metabolism, metal resistance, and organic contaminant degradation, which suggested that the functional gene diversity was higher than was previously thought. Mantel test results indicated that AMD microbial communities are shaped largely by surrounding environmental factors (e.g., S, Mg, and Cu). Functional genes (e.g., *narG* and *norB*) and several key functional processes (e.g., methane generation, ammonification, denitrification, sulfite reduction, and organic contaminant degradation) were significantly ($P < 0.10$) correlated with environmental variables. This study presents an overview of functional gene diversity and the structure of AMD microbial communities and also provides insights into our understanding of metabolic potential in AMD ecosystems.

When sulfide ores are exposed to air, water, and microbes (autotrophic and heterotrophic archaea and bacteria) during the exploration of coals and metal deposits, acid mine drainage (AMD), which usually has low pH and high concentrations of sulfate and metals, is formed (5, 8, 34, 52). AMD is a prevalent, international environmental problem that threatens aquatic life and surrounding ecosystems (11, 15, 21, 24, 26, 33, 43, 44), although it also has been regarded as an excellent model to investigate linkages between microbial communities and geochemistry because of its relatively simple microbial community composition and structure (5, 6, 8). A better understanding of the microbial community diversity and ecology in AMD would provide a scientific foundation not only for the bioremediation of AMD-contaminated environments but also for potential applications for the energy-efficient recovery of valuable metals from mine waste and the removal of sulfur from coal (6, 23).

The phylogenetic diversity of AMD microbial communities has been well studied by 16S rRNA-based approaches, including PCR cloning (36, 54, 55, 57), denaturing gradient gel electrophoresis (DGGE) (10, 20, 21), fluorescence *in situ* hybridization (9, 29, 40), and oligonucleotide microarrays (14, 56). For instance, the DNA sequencing of a natural acidophilic biofilm (pH 0.80) revealed that pathways for carbon and nitrogen fixation and energy generation might occur in these communities and provided insight into survival strategies in such an extreme environment (45). Gene expression and metabolic functions in a natural AMD microbial biofilm also were examined using proteomic approaches (37). A novel cytochrome was found to play an important role in iron oxidation (37). Community genomic and proteomic analyses of the chemoautotrophic iron oxidizers *Leptospirillum rubarum* (group II) and *Leptospirillum ferrodiazotrophum* (group III) showed that both groups had genes involved in carbon fixation and the biosynthesis of vitamins, fatty acids, and biopolymers that were active (16). In another study, the denitrification potential was examined by microcosm incubations using sediment from acidified streams, and results suggested that denitrification occurred in AMD streams and reduce acidity if stimulated within the environment (4). In addition, several studies have focused on the phylogenetic response of AMD microbial communities to geochemical variables, and the results showed that the microbial communities were shaped mainly by the surrounding environmental factors, such as pH, S, and metal ions, including Pb, Zn, Cu, and Fe^{2+} (5, 8, 57). However, although all of the

* Corresponding author. Mailing address: School of Mineral Processing and Bioengineering, Central South University, Changsha, Hunan 410083, China. Phone: 86 (731) 8879815. Fax: 405 325 7552. E-mail: x-mine@mail.csu.edu.cn.

† Supplemental material for this article may be found at http://aem.asm.org/.

▽ Published ahead of print on 19 November 2010.

studies described above revealed some functional potential and activity in AMD microbial communities, little is known about the functional diversity and metabolic potential at the community level, and the relationship between the functional gene structure of AMD microbial communities and the surrounding environmental factors remains unclear. Therefore, a more comprehensive characterization of the AMD microbial communities is necessary to provide a full picture of their functional potential and the underlying mechanisms controlling the AMD microbial community diversity and structure.

GeoChip 2.0, a functional gene-based microarray, contains more than 24,000 probes covering about 10,000 genes in 150 functional groups involved in carbon, nitrogen, sulfur, and phosphorus cycling, metal resistance, and organic contaminant degradation (18). It has been used to analyze microbial communities from different habitats, such as soil (25, 53, 60), water (18, 42, 46, 47), oil fields (28), marine sediment (49), extreme environments (31, 48), bioreactor systems (38), and other habitats (22).

In this study, GeoChip 2.0 was employed to address two key questions. (i) What is the functional gene diversity, structure, and metabolic potential of AMD microbial communities? (ii) How does the functional structure of microbial communities respond to environmental variables? To answer these questions, five AMD samples were obtained from three copper mines in southeast China. Our results showed that, in contrast to previous thoughts, microbial functional gene diversity was quite high in AMD microbial communities, which might be shaped largely by the surrounding environmental factors. This study provides a comprehensive survey of the functional gene diversity and metabolic potential of AMD microbial communities and insights into linkages between AMD microbial communities and environmental factors.

MATERIALS AND METHODS

Site description, sample collection, and geochemical analysis. A total of five AMD samples were collected in April 2007 from three copper mines, Tongchang (TC), Yinshan (YS), and Yongping (YP), in southeast China (see Table S1 and Fig. S1 in the supplemental material). TC is located in a northeast-trending magmatic belt along the southeastern margin of the Yangtze Craton, and it is the largest open-pit copper mine in China and the third largest in the world (17, 27). It has been mined for more than 800 years and currently produces more than 100,000 tons of ore per year (55). Large amounts of mine waste are produced and exposed to air and water, resulting in a significant amount of AMD. YS, ~15 km to the west of TC, is a deep open-pit and multimetal mine located in the Le-De metallogenic belt in Jiangxi Province (58). YP, ~90 km south of TC, is a Cu-Mo deposit located in the Huaiyushan-North Wuyishan Cu-Pb-Zn polymetallic metallogenic belt in southeastern China (27).

Two samples were obtained from both TC and YS. TC_SLS was obtained from an accumulated water pool in the working area of TC, and TC_DWT was obtained from a pump station where AMD, rain water, and accumulated water were pumped from the open pit. YS_WKB was obtained from an accumulated water pool near the tailing dam of YS, and YS_1SJ (mixture of ground water, AMD, and accumulated water at 160 m below ground) was obtained from the no. 1 vertical shaft in YS. The YP sample was obtained from an accumulated water pool in the opencast working zone of YP.

Water (50 liters) was obtained from three different locations within each site at 0 to 10 cm below the water surface. All samples were transferred to the laboratory within 48 h at room temperature. Since low DNA yields were obtained from all samples, the three replicates were combined, resulting in one sample for each site. A total of 50 ml of the original water samples was used for geochemical analysis. The remainder of each water sample was filtered (0.22-μm nylon filters), and the filters were stored at −80°C until used for DNA extraction.

The latitude and longitude of sampling locations were determined using a Garmin eTrex (Garmin, KS). The temperature and pH of each site were measured on site. A total of 20 elements were measured using inductively coupled plasma-atomic emission spectrometry (ICP-AES; PS-6; Baird).

Microbial community DNA isolation and purification. The community DNA was extracted using a freeze-grinding method as described previously (59) and purified using a Promega Wizard DNA clean-up system (Madison, WI) according to the manufacturer's directions. DNA quality was evaluated by the absorbance ratios at A_{260}/A_{280} and A_{260}/A_{230} using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies Inc., Wilmington, DE). Only DNA with A_{260}/A_{280} and A_{260}/A_{230} ratios of >1.7 and 1.8, respectively, were used for further analysis. DNA was quantified using PicoGreen (1) and a FLUOstar Optima (BMG Labtech, Jena, Germany). Purified DNA was stored at −80°C until used.

Microbial community DNA amplification, labeling, microarray hybridization, and scanning. DNA (100 ng) was amplified using a Templiphi kit (GE Healthcare, Piscataway, NJ) with the modification of adding 0.1 μM spermidine and 260 ng · μl^{-1} single-stranded DNA binding protein to improve the amplification efficiency and representativeness (50). The amplified DNA (~2.5 μg) was labeled with the fluorescent dye Cy5 (GE Healthcare) by random priming (46, 50). Labeled DNA was purified with a QIAquick purification kit (Qiagen) and then dried in a SpeedVac (45°C, 45 min; ThermoSavant). Dried, labeled DNA was suspended in hybridization buffer (130 μl; 50% formamide, 3× SSC [1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate], 0.3% SDS, 0.7 μg · μl^{-1} Herring sperm DNA, 0.85 mM dithiothreitol [DTT]), denatured at 98°C for 3 min, and then kept at 65°C until hybridization. Hybridizations were performed at 42°C for 10 h using a Tecan HS4800 Pro hybridization station (Tecan). All samples were amplified, labeled, and hybridized in triplicate as pseudoreplicates. After hybridization, arrays were scanned with a ScanArray 500 microarray scanner (PerkinElmer, Boston, MA) at 633 nm using a laser power of 90% and a photomultiplier tube (PMT) gain of 75%. Scanned images were processed using ImaGene, version 6.1 (BioDiscovery, El Segundo, CA).

GeoChip data preprocessing. Raw data obtained using ImaGene were uploaded to our laboratory's microarray data manager (http://ieg.ou.edu/microarray/) and preprocessed using the data analysis pipeline with the following major steps: (i) spots flagged as poor quality by ImaGene 6.1 and with a signal-to-noise ratio [SNR; SNR = (signal intensity − background)/standard deviation of background] of less than 2.0 were removed; (ii) the normalized intensity of each spot was calculated by dividing the signal intensity of each spot by the mean intensity of the effective spots of the array; (iii) if any replicates had a signal intensity minus mean signal intensity of more than two times the standard deviation, this replicate was removed as an outlier; and (iv) at least two spots out of three replicates were required for a gene to be positive. Preprocessed GeoChip data were used for further statistical analysis.

Statistical analysis. The paired spatial distance matrices (as the spatial distance factor) of all samples were calculated with the latitude and longitude data using a web-based program (http://www.movable-type.co.uk/scripts/latlong.html). Physical (pH and temperature) and chemical (elements) properties of AMDs (as geochemical factors) were z transformed to convert all measurements to the same scale prior to further analysis using the formula $z = (x_i - \bar{x})/s$, where x_i is the sample value, $\bar{x}$ is the mean of all samples, and s is the standard deviation (41).

Principal component analysis (PCA) was performed to analyze the geochemical gradient with CANOCO (version 4.5; Biometris-Plant Research International, The Netherlands). Diversity indices and Mantel tests were performed using R 2.9.1 (http://www.r-project.org/). The Mantel test (30) is a statistical test of the correlation between two matrices and is commonly used in ecology. Hierarchical clustering analysis was performed with CLUSTER (http://rana.lbl.gov/EisenSoftware.htm) and visualized by TREEVIEW (12).

RESULTS

Geochemical properties in sampling sites. The five AMD samples were obtained from five different sites, with TC_SLS and TC_DWT being close to each other, YS_WKB and YS_1SJ being close to each other, and YP being distant from both TC and YS (see Fig. S1 in the supplemental material). TC_SLS had the highest temperature (30°C), which was almost 10°C higher than that of YS_WKB and 13.5°C higher than those of the other three AMDs. The pH was not substantially different among these samples and ranged from 2 to 3 (see Table S1 in the supplemental material). These samples were heterogeneous in terms of geochemical components, but all

TABLE 1. Geochemical properties of sampled AMD

Sample ID	Amount (μg/ml) in sample																			
	S	Fe	Al	Mg	Ca	Zn	Mn	Cu	Si	Na	P	Ni	Co	V	B	Y	La	Zr	Sc	Ba
TC_SLS	10,151	5,158	1,410	562.9	304.7	1.9	13.4	40.6	66.6	6.1	52.7	6.1	4.5	1.5	1.2	0.7	0.6	0.5	0.4	0.2
TC_DWT	1,252	777.8	153.8	97.1	109.9	2.2	20.3	4.9	28	10.7	6.3	1.2	0.4	0.3	0.2	0.2	0.3	0.3	0.07	0.2
YS_WKB	14,034	9,887	2,854	874.8	227.3	455.9	660	491.2	65.7	30.6	65.2	9.5	3.6	3.6	2.2	1.9	1.4	0.7	0.8	0.2
YS_1SJ	1,735	619.9	102	249.1	163.1	571.5	193.7	39.7	20.3	110.4	4.1	2	0.5	0.2	0.2	0.2	0.2	0.2	0.07	0.1
YP	4,713	3,423	314	478.3	590.4	12.2	20.2	80.7	58.8	9.2	24.4	2.6	1.7	0.9	0.8	0.5	0.7	0.5	0.2	0.2

were rich in sulfur (1.25 ~ 14.03 g/liter) and metals, such as Fe (0.62 ~ 9.90 g/liter), Al (0.10 ~ 2.85 g/liter), Mg (0.10 ~ 0.56 g/liter), and Ca (0.11 ~ 0.60 g/liter). YS_WKB and YS_1SJ had much higher concentrations of Mn (0.19 ~ 0.66 g/liter), Na (0.03 ~ 0.11 g/liter), and Zn (0.46 ~ 0.57 g/liter) than other AMDs. This may be because they were close to each other at the same Cu mine (Table 1). In general, YS_WKB had the highest concentration of most elements, especially Cu, S, and Mg, whereas TC_DWT had the lowest (Table 1). These relationships also were indicated by PCA, with 91.3% of the total variance being explained by the first two axes (PC1 and PC2). PC1 explained 72.9% of the total variance, and PC2 explained 18.4% of the total variance (see Fig. S2 in the supplemental material).

Overview of functional gene diversity and structure of AMD microbial communities. The examined microbial communities showed high variations among these five AMD samples as measured by the number of detected genes, overlapping genes between samples, unique genes, and diversity indices (Table 2; also see Table S2 and S4 in the supplemental material). The number of detected genes was 1,051, 985, 653, 185, and 157 for YS_WKB, YP, TC_DWT, YS_1SJ, and TC_SLS, respectively. YP and YS_WKB had the most overlapped genes (708; 49.9%), while YS_1SJ and YP had the fewest (107; 9.6%). As expected, YS_WKB had the highest number of unique genes (260; 23.6%), while TC_SLS had the fewest (13; 7.7%) (Table 2). Simpson's reciprocal diversity index (1/*D*) was highest in YS_WKB and lowest in YS_1SJ (621.27 and 92.35, respectively), with the overall diversity in the following order: YS_WKB > YP > TC_DWT > TC_SLS > YS_1SJ. Similar results were observed with the Shannon-Weaver index (H′). Evenness was comparable (0.53 to 0.59) for the four surface/subsurface samples, while it was lower for the deep underground sample (0.46) (Table 2). Among the 1,530 genes detected in at least one sample, 176 were involved in carbon degradation, 43 in carbon fixation, 31 in methane oxidation, 6 in methane generation, 231 in nitrogen cycling, 322 in metal resistance, 85 in sulfite reduction, and 636 in contaminant degradation (see Table S2 and Fig. S3 in the supplemental material).

Phylogenetically, 1,448 genes were derived from bacteria, 23 from archaea, and 96 from fungi. For bacteria, 246 genes were from *Alphaproteobacteria*, 135 from *Betaproteobacteria*, 356 from *Gammaproteobacteria*, 62 from *Deltaproteobacteria*, 144 from *Actinobacteria*, 43 from *Cyanobacteria*, 106 from *Firmicutes*, and 8 from *Planctomyces* (see Table S3). Taken together, these results indicated the overall functional gene as well as the phylogenetic diversity of these AMD microbial communities appeared to be quite high.

Analysis of detected functional genes. To visualize the patterns of all detected functional genes among these five AMD samples, a cluster analysis was performed (Fig. 1). A total of five different patterns were observed. Group 1 was the smallest cluster, containing only 41 (2.67%) genes detected primarily in TC_SLS. Most of these genes are involved in organic contaminant degradation and metal resistance. Group 2 had 443 (28.90%) genes, primarily from YP, with most of these genes being involved in organic contaminant degradation, carbon degradation, and nitrogen fixation. Group 3 contained 105 (6.85%) genes, primarily in TC_1SJ, and most are involved in denitrification and carbon degradation. Group 4 had 442 (28.83%) genes detected primarily in YS_WKB, and most are involved in metal resistance and dissimilatory sulfite reduction. Group 5 was the largest cluster, with 500 genes (32.62%) primarily from TC_DWT, although many of these genes also were detected in YS_WKB and YP. Most of these genes are involved in ammonification, carbon fixation, and organic contaminant degradation (Fig. 1). To better understand the functional processes of the AMD microbial communities involved in metal resistance, sulfur metabolism, and carbon and nitrogen metabolism, more-detailed analyses of these gene groups were performed.

TABLE 2. Gene overlap, uniqueness, diversity indices, and total detected gene numbers of AMD samples[a]

Unique and overlap genes	TC_SLS	TC_DWT	YS_WKB	YS_1SJ	YP	Shannon Weaver index	Shannon Weaver evenness	1/*D*	Total no. of genes detected
TC_SLS	**13 (7.7)**	*141* (*19.6*)	*151* (*13.5*)	*46* (*14.2*)	*142* (*13.5*)	4.85	0.59	100.25	157
TC_DWT		**90 (13.0)**	*549* (*44.2*)	*94* (*11.8*)	*485* (*39.3*)	6.21	0.53	367.05	653
YS_WKB			**260 (23.6)**	*115* (*9.7*)	*708* (*49.9*)	6.72	0.56	621.27	1,051
YS_1SJ				**70 (35.0)**	*107* (*9.6*)	4.95	0.46	92.35	185
YP					**256 (24.9)**	6.65	0.57	588.6	985

[a] Values in parentheses are percentages. Italicized values indicate the number of overlapping genes between samples, boldface values indicate the number of unique genes in each sample.

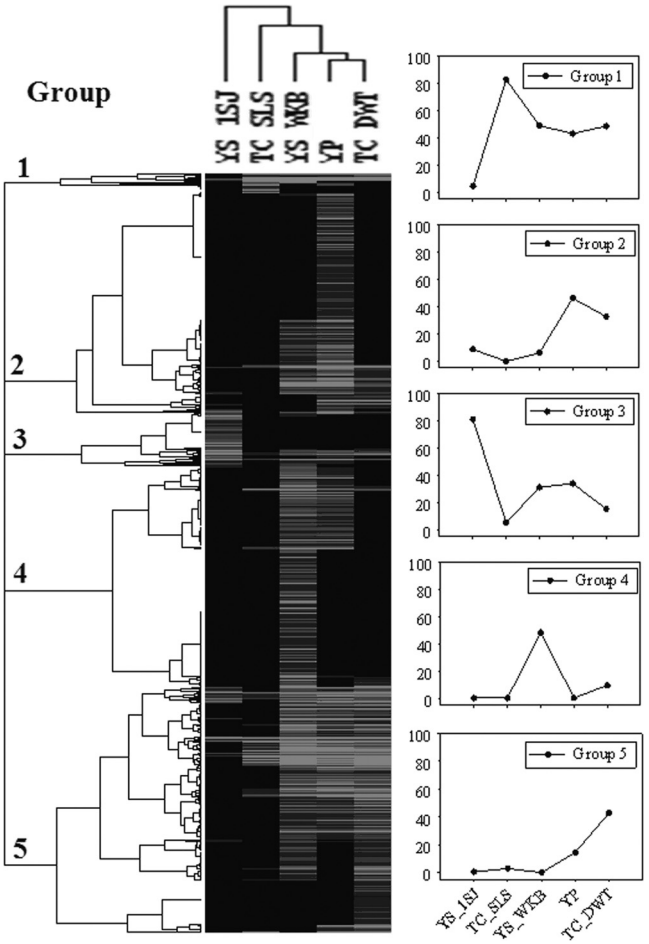

FIG. 1. Hierarchical cluster of the five AMD samples based on the signal intensity of all detected genes. The figure was generated by CLUSTER and visualized by TREEVIEW (http://rana.lbl.gov/EisenSoftware.htm). Black represents no hybridization above background levels, and gray represents positive hybridization. The color intensity indicates differences in hybridization signal. Average signal intensities of these groups for each sample are shown on the right.

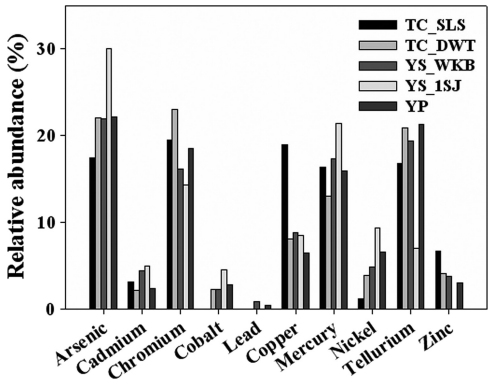

FIG. 2. Relative abundance of detected metal resistance genes. The relative abundance of genes was calculated by dividing the total signal intensity of each group of metal resistance genes by the total intensity of all metal resistance genes detected by GeoChip 2.0. Groups of metal resistance genes include As, Cd, Cr, Co, Cu, Hg, Ni, Te, and Zn.

Metal resistance. Since the AMD environments are rich in various metals, such as Fe, Zn, and Cu, metal resistance genes were examined. A total of 242, 213, 138, 42, and 38 genes involved in metal resistance/reduction were detected in YS_WKB, YP, TC_DWT, TC_SLS, and YS_1SJ, respectively (see Table S2 in the supplemental material). Generally speaking, the relative abundances of Cu and Zn resistance genes were high, while those of As and Ni resistance genes were low in TC_SLS. TC_DWT had a high relative abundance of Cr and Se resistance genes but a low relative abundance of Cd and Hg resistance genes. YS_WKB had high relative abundances of Cd and Pb resistance genes. YS_1SJ had high relative abundances of As, Co, Ni, and Ag resistance genes but low relative abundances of Cr and Te resistance genes. YP had the lowest relative abundance of Cu resistance genes among the five samples (Fig. 2). A Mantel test showed that the abundance of Zn resistance genes was positively correlated with the Zn concentration ($r = 0.427$; $P = 0.049$), although no such significant correlations were observed for other metals and their related gene abundances. All of these results indicate that metal resistance genes were highly diverse among all AMD microbial communities.

Sulfur metabolism. High concentrations of sulfur were observed in all five AMD samples (Table 1), and sulfur metabolism is considered an important process in AMD microbial communities. GeoChip 2.0 contains probes for *dsrA* and *dsrB* to analyze potential sulfur reduction and sulfate-reducing bacterial populations, although it does not have functional markers for sulfur oxidation. *dsrA* and *dsrB* were detected in all samples (YP, 55; YS_WKB, 51; TC_DWT, 27; YS_1SJ, 9; and TC_SLS, 8) (see Table S2). Of 85 *dsrA* genes detected, 26 were from uncultured sulfate-reducing bacteria, 17 from laboratory clones, and 42 from cultured sulfate-reducing bacteria (Fig. 3). Hierarchical cluster analysis showed that YS_1SJ and YS_WKB clustered together, while TC_SLS and TC_DWT clustered together and then with YP, which appeared to be geographically dependent (Fig. 3). Mantel test analysis showed no significant ($r = -0.313$; $P = 0.834$) correlation between the S concentration and the abundance of *dsrA* and *dsrB* genes. These results suggest that S reduction occurs in these AMD microbial communities, and that the sulfate reduction potential is geographically dependent.

Functional genes involved in the C cycle. Since carbon sources are severely limited in AMD, microbial carbon cycling may play an important role in this ecosystem. The metabolic potential for carbon fixation, carbon degradation, methane generation, and methane oxidation were detected in all five samples (Fig. 4a). A key enzyme, D-ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO), involved in carbon fixation via the Calvin cycle, was detected in all of the AMD samples. Although there are 391 *rbcL* probes on GeoChip 2.0, only 15, 12, 9, 2, and 1 genes were detected in YS_WKB, YP, TC_DWT, TC_SLS, and YS_1SJ, respectively. Most of detected genes (16/26) are derived from uncultured bacteria, and 10 were from cultured organisms, including three from *Acidithiobacillus ferrooxidans* and one each from *Thiobacillus* sp., *Rhodospirillum rubrum*, *Thioalkalispira microaerophila*, *Gloeobacter violaceus*, *Bacillus cereus*, *Methylocapsa acidiphila*, and *Syn-*

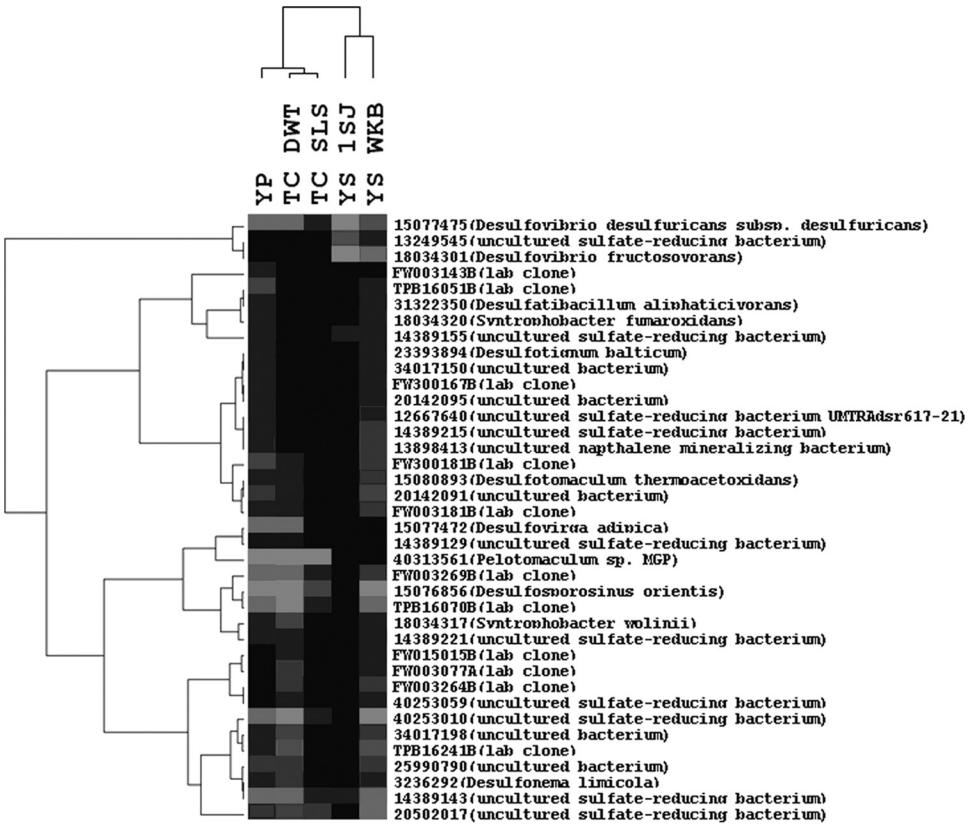

FIG. 3. Hierarchical cluster analysis of all five AMD samples based on the hybridization signal intensity of sulfate reduction genes (*dsrA* and *dsrB*). Genes that were detected only in one sample are not shown. Details for hierarchical cluster analysis are described in the legend to Fig. 1.

echocystis sp. (Fig. 4b). Genes involved in the degradation of cellulose, chitin, and lignin also were detected, and the highest signal intensities were observed in YS_WKB and YP (see Fig. S5 in the supplemental material). For methane metabolism, a higher signal intensity of methane oxidation genes (*mmo* and *pmo*) was detected than that of methane generation genes (*mcr*) in all five AMD samples (see Fig. S5). Most of the *mcr* genes were derived from uncultured archaea, which suggests that archaea play an important role in methane cycling in AMD. The results described above suggest that most functional processes in the carbon cycle (e.g., carbon fixation and methane oxidation) exist in AMD systems.

Functional genes involved in the N cycle. A total of 35 *nif* genes for nitrogen fixation, 78 genes (*ureC* and *gdh*) for ammonification, 10 genes (*nasA*) for assimilatory nitrate/nitrite reduction, 38 genes (*amoA*, *amoB* and *hao*) for nitrification, and 70 genes (*narG*, *nirS*, *nirK*, *narB*, *norB*, and *nosZ*) for denitrification were detected across the five AMD samples (Fig. 5; also see Table S2 in the supplemental material). A total of 33 *nifH* gene sequences were detected, with 27 derived from uncultured microorganisms and six from cultured organisms, including two from *Calothrix* sp. and one each from *Chlorogloeopsis* sp., *Paenibacillus macerans*, *Vibrio cincinnatiensis*, and *Anabaena variabilis* (see Fig. S4). Fifteen *amoA* genes were detected, with 2 derived from *Nitrosomonas* sp. strains C-113a and AL212, 2 from *Nitrosococcus oceani*, and 11 from uncultured bacteria.

Six functional genes from denitrification pathways are represented on GeoChip 2.0. Of these, 23 *narG*, 17 *norB*, 15 *nirK*, 13 *nirS*, 13 *nosZ*, and 5 *narB* genes were detected in the AMD samples, and most (64.28%) were from as-yet uncultivated bacteria. Although a low abundance of genes involved in other nitrogen processes were observed in YS_1SJ, this sample had a relatively high abundance of the genes involved in denitrification, which may be consistent with its anaerobic environment (Fig. 5). The results described above indicate that most of the functional genes involved in the nitrogen cycle are present in AMD systems.

Relationship between microbial communities and environmental factors. To examine if spatial distance and environmental factors affect the structure of AMD microbial communities, a Mantel test was performed. Our analysis showed that physical and geochemical variables had a significantly positive correlation ($r = 0.6107$, $P = 0.015$) with microbial communities, although spatial distance was not significantly correlated with changes in microbial communities. These results indicated that AMD microbial communities are shaped largely by the surrounding environmental factors instead of the spatial distance.

To further determine the most key individual environmental factors affecting the microbial community functional structure, Mantel tests of all five AMDs were performed with 20 individual environmental variables. The results showed that the gradients of S, Mg, Cu, Ni, Co, B, and La were significantly ($P < 0.05$) correlated with the microbial functional structures, while

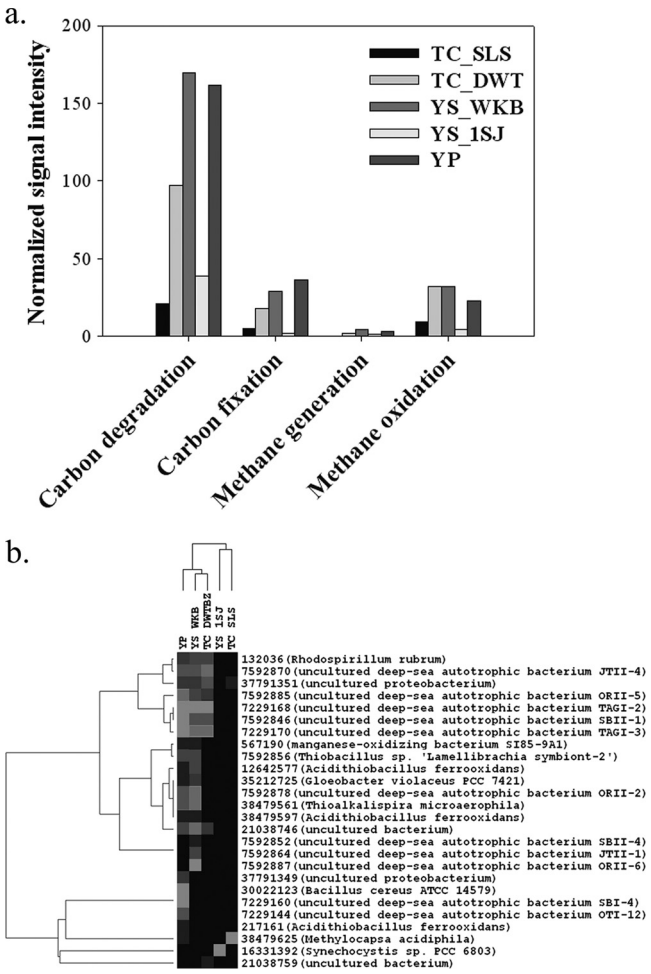

FIG. 4. Distribution of carbon cycling genes. (a) Normalized signal intensity of detected genes involved in carbon degradation, carbon fixation, methane generation, and methane oxidation. (b) Hierarchical clustering of carbon fixation genes. Details of hierarchical cluster analysis are described in the Fig. 1 legend.

Fe, Al, P, V, Y, Zr, and Sc were marginally significant ($P < 0.1$) (Table 3), suggesting that these metals play important roles in shaping microbial community functional structure in AMD.

Mantel tests were performed further to examine the relationships between various functional gene groups and environmental factors. Significant ($P < 0.05$) or marginally significant ($P < 0.10$) relationships between environmental factors and various functional gene groups were observed: methane generation ($r = 0.4899$, $P = 0.0319$), ammonification ($r = 0.5822$, $P = 0.0299$), denitrification ($r = 0.4332$, $P = 0.0779$), sulfite reduction ($r = 0.5685$, $P = 0.0789$), and contaminant degradation ($r = 0.6027$, $P = 0.019$) (see Table S5). All of the results described above indicate that AMD microbial communities and various functional gene groups are shaped largely by the surrounding environmental factors.

DISCUSSION

AMD represents a widespread environmental problem but also provides an ideal ecosystem model to better understand

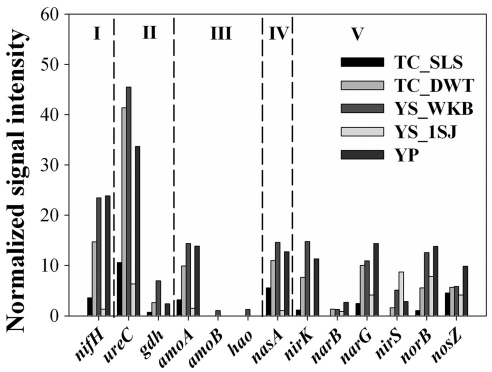

FIG. 5. Normalized signal intensity of functional genes involved in the nitrogen cycle. The percentages were calculated by dividing the total signal intensity for each group by the total intensity of all nitrogen-cycling genes detected by GeoChip 2.0. These genes are classified into five groups: I, nitrogen fixation, including *nifH* encoding nitrogenase reductase; II, ammonification, including *gdh* for glutamate dehydrogenase and *ureC* encoding urease; III, nitrification, including genes *amoA*, encoding ammonia monooxygenase, and *hao*, for hydroxylamine oxidoreductase; IV, assimilatory nitrate/nitrite reduction, including *nasA* encoding nitrate reductase; and V, denitrification, including *narG* for nitrate reductase, *nirS* for nitrite reductase, *norB* for nitric oxidereductase, *nosZ* for nitrous oxide reductase, and *nirK* encoding nitrite reductase with denitrification activity.

the structure and function of microbial communities and their linkages with environmental variables and ecosystem functioning. Several studies have detected the activity of or genes responsible for key functional processes in AMD or AMD-like environments. Denitrification has been shown to be actively

TABLE 3. Relationship of whole microbial community functional structure to individual environmental variables revealed by Mantel test[a]

Environmental variable	r_M[b]	P
pH	0.1773	0.4016
Temp (°C)	−0.3204	0.8092
S	0.5189	**0.046**
Fe	0.5462	**0.0719**
Al	0.4711	**0.0939**
Mg	0.6964	**0.018**
Ca	0.0546	0.4456
Zn	−0.3892	0.8621
Mn	0.1161	0.4465
Cu	0.6964	**0.0109**
Si	0.1366	0.3676
Na	−0.3209	0.8022
P	0.4506	**0.0999**
Ni	0.5189	**0.033**
Co	0.5189	**0.048**
V	0.4711	**0.0709**
B	0.5754	**0.042**
Y	0.5754	**0.0539**
La	0.6148	**0.035**
Zr	0.5089	**0.0709**
Sc	0.4996	**0.0789**
Ba	−0.1201	0.8132

[a] The signal intensity of all functional genes among five samples was used as the first matrix; the normalized environmental variables were used as the second matrix. Boldface values indicate significant P values (<0.1).

[b] r_M, Mantel's correlation coefficient.

present in AMD microbial communities (4). Genes for several functional processes, including carbon and nitrogen fixation, energy generation, protein refolding, oxidative stress, and iron oxidation, have been detected in a natural acidophilic biofilm (37, 45). In addition, a recent study has shown that AMD microbial communities are well adapted to their extreme habitats and can achieve primary production rates comparable to those of nonextreme environments (7). In this study, our analysis showed that most of the functional processes targeted by GeoChip 2.0, including carbon and nitrogen fixation, carbon degradation, methane metabolism, ammonification, nitrification, denitrification, nitrogen reduction, sulfur reduction, metal resistance and reduction, and organic contaminant degradation, were detected in all of the AMD samples examined in this study. Although only five samples from three copper mines were used for this study, they are typical AMDs and covered both surface and subsurface (at a depth of 160 m) environments. As such, results from this study should provide a more comprehensive view of the functional structure and potential activity of AMD microbial communities in general. While AMD generally is considered to be a simple system, this comprehensive survey detected almost all major metabolic processes found in other ecosystems, suggesting that the functional gene diversity and metabolic potential of AMD microbial communities is higher than previously thought.

AMD microbial communities contain a variety of microorganisms, including lithoautotrophs, organohetertrophs, lithoheterotrophs, and anaerobes, and a subset of these organisms must fix carbon and nitrogen to drive both carbon and nitrogen cycles. *Thiobacillus* species, including *Acidithiobacillus ferrooxidans*, are probably the dominant group of CO_2-fixing microorganisms at lower temperatures (<30°C) and higher pH (>2) (5). The samples examined in this study were from environments with similar temperature and pH ranges (<30°C, pH 2 to 3), and four *rbcL* genes derived from *Thiobacillus* spp. were detected. Another 24 *rbcL* genes also were detected, which were derived from cultured (6) and uncultured (16) organisms. In lower pH (<2.0), higher temperature (>30°C) communities, autotrophic taxa include *Leptospirillum* spp., *Ferroplasma* spp., *Sulfobacillus* spp., *Ferromicrobium* spp., and *Acidimicrobium* spp. (5), but no autotrophic taxa were detected in the low-temperature samples from the current study. The signal intensities of genes involved in methane oxidation and carbon degradation were similar to or higher than that for carbon fixation genes but were low for genes involved in methane generation in all samples.

The nitrogen cycle is one of the most well-known and important biogeochemical cycles, but little is known regarding how these processes are affected by extreme environments such as AMD (4). Although no nitrogen fixation has been directly observed in AMD systems (5), in this study, GeoChip detected 33 *nifH* sequences, 27 derived from uncultured organisms and 6 from cultured organisms, including *Calothrix* sp., which is able to fix nitrogen (19). This suggests that nitrogen fixation is a primary process in AMD systems, although the further functional assay of nitrogen fixation is needed. While AMD systems generally are exposed to O_2, which inhibits nitrogen fixation, a previous study showed that *A. ferrooxidans* might overcome this problem by using tetrathionate as an electron donor and ferric iron (rather than O_2) as an electron acceptor for nitrogen fixation (35). Also, denitrification is thought to be an important geochemical process in AMD systems (4). This is consistent with our GeoChip data, which showed that all six denitrification functional gene markers (e.g., *nirS*, *nirK*, *norB*, and *nosZ*) were detected. In addition, nitrification genes were detected, with two *amoA* genes derived from *Nitrosomonas* sp. strains C-113a and AL212 and two from *Nitrosococcus oceani*. Interestingly, a high abundance of genes involved in ammonification (*ureC* and *gdh*) was detected. Therefore, this study indicates that most of the functional genes involved in carbon and nitrogen cycling are present in AMD systems. However, the activity of these genes in AMD systems and the organisms involved need to be experimentally validated in the future.

AMD systems also are dominated by the oxidation and reduction of Fe and S (2, 5, 37, 45). Although sulfate reduction genes (*dsrA* and *dsrB*) were detected, GeoChip 2.0 does not contain any specific functional genes involved in iron oxidation or reduction or sulfur oxidation.

It appears that environmental factors (e.g., temperature, pH, S, and metals) and carbon and nitrogen sources have fundamental impacts on AMD microbial community structures. Previous studies have shown that pH, temperature, Fe, S, Pb, Zn, Cu, and Ca are important drivers shaping microbial community structure in AMD systems (5, 8, 58), and AMD isolates are able to grow at a range of temperatures (14 to 65°C), so a change in environment temperature would lead to changes in microbial community structure and function (5). For example, previous studies showed that the growth rate of *A. ferrooxidans* outcompeted that of *Leptospirillum ferrooxidans* only below 14°C (39), and that some *Sulfobacillus* isolates are able to grow at up to 65°C (3). pH also is considered an important factor affecting microbial community diversity and structure in AMD (5, 57). However, temperature and pH did not appear to be important environmental factors shaping the functional structure of AMD microbial communities in this study, most likely because these AMD samples had a similar temperature (<30°C) and pH range (2.0 ~ 3.0). In addition, previous studies indicated that Ca, Co, Na, Fe, Mg, P, and K were essential elements, and that Mn, Cr, Cu, and Zn were required as trace elements for cell growth in AMD (13, 32, 51). It is not surprising that S, Mg, Cu, Ni, Co, B, and La were significantly correlated with the functional structure of AMD microbial communities in this study.

In summary, both microbial functional and phylogenetic communities were examined in this study, and almost all genes involved in carbon, nitrogen, and sulfur cycling and metal resistance were detected by GeoChip 2.0, indicating that the functional diversity in AMD ecosystems is greater than previously thought. AMD microbial communities may be shaped largely by surrounding environmental factors. However, the DNA-based GeoChip analysis may detect only the functional potential of microbial communities. Like many studies using high-throughput technologies, GeoChip results provide rich sources of new hypotheses for potential community functions, especially novel functions. Similarly, GeoChip detected functional genes, but not necessarily the actual populations. For instance, some genes derived from alkaliphiles were detected, which suggests that the populations carrying those genes could exist in AMD systems, but it does not mean that they are

alkaliphiles. To validate those novel functional processes, populations, and/or activities, more systematic, in-depth analyses (e.g., shotgun metagenome sequencing and functional activity assays) are needed.

ACKNOWLEDGMENTS

This work was supported by the National Basic Research Program of China (no. 2004CB619201), the Funds for Creative Research Group of China (no. 50621063), the National Natural Science Foundation of China (no. 50321402 and 50774102), the Graduate Education Innovative Program of Central South University (no. 1343-77341), the U.S. Department of Energy, Office of Science, under the Environmental Remediation Science Program (DOE-ER64125), and the Genomics: GTL program through contract DE-AC02-05CH11231 between the Lawrence Berkeley National Laboratory and the U.S. Department of Energy.

REFERENCES

1. **Ahn, S. J., J. Costa, and J. R. Emanuel.** 1996. PicoGreen quantitation of DNA: effective evaluation of samples pre- or post-PCR. Nucleic Acids Res. **24:**2623–2625.
2. **Allen, E. E., and J. F. Banfield.** 2005. Community genomics in microbial ecology and evolution. Nat. Rev. Microbiol. **3:**489–498.
3. **Atkinson, T., et al.** 2000. A microbiological survey of Montserrat Island hydrothermal biotopes. Extremophiles **4:**305–313.
4. **Baeseman, J. L., R. L. Smith, and J. Silverstein.** 2006. Denitrification potential in stream sediments impacted by acid mine drainage: effects of pH, various electron donors, and iron. Microb. Ecol. **51:**232–241.
5. **Baker, B. J., and J. F. Banfield.** 2003. Microbial communities in acid mine drainage. FEMS Microbiol. Ecol. **44:**139–152.
6. **Banfield, J. F., N. C. Verberkmoes, R. L. Hettich, and M. P. Thelen.** 2005. Proteogenomic approaches for the molecular characterization of natural microbial communities. Omics **9:**301–333.
7. **Belnap, C. P., et al.** 2010. Cultivation and quantitative proteomic analyses of acidophilic microbial communities. ISME J. **4:**520–530.
8. **Bond, P. L., G. K. Druschel, and J. F. Banfield.** 2000. Comparison of acid mine drainage microbial communities in physically and geochemically distinct ecosystems. Appl. Environ. Microbiol. **66:**4962–4971.
9. **Bond, P. L., S. P. Smriga, and J. F. Banfield.** 2000. Phylogeny of microorganisms populating a thick, subaerial, predominantly lithotrophic biofilm at an extreme acid mine drainage site. Appl. Environ. Microbiol. **66:**3842–3849.
10. **Demergasso, C. S., et al.** 2005. Molecular characterization of microbial populations in a low-grade copper ore bioleaching test heap. Hydrometallurgy **80:**241–253.
11. **Dsa, J. V., K. S. Johnson, D. Lopez, C. Kanuckel, and J. Tumlinson.** 2008. Residual toxicity of acid mine drainage-contaminated sediment to stream macroinvertebrates: relative contribution of acidity versus metals. Water Air Soil Poll. **194:**185–197.
12. **Eisen, M. B., P. T. Spellman, P. O. Brown, and D. Botstein.** 1998. Cluster analysis and display of genome-wide expression patterns. Proc. Natl. Acad. Sci. U. S. A. **95:**14863–14868.
13. **Finney, L. A., and T. V. O'Halloran.** 2003. Transition metal speciation in the cell: Insights from the chemistry of metal ion receptors. Science **300:**931–936.
14. **Garrido, P., et al.** 2008. An oligonucleotide prokaryotic acidophile microarray: its validation and its use to monitor seasonal variations in extreme acidic environments with total environmental RNA. Environ. Microbiol. **10:**836–850.
15. **Gerhardt, A., L. J. de Bisthoven, K. Guhr, A. M. V. M. Soares, and M. J. Pereira.** 2008. Phytoassessment of acid mine drainage: Lemna gibba bioassay and diatom community structure. Ecotoxicology **17:**47–58.
16. **Goltsman, D. S. A., et al.** 2009. Community genomic and proteomic analyses of chemoautotrophic iron-oxidizing "*Leptospirillum rubarum*" (group II) and "*Leptospirillum ferrodiazotrophum*" (group III) bacteria in acid mine drainage biofilms. Appl. Environ. Microbiol. **75:**4599–4615.
17. **He, M., Z. Wang, and H. Tang.** 1997. Spatial and temporal patterns of acidity and heavy metals in predicting the potential for ecological impact on the Le An river polluted by acid mine drainage. Sci. Total Environ. **206:**67–77.
18. **He, Z., et al.** 2007. GeoChip: a comprehensive microarray for investigating biogeochemical, ecological and environmental processes. ISME J. **1:**67–77.
19. **Kallas, T., M. C. Rebiere, R. Rippka, and N. T. Demarsac.** 1983. The structural nif genes of the cyanobacteria Gloeothece sp. and Calothrix sp. share homology with those of Anabaena sp., but the Gloeothece genes have a different arrangement. J. Bacteriol. **155:**427–431.
20. **Karnachuk, O. V., et al.** 2009. Bacteria of the sulfur cycle in the sediments of gold mine tailings, Kuznetsk Basin, Russia. Microbiology **78:**483–491.
21. **Kim, J., et al.** 2009. Influence of acid mine drainage on microbial communities in stream and groundwater samples at Guryong Mine, South Korea. Environ. Geol. **58:**1567–1574.
22. **Kimes, N. E., J. D. Van Nostrand, E. Weil, J. Z. Zhou, and P. J. Morris.** 2010. Microbial functional structure of Montastraea faveolata, an important Caribbean reef-building coral, differs between healthy and yellow-band diseased colonies. Environ. Microbiol. **12:**541–556.
23. **Krebs, W., C. Brombacher, P. P. Bosshard, R. Bachofen, and H. Brandl.** 1997. Microbial recovery of metals from solids. FEMS Microbiol. Rev. **20:**605–617.
24. **Lear, G., D. Niyogi, J. Harding, Y. M. Dong, and G. Lewis.** 2009. Biofilm bacterial community structure in streams affected by acid mine drainage. Appl. Environ. Microbiol. **75:**3455–3460.
25. **Leigh, M. B., et al.** 2007. Biphenyl-utilizing bacteria and their functional genes in a pine root zone contaminated with polychlorinated biphenyls (PCBs). ISME J. **1:**134–148.
26. **Levings, C. D., et al.** 2004. Effects of acid mine drainage on the estuarine food web, Britannia Beach, Howe Sound, British Columbia, Canada. Hydrobiologia **525:**185–202.
27. **Li, X. F., and M. Sasaki.** 2007. Hydrothermal alteration and mineralization of middle Jurassic Dexing porphyry Cu-Mo deposit, southeastern China. Resource Geol. **57:**409–426.
28. **Liang, Y. T., et al.** 2009. Microarray-based analysis of microbial functional diversity along an oil contamination gradient in oil field. FEMS Microbiol. Ecol. **70:**324–333.
29. **Mahmouda, K. K., L. G. Leducb, and G. D. Ferroni.** 2005. Detection of *Acidithiobacillus ferrooxidans* in acid mine drainage environments using fluorescent in situ hybridization (FISH). J. Microbiol. Methods **61:**33–45.
30. **Mantel, N.** 1967. Detection of disease clustering and a generalized regression approach. Cancer Res. **27:**209–220.
31. **Mason, O. U., et al.** 2009. Prokaryotic diversity, distribution, and insights into their role in biogeochemical cycling in marine basalts. ISME J. **3:**231–242.
32. **Morel, F. M. M., and N. M. Price.** 2003. The biogeochemical cycles of trace metals in the ocean. Science **300:**944–947.
33. **Nordstrom, D. K., and C. N. Alpers.** 1999. Negative pH, efflorescent mineralogy, and consequences for environmental restoration at the Iron Mountain superfund site, California. Proc. Natl. Acad. Sci. U. S. A. **96:**3455–3462.
34. **Nordstrom, D. K., C. N. Alpers, C. J. Ptacek, and D. W. Blowes.** 2000. Negative pH and extremely acidic mine waters from Iron Mountain, California. Environ. Sci. Technol. **34:**254–258.
35. **Norris, P. R., J. C. Murrell, and D. Hinson.** 1995. The potential for diazotrophy in iron-oxidizing and sulfur-oxidizing acidophilic bacteria. Arch. Microbiol. **164:**294–300.
36. **Qiu, G. Z., et al.** 2008. Archaeal diversity in acid mine drainage from Dabaoshan Mine, China. J. Basic Microbiol. **48:**401–409.
37. **Ram, R. J., et al.** 2005. Community proteomics of a natural microbial biofilm. Science **308:**1915–1920.
38. **Rodríguez-Martínez, E. M., E. X. Perez, C. W. Schadt, J. Zhou, and A. A. Massol-Deya.** 2006. Microbial diversity and bioremediation of a hydrocarbon-contaminated aquifer (Vega Baja, Puerto Rico). Int. J. Environ. Res. Public Health **3:**292–300.
39. **Sand, W., K. Rohde, B. Sobotke, and C. Zenneck.** 1992. Evaluation of Leptospirillum ferrooxidans for leaching. Appl. Environ. Microbiol. **58:**85–92.
40. **Schippers, A., A. A. Nagy, D. Kock, F. Melcher, and E.-D. Gock.** 2008. The use of FISH and real-time PCR to monitor the biooxidation and cyanidation for gold and silver recovery from a mine tailings concentrate (Ticapampa, Peru). Hydrometallurgy **94:**77–81.
41. **Sokal, R. R., and F. J. Rohlf.** 1994. Biometry: the principles and practices of statistics in biological research, 3rd ed. W. H. Freeman, New York, NY.
42. **Tas, N., et al.** 2009. Tracking functional guilds: "Dehalococcoides" spp. in European river basins contaminated with hexachlorobenzene. Appl. Environ. Microbiol. **75:**4696–4704.
43. **Tiwary, R. K.** 2001. Environmental impact of coal mining on water regime and its management. Water Air Soil Poll. **132:**185–199.
44. **Tremaine, S. C., and A. L. Mills.** 1991. Impact of water column acidification on protozoan bacterivory at the lake sediment-water interface. Appl. Environ. Microbiol. **57:**775–784.
45. **Tyson, G. W., et al.** 2004. Community structure and metabolism through reconstruction of microbial genomes from the environment. Nature **428:**37–43.
46. **Van Nostrand, J. D., et al.** 2009. GeoChip-based analysis of functional microbial communities during the reoxidation of a bioreduced uranium-contaminated aquifer. Environ. Microbiol. **11:**2611–2626.
47. **Waldron, P. J., et al.** 2009. Functional gene array-based analysis of microbial community structure in groundwaters with a gradient of contaminant levels. Environ. Sci. Technol. **43:**3529–3534.
48. **Wang, F. P., et al.** 2009. GeoChip-based analysis of metabolic diversity of microbial communities at the Juan de Fuca Ridge hydrothermal vent. Proc. Natl. Acad. Sci. U. S. A. **106:**4840–4845.
49. **Wu, L., L. Kellogg, A. H. Devol, J. M. Tiedje, and J. Zhou.** 2008. Microarray-based characterization of microbial community functional structure and heterogeneity in marine sediments from the gulf of Mexico. Appl. Environ. Microbiol. **74:**4516–4529.
50. **Wu, L. Y., X. Liu, C. W. Schadt, and J. Z. Zhou.** 2006. Microarray-based

analysis of subnanogram quantities of microbial community DNAs by using whole-community genome amplification. Appl. Environ. Microbiol. **72:**4931–4941.
51. **Xie, B., D. Xi, and J. Chen.** 2002. Effects of heavy metals on activated sludge microorganism. J. DongHua Univ. **19:**132–137.
52. **Xie, X. H., S. M. Xiao, and J. S. Liu.** 2009. Microbial communities in acid mine drainage and their interaction with pyrite surface. Curr. Microbiol. **59:**71–77.
53. **Yergeau, E., S. Kang, Z. He, J. Zhou, and G. A. Kowalchuk.** 2007. Functional microarray analysis of nitrogen and carbon cycling genes across an Antarctic latitudinal transect. ISME J. **1:**163–179.
54. **Yin, H., et al.** 2008. Molecular diversity of 16S rRNA and gyrB genes in copper mines. Arch. Microbiol. **189:**101–110.
55. **Yin, H., et al.** 2008. Bacterial diversity based on 16S rRNA and gyrB genes at Yinshan mine, China. Syst. Appl. Microbiol. **31:**302–311.
56. **Yin, H. Q., et al.** 2007. Development and evaluation of 50-mer oligonucleotide arrays for detecting microbial populations in acid mine drainages and bioleaching systems. J. Microbiol. Methods **70:**165–178.
57. **Yin, H. Q., et al.** 2008. Microbial community diversity and changes associated with a mine drainage gradient at the Dexing copper mine, China. Aquat. Microb. Ecol. **51:**67–76.
58. **Zhang, D., G. Xu, W. Zhang, and S. D. Golding.** 2007. High salinity fluid inclusions in the Yinshan polymetallic deposit from the Le-De metallogenic belt in Jiangxi Province, China: their origin and implications for ore genesis. Ore Geol. Rev. **31:**247–260.
59. **Zhou, J., M. A. Bruns, and J. M. Tiedje.** 1996. DNA recovery from soils of diverse composition. Appl. Environ. Microbiol. **62:**316–322.
60. **Zhou, J. Z., S. Kang, C. W. Schadt, and C. T. Garten.** 2008. Spatial scaling of functional gene diversity across various microbial taxa. Proc. Natl. Acad. Sci. U. S. A. **105:**7768–7773.