



Acid mine drainage affects the diversity and metal resistance gene profile of sediment bacterial community along a river

Xiaohui Zhang^{a, b, c}, Song Tang^{d, e, f, *}, Mao Wang^{g, **}, Weimin Sun^h, Yuwei Xie^{a, i}, Hui Peng^j, Aimin Zhong^k, Hongling Liu^{a, b, c, ***}, Xiaowei Zhang^{a, b, c}, Hongxia Yu^{a, b, c}, John P. Giesy^{a, i}, Markus Hecker^{f, i}

^a State Key Laboratory of Pollution Control and Resource Reuse, School of the Environment, Nanjing University, Nanjing, Jiangsu 210023, China

^b Jiangsu Key Laboratory of Environmental Safety and Health Risk of Chemicals, Nanjing, Jiangsu 210023, China

^c Research Center for Environmental Toxicology & Safety of Chemicals, Nanjing University, Nanjing, Jiangsu 210023, China

^d National Institute of Environmental Health, Chinese Center for Disease Control and Prevention, Beijing 100021, China

^e Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing, Jiangsu 211166, China

^f School of Environment and Sustainability, University of Saskatchewan, Saskatoon, SK S7N 5C3, Canada

^g School of Public Health, Sun Yat-sen University, Guangzhou, Guangdong 510080, China

^h Guangdong Key Laboratory of Integrated Agro-environmental Pollution Control and Management, Guangdong Institute of Eco-environmental Science & Technology, Guangzhou 510650, China

ⁱ Toxicology Centre, University of Saskatchewan, Saskatoon, SK S7N 5B3, Canada

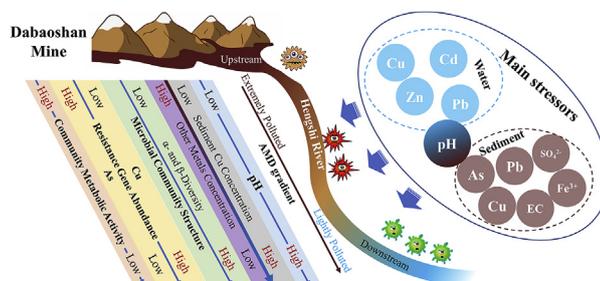
^j Department of Chemistry and School of the Environment, University of Toronto, Toronto, Ontario M5S 3H6, Canada

^k The Centre for Disease Control and Prevention of Wengyuan County, Shaoguan, Guangdong 512600, China

HIGHLIGHTS

- Diversities of bacterial community reflected changes in geochemical conditions.
- The major factor that influenced compositions of bacterial communities was sediment electricity conductivity.
- A decreasing trend of metabolic genes abundance with decreasing pollution.
- Metal concentrations and abundance of corresponding resistance genes are relevant.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 7 May 2018

Received in revised form

28 October 2018

Accepted 29 October 2018

Available online 1 November 2018

Handling Editor: Martine Leermakers

ABSTRACT

Acid mine drainage (AMD) is one of the most hazardous byproducts of some types of mining. However, research on how AMD affects the bacterial community structure of downstream riverine ecosystems and the distribution of metal resistance genes (MRGs) along pollution gradient is limited. Comprehensive geochemical and high-throughput next-generation sequencing analyses can be integrated to characterize spatial distributions and MRG profiles of sediment bacteria communities along the AMD-contaminated Hengshi River. We found that (1) diversities of bacterial communities significantly and gradually increased along the river with decreasing contamination, suggesting community composition reflected

* Corresponding author. No.7 Panjiayuan Nanli, Chaoyang District, Beijing 100021, China.

** Corresponding author. No.74 Zhongshan 2nd Road, School of Public Health, Sun Yat-Sen University, Yuexiu District, Guangzhou, Guangdong 510080, China.

*** Corresponding author. School of the Environment, Nanjing University, Nanjing, Jiangsu 210023, China.

E-mail addresses: tangsong@nieh.chinacdc.cn (S. Tang), wangmao@mail.sysu.edu.cn (M. Wang), hlliu@nju.edu.cn (H. Liu).

Keywords:

AMD gradient
Daobaoshan mine
16S rDNA sequencing
Comparative metagenomics
Electrical conductivity
pH

changes in geochemical conditions; (2) relative abundances of phyla *Proteobacteria* and genus *Halomonas* and *Planococcaceae* that function in metal reduction decreased along the AMD gradient; (3) low levels of sediment salinity, sulfate, aquatic lead (Pb), and cadmium (Cd) were negatively correlated with bacterial diversity despite pH was in a positive manner with diversity; and (4) arsenic (As) and copper (Cu) resistance genes corresponded to sediment concentrations of As and Cu, respectively. Altogether, our findings offer initial insight into the distribution patterns of sediment bacterial community structure, diversity and MRGs along a lotic ecosystem contaminated by AMD, and the factors that affect them.

© 2018 Elsevier Ltd. All rights reserved.

1. Introduction

Microorganisms living in freshwater sediment are regarded as the main transformers of biogeochemical nutrients and contaminants at the water-sediment interface of rivers and streams (Madsen, 2011; Ruiz-González et al., 2015). Yet, in the case of open lotic systems such as rivers, sediment microbial communities are characterized by significant fluctuations in absolute and relative numbers of taxa (Read et al., 2015). Abundance and diversity of microbes are closely linked to temporal and spatial changes in physicochemical (e.g. pH, temperature and salinity) and biotic variables (e.g. abundances of zooplankton and heterotrophic nanoflagellates), nutrients (e.g. dissolved organic carbon, nitrogen and phosphorus) and pollutants (e.g. pesticides and metals) in the environment. Therefore, composition of a microbial community can be employed to assess local environmental and regional conditions, and serve as a sensitive bio-indicator of pollution (Souffreau et al., 2015; Ibekwe et al., 2016; Sun et al., 2016).

Acid mine drainage (AMD) refers to acidic runoff originating from active or abandoned mining sites and is a global environmental problem. The low pH and high concentrations of sulfate and dissolved toxic metals in AMD are significant threats to the surrounding environments. However, AMD can promote opportunities for certain forms of life. Some metabolically active microbes are well adapted to extremely toxic and acidic environments (Johnson and Hallberg, 2008) and have significant potential for AMD bioremediation (Johnson and Hallberg, 2005). These microbes have developed a variety of ways to cope with excessive metal-loaded niches, including efflux-mediated metal transport, metal exclusion by permeability barrier, intracellular and extracellular sequestration, and enzymatic detoxification (Denef et al., 2010).

Recent advances in high-throughput sequencing (HTS) enable in-depth coverage of occurrence, diversity, distribution and interaction patterns of microbial consortia in AMD ecosystems (Gonzalez-Toril et al., 2003). A wide variety of acidophilic and metal-tolerant microorganisms have been identified and their metabolic capacities and functions have been characterized (Baker and Banfield, 2003). High-throughput metagenomics is also regarded as a powerful way to unveil the gene pool such as the occurrence of metal resistance genes (MRGs) in complex communities, revealing essential biological processes and resistant gene elements as survival strategies under an extreme condition in a high-resolution and culture-independent manner. These properties had been used in environmental monitoring and assessment, from which the results demonstrated that the activities of MRGs expressed in microbes were affected by AMD (Chen et al., 2015a). However, to date, few studies have explored the spatial dynamics in diversity of microorganisms and abundance distribution of MRGs along the pollution gradient of an AMD-impacted riverine ecosystem.

The Hengshi River provides an excellent opportunity to address the above issue. It is located (approximately 25 km) southeast of Shaoguan, Guangdong Province, China, and was mainly used as a

source for drinking water and agricultural irrigation (Zhou and Xia, 2010). In 1958, a large-scale and open-pit multi-metal sulphide Dabaoshan mine (DBS) was built upstream of the river and has been fully operational since the 1970s. The mine discharges large quantities of acidic (pH~2.5) and metallic runoff containing copper (Cu), cadmium (Cd), lead (Pb), and zinc (Zn) as well as sulfate (SO_4^{2-}) into the Hengshi River (Chen et al., 2007, 2015c; Lin et al., 2007). Compared to the extremely contaminated upstream (pH~2.5), downstream areas are moderately to minimally contaminated (pH~6). This mining area has attracted global attention because it has caused adverse effects to surrounding ecosystems and irreversible health damages to local residents (Wang et al., 2011; Larson, 2014). Since the composition of AMD is remarkably variable among sites, depending on multiple factors, treatment solutions for remediation are generally location-specific. Developing efficient bioremediation strategies *in situ* for Hengshi River requires a comprehensive understanding of the geochemical factors influencing community structures and the metabolic potentials of indigenous microbial communities.

With the present work, geochemical analytical and HTS approaches were combined to systematically unravel differences in composition and abundance distribution of MRGs of indigenous bacterial communities along the continuously decreasing AMD gradient of Hengshi River. Specifically, we hypothesize (1) an increase in diversity of sediment bacteria community along the AMD gradient; (2) the upstream has more tolerant assemblages with functions in acid resistance, metal reduction and sulfate cycling than the downstream does; (3) changes in the structure of the bacterial community according to the variations of certain geochemical variables along the river; and (4) habitat-specific functional fingerprints (relative abundance of MRGs in bacterial community) correspond to the characteristics of the sampled environment niche (concentrations of metals).

2. Materials and methods

2.1. Samples collection and processing

The DBS mine (24°34'28" N; 113°43'42" E) and the Hengshi and Reference Rivers are located in Wengyuan, Shaoguan City, Guangdong Province, China (Fig. 1). This area has a subtropically humid monsoon climate, with an annual average temperature of 20 °C and precipitation of 1800 mm (Zhao et al., 2012). The DBS Mine is the largest meso-hypothermal deposit of polymetallic sulfide in South China. Over the 60 years the mine has been in operation, waste has accumulated and a dam was built across the valley to intercept floodwaters and retain mud that was transported from stockpiles of waste rock, forming a 12.81 km² lake (AMD pollution source) (Lin et al., 2007; Chen et al., 2015c). Acidic water overflows the dam at a rate of 0.01–0.12 m³/s into Hengshi River throughout the year (Chen et al., 2015c).

A total of 27 sediment and 27 water samples were collected in

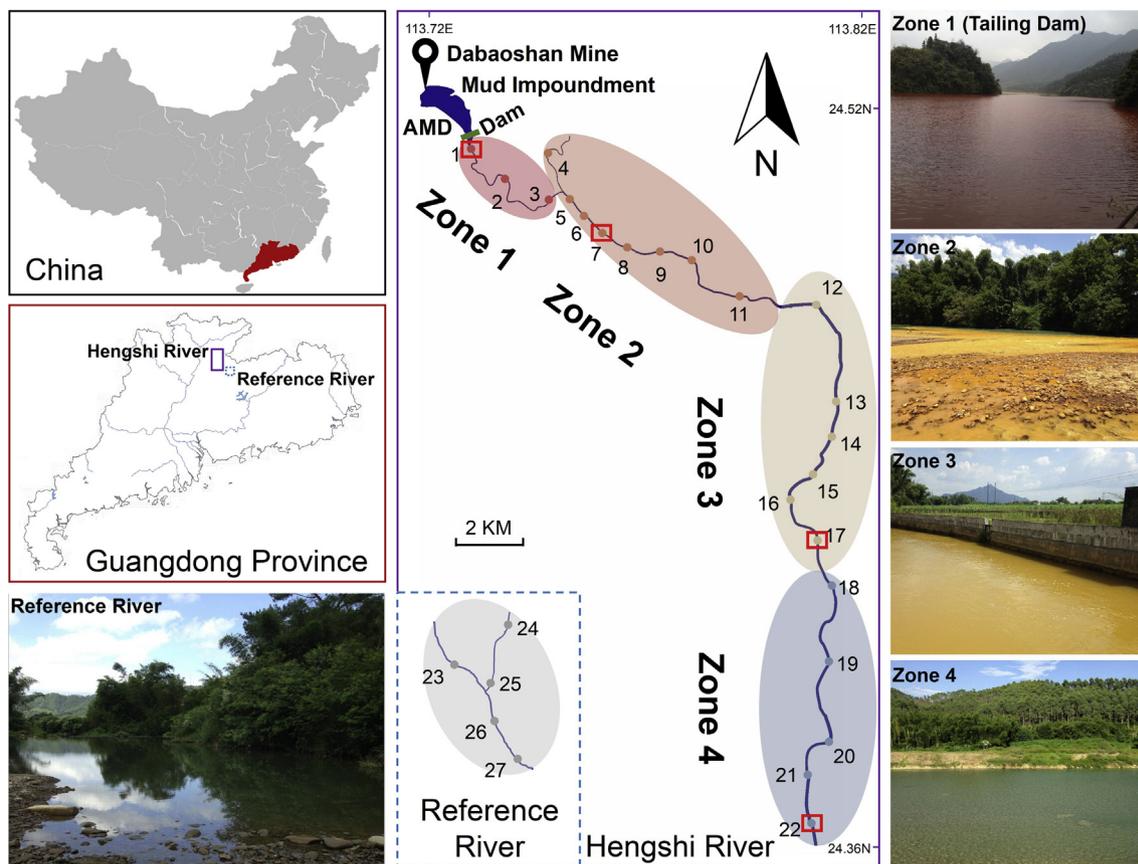


Fig. 1. Locations and representative field photos of sampling sites in Dabaoshan Mine, mud impoundment, tailing dam (AMD pollution source), AMD-contaminated Hengshi River, and Reference River in Guangdong Province, China. Sampling sites of Hengshi River were divided into four zones based on geodistance and AMD pollution levels. Zone 1-Extremely polluted (1–3); Zone 2-Heavily polluted (4–11); Zone 3-Moderately polluted (12–17); and Zone 4-Lightly polluted (18–22). Five uncontaminated samples (23–27) were collected from a Reference River near Hengshi River. In addition, four sediment samples 1, 7, 17 and 22 were selected from four zones, and sediment was collected in October 2015 for metagenomics analysis. For comparative metagenomics, four typical sediments (1, 7, 17 and 22 in the red box) were selected and sequenced to represent four different zones. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

October 2014. According to the distance to the tailing dam, 22 sampling sites along Hengshi River were divided into four zones: Zone 1-Extremely polluted (1–3); Zone 2-Heavily polluted (4–11); Zone 3-Moderately polluted (12–17); and Zone 4-Lightly polluted (18–22). Five samples (23–27) were collected from an uncontaminated Qingyunshan River. In addition, four sediment samples 1, 7, 17 and 22 were selected from four zones, and sediment was collected in October 2015 for metagenomics analysis. Triplicate surface sediment (top~5 cm, approximately 500 g) was collected per site using a grab sampler (HYDRO-BIOS Apparatebau GmbH, Kiel-Holtenua, Germany). Sediment from each site was pooled, homogenized, and stored in sterile polypropylene bags and placed immediately on dry ice. Water samples were also collected and stored in 500 mL sterile tubes on dry ice. All samples were returned to the laboratory immediately and kept at -80°C until processing.

2.2. Geochemical analysis

Analytical methods of sediment grain size, hydrogen peroxidase (HP), electrical conductivity (EC), pH, total organic carbon (TOC), metals, and sulfate (SO_4^{2-}) are provided in [Supporting Information](#).

2.3. DNA extraction, sequencing and data processing

Genomic DNA was extracted from 0.25 g of homogenized sediment using the MoBio PowerSoil DNA Kit (MoBio Laboratories Inc., Carlsbad, CA, USA) according to the manufacturer's instructions.

The modified primer set 314F/518R (Klindworth et al., 2012) was used to amplify 200 bp of the V3 region of bacterial 16S rRNA genes. Libraries were built on an Ion Torrent Personal Genome Instrument in-house. Low quality reads and sequence adaptors were trimmed and then analyzed with the Quantitative Insights Into Microbial Ecology (QIIME) toolkit (Caporaso et al., 2010) and UPARSE pipeline (Edgar, 2013). A representative sequence was chosen from each operational taxonomic units (OTUs) and classified by operating a Ribosomal Database Project (RDP) Classifier (Wang et al., 2007) with a confidence greater than 80% against Greengene database (DeSantis et al., 2006). Detailed methods for PCR amplification, sequencing and bioinformatic analysis are provided in [Supporting Information](#).

2.4. Comparative metagenomics

Shotgun libraries were constructed and paired-end sequencing was performed on an Illumina HiSeq-2500 platform at Novogene (Beijing, China). Approximately 5 Gb of raw sequence data was generated from each library. After filtering and removing sequence segments that were shorter than 500 bp, high-quality scaffolds were used for gene function prediction. Open Reading Frame (ORF) prediction, redundancy removal, gene catalogue alignment and filtering of low quality reads were performed successively on scaffolds to get gene catalogues (unigenes) for function annotation. Unigenes were compared against the KEGG, COG and BacMat (<http://bacmet.biomedicine.gu.se/index.html>) databases using

DIAMOND (blastp, $\text{evalue} \leq 10^{-5}$) (Buchfink et al., 2015). Detailed metagenomic analysis is provided in Supporting Information. Raw sequences of 16S and metagenomics have been made available in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (Accession SRP094407).

2.5. Statistical analysis

Statistical analyses were completed using R software v3.3.1. Alpha and beta diversities were estimated using the “phyloseq” (McMurdie and Holmes, 2013) package. Differences in measured endpoints among different zones and between Hengshi and Reference Rivers were analyzed using a Kruskal-Wallis rank sum test. Differences were considered significant when $p < 0.05$. Random Forests (RF) analysis was employed to determine the importance of each geochemical parameter on alpha diversity (Chao1). Detailed statistical analysis is provided in Supporting Information.

3. Results

3.1. Geochemical analysis along the river

Concentrations of geochemical parameters varied significantly along Hengshi River as well as between Hengshi and Reference Rivers (Tables S1 and S3 and Fig. 2). Among all four zones, the results of downstream Zone 4 were the most similar to those of the Reference River. A significant increase in pH was observed along the river, while a significant decrease in electrical conductivity (Sd. EC), total water or sediment concentrations of As, Cd, Zn and Pb, and sediment concentrations of total Fe (Sd.Total.Fe), ferrous (Sd.Ferric), and SO_4^{2-} (Sd.Sulfate) were found upstream compared to those

found downstream (Fig. 2). However, contrary to the decreasing Cu concentration in the water, a significant upward Cu concentration in sediments was found along the AMD gradient.

3.2. Bacterial community composition and diversity along the river

A total of 696,071 reads were obtained for all 27 samples and clustered into 5276 OTUs. 685 OTUs were shared among four zones (Fig. S1A). Of all classifiable sequences, 10 major phyla were identified. *Firmicutes*, *Proteobacteria* and *Actinobacteria* were the predominant phyla accounting for 40.59%, 21.67% and 19.37% of all reads, respectively (Figs. S1B and S2). However, the proportion of these phyla differed among sites. Other phyla such as *Actinobacteria*, *Acidobacteria*, *Bacteroidetes*, and *Nitrospirae* that have been reported in AMD environments were also detected. Among these phyla, relative abundances of *Proteobacteria* decreased gradually along the river (Fig. S1B).

At the genus level, relative abundances of *Halomonas*, *Planococcaceae* and *Bacillales* decreased with decreasing pollution gradients, while *Clostridium* exhibited an opposite trend with AMD gradients (Fig. S2). Co-occurrence network analysis generated both positive and negative correlations within all genera (relative abundance > 0.5%) of Hengshi River (Fig. S3A), which resulted in an interaction network consisting of 55 nodes and 236 edges. Four major modules were identified: Mod 1 included *Acidocella*, *Acidiphilium*, *Lactobacillus*, and *Gallionella*; Mod 2 included *Propionisimonas*, *Rhodococcus*, *Caloramator*, *Paenibacillus*, *Mycobacterium*, *Streptomyces* and *Desulfosporosinus*; Mod 3 included *Corynebacterium*, *Staphylococcus*, *Lactococcus*, *Idiomarina*, *Halomonas*, *Comamonas*, and *Flavobacterium*; and Mod 4 included *Arthronema*, *Hydrogenophaga*, *Sediminibacterium*, *Rhodoplanes*, *Opitutus*, *Methylobacterium*, *Leptolyngbya*, and *Pseudanabaena*. In addition,

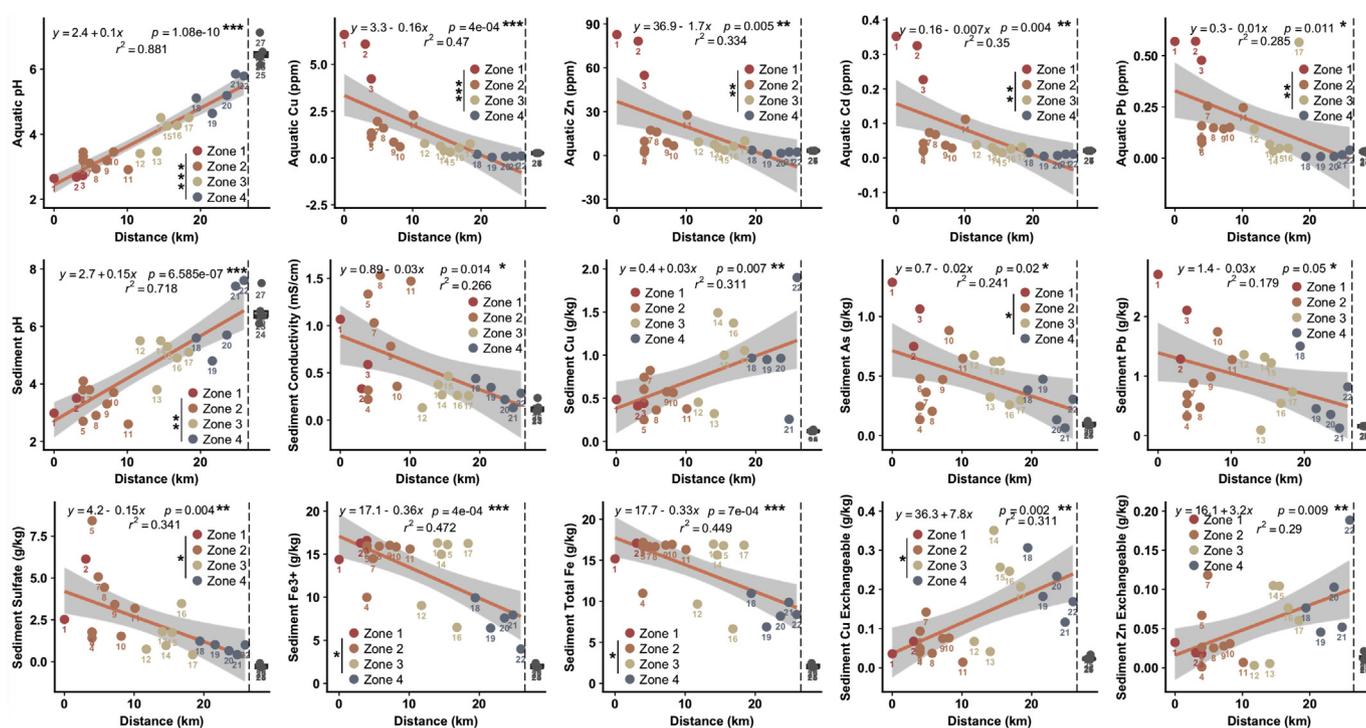


Fig. 2. Significant changes in key physiochemical parameters along Hengshi River. Linear regressions between the physiochemical parameters of water and sediment samples from Hengshi River and the geodistance to the AMD pollution source (km) are given. Shaded areas are the 95% confidence intervals (95% CI) for each model. The equation, adjusted r^2 and p values for the specific linear regressions are given in each panel. Different colors of the dots indicated the samples from different zones. The Kruskal-Wallis test was performed among different zones. Significance was determined at p values $< 0.05^*$, $< 0.01^{**}$, and $< 0.001^{***}$. Whisker boxplot (in the right of each panel) shows the values of the samples from Reference River. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Hydrogenophaga, *Lactobacillus*, *Meiothermus*, *Acidocella*, *Sediminibacterium*, *Thiomonas*, *Acidiphilium*, *Prevotella*, *Flavobacterium*, *Comamonas*, and *Halomonas* significantly predominated in Hengshi sediments compared to Reference sediments (Fig. 53B).

The least alpha diversity (Chao1) was observed in Zone 1, and diversities significantly and gradually increased with the decreasing AMD gradients (Fig. 3A). This trend was further supported by beta diversity (Fig. 3B) as bacterial communities from each zone exhibited significantly different patterns of composition.

3.3. Relationships between geochemical parameters and bacterial community

We hypothesized that the variations of bacterial communities along the Hengshi River were caused by innate geochemical parameters. The RF model was employed to determine the relative importance of geochemical parameters to community variation (Sun et al., 2017). The top five factors that affected the Chao1 index were Sd.EC, Sd.pH, Aq.Pb, Sd.Sulfate and Aq.Cd (Fig. 4A). Partial dependence plots (Fig. 4B) revealed that diversity dramatically decreased once their concentrations increased from 0 to certain threshold values (~0.5 mS/cm for Sd.EC, ~0.2 mg/L for Aq.Pb, ~2000 mg/kg for Sd.Sulfate, and ~0.08 mg/L for Aq.Cd). There are no conspicuous effects on alpha diversity above the thresholds as indicated by the horizontal line. However, diversity sharply increased when pH > 3.

3.4. Comparative metagenomics among four zones

In order to comprehensively compare distributions of genes and their abundances in various functional categories, assembled sequences in four metagenome libraries were searched against KEGG and BacMet databases. Functional abundance analysis of KEGG showed Zone 1 had a greater relative abundance (about 1.5-fold) of gene families involved in the Metabolism category compared to

that of the other three zones, and a decreasing trend was observed with decreasing pollution (Fig. 5A). Annotations of BacMet showed predominant MARGs were for resistance toward As, Cu and Zn, and their distribution patterns varied among all four zones. More specifically, bacterial communities upstream contained more genes for As detoxification/resistance (e.g. *arsB*, *arsC*, *arsM* and *arsR*), while downstream harbored more genes involved in Cu detoxification (e.g. *copA*, *copB*, *copR*, *copS*, *cusA/ybdE*, *dnaK*, *crdA*, *baeR*, *baeS*, *corR* and *corS*) (Fig. 5B). Most of As and Cu resistance genes were contributed by *Acidobacteria*, *Actinobacteria* and *Proteobacteria*, and the contribution ratio of each phyla varied greatly among the zones (Fig. 5C). Overall, sequences related to As and Cu resistance were mostly attributed to *Proteobacteria* in Zones 1 and 2, while many other phyla dominated in Zone 4. The abundance of As- and Cu-related genes shifted in the same pattern along the Sd.As and Sd.Cu gradient, indicating the adaptation of bacterial communities was in accordance to the contaminate condition.

4. Discussion

We aimed to characterize the spatial dynamics of the structure and function of bacterial communities and their response to AMD contamination along a riverine ecosystem. The elucidation of bacterial population dynamics at a diverse array of polluted sites improves predictive power to the diversity patterns in AMD-impacted ecosystems. RF model was employed to quantify the importance of geochemical parameters to community variation. Moreover, comparative metagenomics were applied to investigate the adaptive strategies and metabolic capabilities of communities in different AMD gradients. Various microbes thrived in contaminated areas, and positive correlations of MARGs with relevant metals suggest a functional potential for metal biotransformation and resistance by local bacterial communities.

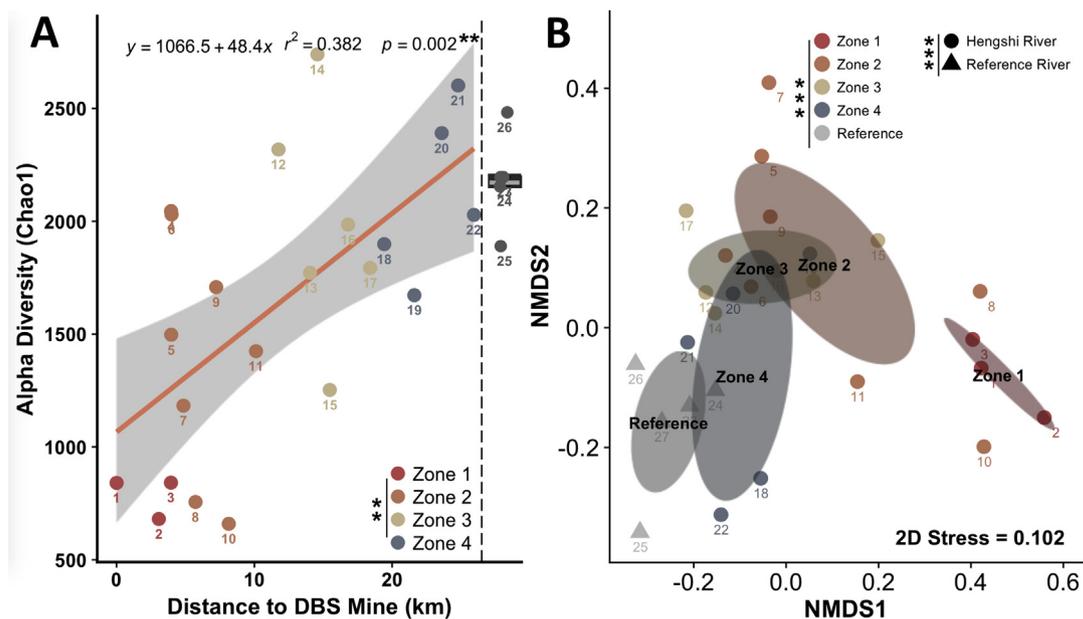


Fig. 3. (A) Alpha diversity (Chao1) along the Hengshi River. Linear regressions between alpha diversity and geodistance to AMD pollution source (km) are given. Shaded areas are 95% CI for each model. The equation, adjusted r^2 and p values for each linear regression are given in each panel. The Kruskal-Wallis test was performed among zones. Significance was determined at p values < 0.01**. Whisker boxplot shows the alpha diversity of Reference River. (B) Beta diversity of bacterial OTUs shifts along the Hengshi River. Non-metric multidimensional scaling (NMDS) visualizes the relative dissimilarities (Bray-Curtis) of normalized read counts among different zones. Each point corresponds to bacterial reads of a sample and ovals represent the 95% confidence ellipse around the centroids of zones. Permutational multivariate analysis of variance test (PERMANOVA, $n = 9999$) further showed that bacterial OTUs between the Hengshi and Reference Rivers and among four zones are significantly different (Adonis test-Rivers $p = 4e-04^{***}$, -Zones $p = 2e-04^{***}$).

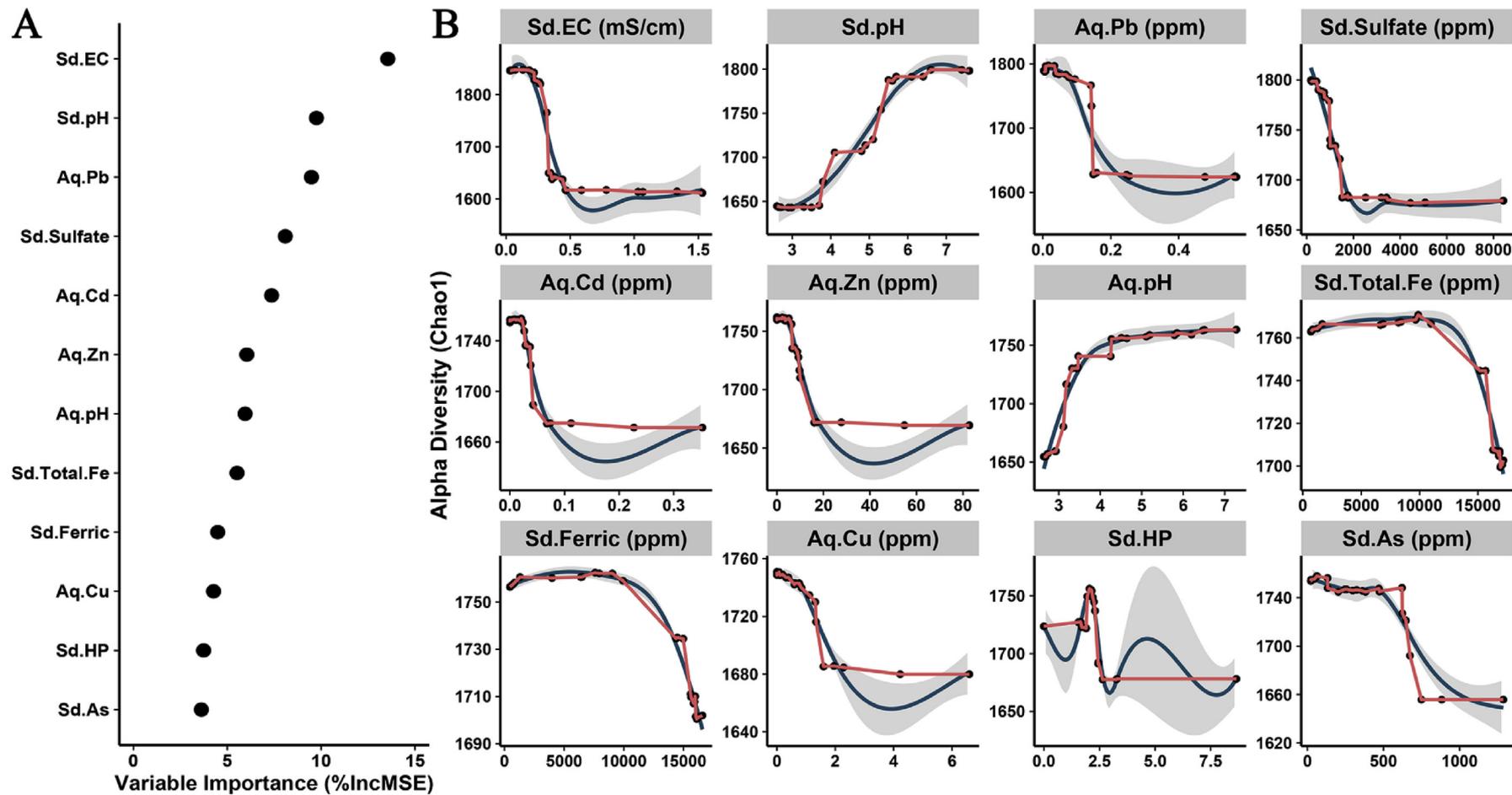


Fig. 4. (A) Variable importance of physiochemical parameters on the alpha diversity (Chao1) as determined by the Random Forest (RF) model. (B) Partial dependence plots of physiochemical parameters for RF predictions of Chao1. Partial dependence of a given parameter is the dependence of the probability after averaging out the influence of other parameters in the model. Red line is the partial dependence data line. Blue line with shaded areas is the local polynomial regression fitting trend line (LOESS). Shaded areas are the 95% CI for each model. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

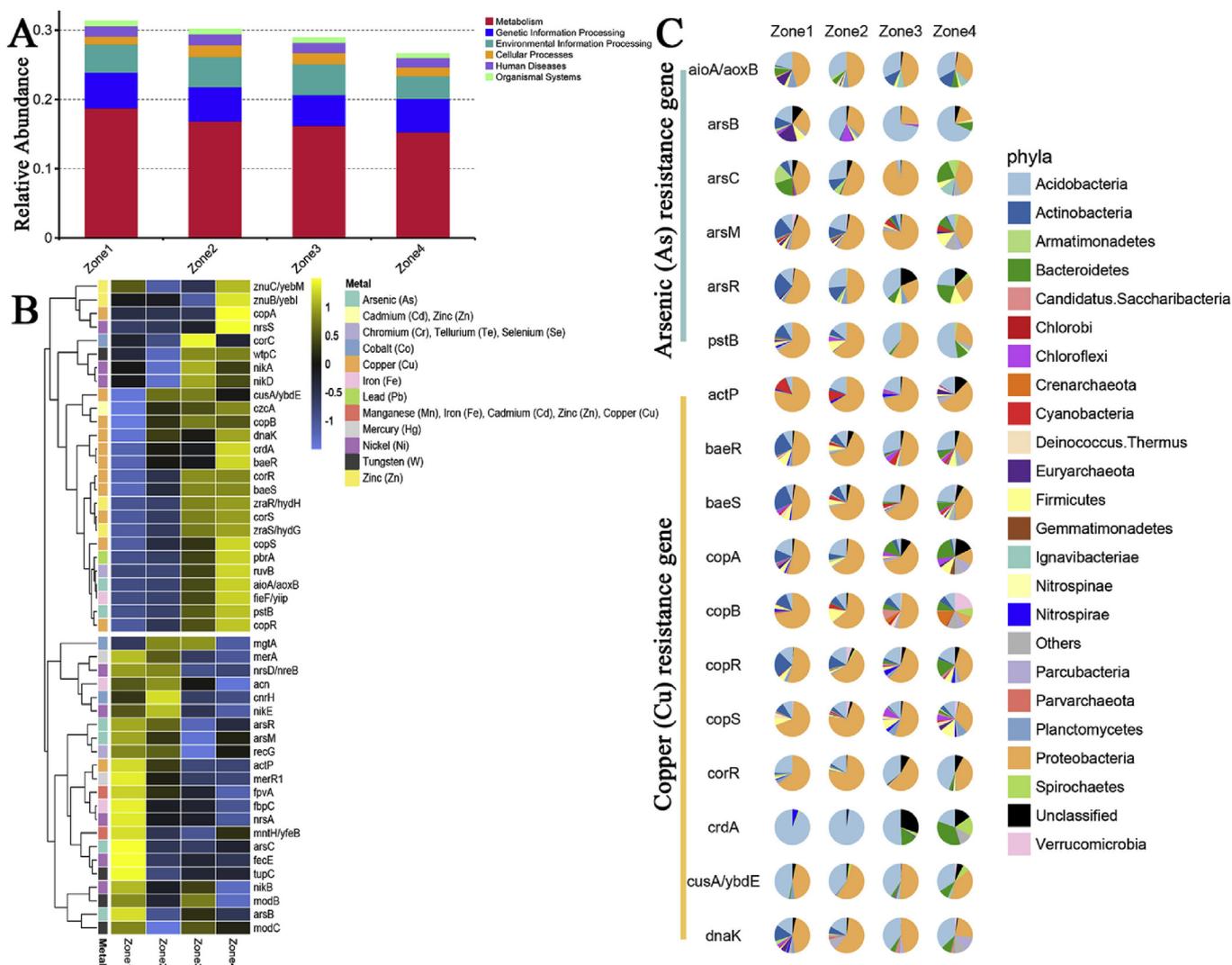


Fig. 5. Functional abundance profile analysis of four zones based on (A) KEGG and (B) BacMet databases. (B) Yellow color indicated high relative abundance level of metal resistance genes (MRGs) and blue indicated low relative abundance level of MRGs among four zones. MRGs responsible for different metal detoxification were colored differently. (C) Distribution of arsenic (As) and copper (Cu) resistance genes among microbial phyla in four zones of Hengshi River. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

4.1. Spatial variations in geochemical data along the river

Most water or sediment parameters demonstrated a significant decreasing trend along the river, which was consistent with previous investigations (Lin et al., 2007; Zhao et al., 2012; Chen et al., 2015c). This trend was due to natural attenuation processes, such as secondary mineral adsorption and co-precipitation, carbonate mineral buffering, and uncontaminated water dilution.

4.2. Changes in composition of bacterial communities along the river

Bacterial communities in the Hengshi River were dominated by certain phyla including Actinobacteria, Firmicutes and Proteobacteria (Fig. S1B), which were similar to previous AMD systems (Kuang et al., 2013a; Liu et al., 2014). A significant shift in bacteria assemblages along the pollution gradient was observed. Although up-stream of the Hengshi River had relatively few lineages, a number of tolerant assemblages with specific functions in acid resistance, metal reduction and sulfate cycling burgeoned under these

extreme conditions, which may have potential in AMD bioremediation. Some lineages including Planococcaceae, Actinomycetales, Halomonas, Bacillales, Acidimicrobiales, Methylophilales, Bacillaceae, Acidithiobacillus, Acidimicrobium, Ferrovum, Leptospirillum and Acetobacteraceae thrived in Zones 1 and 2, but gradually dwindled in Zones 3 and 4 (Figs. S2 and S5). Specifically, Planococcaceae are involved in calcite precipitation and accelerating immobilization, and precipitation of Cu, Cd and Pb (Yang et al., 2016). Comamonas and Flavobacterium are involved in the cycling of nitrogen and sulfur, serving as a nitrogen source for microbial communities (Chang et al., 2002). The genera of reducing ferric Fe (FeOB, Acidocella, Acidiphilium, Meiothermus and Lactobacillus) (Johnson and Bridge, 2002; Johnson and Hallberg, 2003; Baldi et al., 2009; Yang et al., 2014), oxidizing ferrous Fe (FeRB, Acidithiobacillus, Ferrovum, Leptospirillum and Sediminibacterium) (Li et al., 2016), oxidizing sulfur (Acidithiobacillus, Acidiphilium and Thiomonas) (Coupland et al., 2004; Arsène-Plöetz et al., 2010; Auld et al., 2013), and oxidizing arsenite (Halomonas and Thiomonas) (Lin et al., 2012; Auld et al., 2013) were identified.

Sulfate reduction plays a crucial role in AMD bioremediation,

leading to the consumption of protons, increase in pH, and generation of sulfide, which precipitates metals (Johnson and Hallberg, 2005). In contrast to dissimilatory iron reduction, few acidophilic bacteria can reduce sulfate. Optimum growth conditions for sulfate-reducing bacteria (SRB) are pH levels from 5 to 9 (Sun et al., 2015). Lower pH levels of Zones 1 and 2 likely suppressed growth of SRB, and therefore, two genera of *Desulfocapsa* and *Desulfobulbus*, which reduce sulfate, exhibited greater abundances in Zones 3 and 4 (Fig. S5). Relative abundances of *Bacillus* (genus), *Clostridiaceae* (family), *Clostridium* (order), *Clostridiales* (genus), *Gaiellaceae* (family), and *Peptostreptococcaceae* (family) were negatively correlated with the AMD gradient (Figs. S2 and S3B), indicating they are not tolerant to low pH and metal pollution.

The Hengshi River can be characterized by the presence of a wide spectrum of Fe- and S-metabolizing bacteria. Bioreactor systems using these acidophilic and sulfidogenic bacteria that are indigenous to mine-impacted environments could remove and precipitate metals from mine waters. Therefore, phylogenetically divergent lineages coexisting immediately downstream of mining operations may have the potential for *in situ* natural attenuation of Hengshi AMD. Further investigation is needed to explore the significance of these phylotypes in natural attenuation of AMD.

4.3. Relationships between geochemical variables and bacterial community diversity

Compositions of bacterial communities changed significantly with AMD gradients, indicating variations in community structures reflecting the geochemical properties and microbial processes in contaminated sediments. Alpha diversity of Zone 1 sediments was the least, suggesting extreme conditions decreased the diversity, and lower pH levels and greater concentrations of metals and sulfate represented stressors that affect adaptations of bacteria. Both alpha and beta diversities were significantly and positively correlated with pH, confirming pH affected local bacterial communities and their associated ecosystem function (Kuang et al., 2013b). This result might be due to the pH levels directly imposing a strongly physiological constraint on bacteria by reducing the net growth of individual lineages that are unable to survive under highly acidic conditions.

The RF results further demonstrated that Sd.EC, Sd.Sulfate, Sd.Total.Fe, Sd.Ferric and metals in water (Aq.Pb, Aq.Cd, Aq.Zn, and Aq.Cu) were the main contributors influencing diversity, which is consistent with previous AMD studies (Mendez et al., 2008; Wakelin et al., 2012; Yang et al., 2014). Salinity (as EC) was the most important factor influencing alpha diversity, which proved salinity influenced the ecological distribution of prokaryotic taxa along diverse environmental gradients (Kuyucak, 2002; Lozupone and Knight, 2007). Our findings also revealed there are concordant relationships between geochemical predictors and biological information in sediment. Of which, Sd.EC, Aq.Pb, Sd.Sulfate, Aq.Cd, Aq.Zn and Aq.Cu were negatively correlated with bacterial diversity at low concentrations, whereas Sd.Total.Fe, Sd.Ferric and Sd.As were negatively correlated at higher concentrations, which was likely because metals can bind to vital cellular structural proteins, enzymes and nucleic acids, interfering with their normal functioning and leading to toxicity (Olaniran et al., 2013; Edgcomb et al., 2016). Ferric/Ferrous was usually considered as nutritional element and As exists in its residual phase in sediment. Hence, Sd.Total.Fe, Sd.Ferric, and Sd.As had no adverse effects on diversity at low concentrations. However, the threshold values should be validated in future laboratory studies. In AMD environments, pH, metal and sulfur compounds were likely to influence bacterial communities, primarily by favoring distributions of acidophilic and metal- and S-metabolizing microorganisms. This is consistent with the observation that

acidophilic and metal-metabolizing bacteria were more frequently detected in the AMD-contaminated creek. Meanwhile, the SRB genera *Desulfocapsa* and *Desulfobulbus* (Fig. S2), which exhibited increased abundances in Zones 3 and 4, might contribute to decreases in sulfate concentrations downstream.

4.4. Comparative metagenomics among four zones

Identification of the presence and abundances of environment-specific genes through gene-centric metagenomics provided valuable insights into adaptive strategies, metabolic capabilities, and evolutionary processes of microbes along the AMD gradient. Functional abundance analysis of COGs (Fig. S5) revealed metabolism dominated among all groups, which was similar to a previous study (Chen et al., 2015b). KEGG analysis showed that abundances of genes encoding for metabolism in Zone 1 were the greatest, indicating metabolic activity of taxa was necessary to counteract extreme contamination upstream.

4.5. Abundance patterns of MRGs among four zones

Long-term presence of high concentrations of toxic metals appears to have promoted bacterial resistance to metals in Hengshi. Four resistance systems involved in As metabolism have been identified including *aio* encoding arsenite oxidation, *arr* encoding arsenate respiration, and *ars* encoding arsenate reduction and arsenite methylation (Cai et al., 2013). Relative abundances of *arsB*, *arsC*, *arsM* and *arsR* gradually decreased along AMD gradient, which was consistent with sediment As concentrations. However, relative abundances of *aioA/aoxB* showed opposite results. *arsB* encodes for a membrane pump that functions alone or with ATPase *arsA* to transport As(III) out of cells. *arsC* arsenate reductase links arsenate to the efflux pump (Dopson et al., 2001; Valdes et al., 2009). *arsR* bound to As(III) acts as an arsenite-responsive repressor, whereas *aioAB*, *arsC*, *arrAB* and *arsM* are core enzymes for oxidation, reduction, respiration and methylation of As, respectively. Hence, resistance to As in Hengshi River involved a combination of two basic mechanisms, and distinct strategies were adopted among bacterial communities of four zones.

For Cu, five resistance systems, including *cop*-(ATP-dependent Cu transporters), *cue*-(Cu efflux), *cus*-(Cu sensing), *pco*-systems, and *cop*-(homologous to *pco*-), safeguard cell compartments from Cu-induced oxidative damage (Rensing and Grass, 2003). Due to higher concentrations of Cu in downstream sediments, *copA*, *copB*, *copR*, *copS*, *cusA/ybdE*, *corR* and *corS* exhibited greater abundances in Zones 3 and 4 compared to Zones 1 and 2. Among these, *copA* encodes for an uptake P-type ATPase (Costa et al., 2012), forming a multicopper oxidase contributing to tolerance and homeostasis of Cu (Petersen and Moller, 2000; Hall et al., 2008), and *copB* encodes a P-type efflux ATPase for homeostasis (Ng et al., 2012). *cus* encode an RND-type carrier that transports Cu out of cells (Navarro et al., 2009). In addition, *corR* and *corS* regulate multicopper oxidases of *cuoA*, *cuoB*, *cuoC*, and P-type ATPases of *copA* and *copB*. In summary, diverse resistance capacities to As and Cu were found along the pollution gradient. This finding was reasonable since concentrations of As were greater in upstream sediments while Cu concentrations were greater in downstream sediments. Based on MRG functions, the mechanisms of bacterial resistance to metal include (i) converting metal ions to a less toxic form and (ii) active transport/efflux system to cope with As and Cu in Hengshi sediments.

Relative abundances of MRGs were highly diverse among bacterial communities from a variety of pollution levels, which significantly extends our knowledge of interactions and resistance mechanisms between bacteria and metals. The clear trends relating concentrations of As and Cu in sediments to abundances of

respective resistance genes (to the extent of having predictive power) provided first insights into the relationships between functional traits of bacterial community and geochemistry of the surrounding environment. However, it needs to be acknowledged that metagenomes used here only represent snapshots of functional potential of local microbial communities since the presence of a gene does not equate to an ecosystem function. Future quantitative metatranscriptomic and metaproteomic analyses will offer a route to link genetic potential with activity and to provide deeper insight into ecological and evolutionary questions regarding the Hengshi River that are currently only being characterized using phylogenetic markers and gene surveys. Integration of these methods with cultivation-dependent methods will further our understanding of microbial and AMD-impacted ecosystem functioning.

Declaration of interest

The authors of this manuscript report no conflicts of interest. The authors alone are responsible for the content and writing of this article.

Submission declaration

The work described in the manuscript has not been previously published.

Acknowledgments

This project was funded by the Startup Funding of National Institute of Environmental Health to Dr. Tang, the High-level Leading Talent Introduction Program of GDAS to Prof. Sun, and the National Natural Science Foundation of China No. 21707132 to Dr. Tang, No. 81102097 to Prof. Wang and Nos. 21677073 and 21377053 to Prof. Liu. Prof. Giesy and Prof. Hecker were supported by the Canada Research Chair program.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.chemosphere.2018.10.210>.

References

- Arsène-Pløetze, F., Koechler, S., Marchal, M., Coppée, J.-Y., Chandler, M., Bonnefoy, V., Brochier-Armanet, C., Barakat, M., Barbe, V., Battaglia-Brunet, F., 2010. Structure, function, and evolution of the *Thiomonas* spp. genome. *PLoS Genet.* 6, e1000859.
- Auld, R.R., Myre, M., Mykytczuk, N.C., Leduc, L.G., Merritt, T.J., 2013. Characterization of the microbial acid mine drainage microbial community using culturing and direct sequencing techniques. *J. Microbiol. Methods* 93, 108–115.
- Baker, B.J., Banfield, J.F., 2003. Microbial communities in acid mine drainage. *FEMS Microbiol. Ecol.* 44, 139–152.
- Baldi, F., Marchetto, D., Battistel, D., Daniele, S., Faleri, C., De Castro, C., Lanzetta, R., 2009. Iron-binding characterization and polysaccharide production by *Klebsiella oxytoca* strain isolated from mine acid drainage. *J. Appl. Microbiol.* 107, 1241–1250.
- Buchfink, B., Xie, C., Huson, D.H., 2015. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59–60.
- Cai, L., Yu, K., Yang, Y., Chen, B.W., Li, X.D., Zhang, T., 2013. Metagenomic exploration reveals high levels of microbial arsenic metabolism genes in activated sludge and coastal sediments. *Appl. Microbiol. Biotechnol.* 97, 9579–9588.
- Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Pena, A.G., Goodrich, J.K., Gordon, J.I., 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336.
- Chang, Y.H., Han, J.L., Chun, J.S., Lee, K.C., Rhee, M.S., Kim, Y.B., Bae, K.S., 2002. *Comamonas koreensis* sp. nov., a non-motile species from wetland in Woopo, Korea. *Int. J. Syst. Evol. Microbiol.* 52, 377–381.
- Chen, A., Lin, C., Lu, W., Wu, Y., Ma, Y., Li, J., Zhu, L., 2007. Well water contaminated by acidic mine water from the Dabaoshan Mine, South China: chemistry and toxicity. *Chemosphere* 70, 248–255.
- Chen, L.-x., Hu, M., Huang, L.-n., Hua, Z.-s., Kuang, J.-l., Li, S.-j., Shu, W.-s., 2015a. Comparative metagenomic and metatranscriptomic analyses of microbial communities in acid mine drainage. *ISME J.* 9, 1579–1592.
- Chen, L.-x., Hu, M., Huang, L.-n., Hua, Z.-s., Kuang, J.-l., Li, S.-j., Shu, W.-s., 2015b. Comparative metagenomic and metatranscriptomic analyses of microbial communities in acid mine drainage. *ISME J.* 9, 1579–1592.
- Chen, M.Q., Lu, G.N., Guo, C.L., Yang, C.F., Wu, J.X., Huang, W.L., Yee, N., Dang, Z., 2015c. Sulfate migration in a river affected by acid mine drainage from the Dabaoshan mining area, South China. *Chemosphere* 119, 734–743.
- Costa, J.S.D., Kothe, E., Abate, C.M., Amoroso, M.J., 2012. Unraveling the *Amycolatopsis tucumanensis* copper-resistome. *Biometals* 25, 905–917.
- Coupland, K., Battaglia-Brunet, F., Hallberg, K.B., Dictor, M.C., Garrido, F., Johnson, D.B., 2004. Oxidation of Iron, Sulfur and Arsenic in Mine Waters and Mine Wastes: an Important Role for Novel *Thiomonas* Spp. National Technical University of Athens, Zografou, Greece, pp. 639–646.
- Denef, V.J., Mueller, R.S., Banfield, J.F., 2010. AMD biofilms: using model communities to study microbial evolution and ecological complexity in nature. *ISME J.* 4, 599–610.
- DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., Huber, T., Dalevi, D., Hu, P., Andersen, G.L., 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* 72, 5069–5072.
- Dopson, M., Lindström, B.E., Hallberg, K.B., 2001. Chromosomally encoded arsenical resistance of the moderately thermophilic acidophile *Acidithiobacillus caldus*. *Extremophiles* 5, 247–255.
- Edgar, R.C., 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* 10, 996–998.
- Edgcomb, V.P., Pachiadaki, M.G., Mara, P., Kormas, K.A., Leadbetter, E.R., Bernhard, J.M., 2016. Gene expression profiling of microbial activities and interactions in sediments under haloclines of E. Mediterranean deep hypersaline anoxic basins. *ISME J.* 10, 2643–2657.
- Gonzalez-Toril, E., Llobet-Brossa, E., Casamayor, E.O., Amann, R., Amils, R., 2003. Microbial ecology of an extreme acidic environment, the Tinto river (vol 69, pg 4853, 2003). *Appl. Environ. Microbiol.* 69, 6959–6959.
- Hall, S.J., Hitchcock, A., Butler, C.S., Kelly, D.J., 2008. A Multicopper oxidase (Cj1516) and a CopA homologue (Cj1161) are major components of the copper homeostasis system of *Campylobacter jejuni*. *J. Bacteriol.* 190, 8075–8085.
- Ibekwe, A.M., Ma, J., Murinda, S.E., 2016. Bacterial community composition and structure in an Urban River impacted by different pollutant sources. *Sci. Total Environ.* 566–567, 1176–1185.
- Johnson, D., Bridge, T., 2002. Reduction of ferric iron by acidophilic heterotrophic bacteria: evidence for constitutive and inducible enzyme systems in *Acidiphilium* spp. *J. Appl. Microbiol.* 92, 315–321.
- Johnson, D.B., Hallberg, K.B., 2003. The microbiology of acidic mine waters. *Res. Microbiol.* 154, 466–473.
- Johnson, D.B., Hallberg, K.B., 2005. Acid mine drainage remediation options: a review. *Sci. Total Environ.* 338, 3–14.
- Johnson, D.B., Hallberg, K.B., 2008. Carbon, iron and sulfur metabolism in acidophilic micro-organisms. *Adv. Microb. Physiol.* 54, 201–255.
- Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., Glöckner, F.O., 2012. Evaluation of General 16S Ribosomal RNA Gene PCR Primers for Classical and Next-generation Sequencing-based Diversity Studies. *Nucleic Acids Research*, p. gks808.
- Kuang, J.-L., Huang, L.-N., Chen, L.-X., Hua, Z.-S., Li, S.-J., Hu, M., Li, J.-T., Shu, W.-S., 2013a. Contemporary environmental variation determines microbial diversity patterns in acid mine drainage. *ISME J.* 7, 1038–1050.
- Kuang, J.L., Huang, L.N., Chen, L.X., Hua, Z.S., Li, S.J., Hu, M., Li, J.T., Shu, W.S., 2013b. Contemporary environmental variation determines microbial diversity patterns in acid mine drainage. *ISME J.* 7, 1038–1050.
- Kuyucak, N., 2002. Role of microorganisms in mining: generation of acid rock drainage and its mitigation and treatment. *EJMP & EP (Eur. J. Miner. Process. Environ. Protect.)* 2, 179–196.
- Larson, C., 2014. China gets serious about its pollutant-laden soil. *Science* 343, 1415–1416.
- Li, X., Wang, H., Hu, X., Hu, C., Liao, L., 2016. Characteristics of corrosion sales and biofilm in aged pipe distribution systems with switching water source. *Eng. Fail. Anal.* 60, 166–175.
- Lin, C., Wu, Y., Lu, W., Chen, A., Liu, Y., 2007. Water chemistry and ecotoxicity of an acid mine drainage-affected stream in subtropical China during a major flood event. *J. Hazard Mater.* 142, 199–207.
- Lin, Y., Fan, H., Hao, X., Johnstone, L., Hu, Y., Wei, G., Alwathnani, H.A., Wang, G., Rensing, C., 2012. Draft genome sequence of *Halomonas* sp. strain HAL1, a moderately halophilic arsenite-oxidizing bacterium isolated from gold-mine soil. *J. Bacteriol.* 194, 199–200.
- Liu, J., Hua, Z.-S., Chen, L.-X., Kuang, J.-L., Li, S.-J., Shu, W.-S., Huang, L.-N., 2014. Correlating microbial diversity patterns with geochemistry in an extreme and heterogeneous environment of mine tailings. *Appl. Environ. Microbiol.* 80, 3677–3686.
- Lozupone, C.A., Knight, R., 2007. Global patterns in bacterial diversity. *Proc. Natl. Acad. Sci. USA* 104, 11436–11440.
- Madsen, E.L., 2011. Microorganisms and their roles in fundamental biogeochemical cycles. *Curr. Opin. Biotechnol.* 22, 456–464.
- McMurdie, P.J., Holmes, S., 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8, e61217.
- Mendez, M.O., Neilson, J.W., Maier, R.M., 2008. Characterization of a bacterial

- community in an abandoned semiarid lead-zinc mine tailing site. *Appl. Environ. Microbiol.* 74, 3899–3907.
- Navarro, C.A., Orellana, L.H., Mauriaca, C., Jerez, C.A., 2009. Transcriptional and functional studies of *Acidithiobacillus ferrooxidans* genes related to survival in the presence of copper. *Appl. Environ. Microbiol.* 75, 6102–6109.
- Ng, S.P., Palombo, E.A., Bhavne, M., 2012. Identification of a copper-responsive promoter and development of a copper biosensor in the soil bacterium *Achromobacter* sp. AO22. *World J. Microbiol. Biotechnol.* 28, 2221–2228.
- Olaniran, A.O., Balgobind, A., Pillay, B., 2013. Bioavailability of heavy metals in soil: impact on microbial biodegradation of organic compounds and possible improvement strategies. *Int. J. Mol. Sci.* 14, 10197–10228.
- Petersen, C., Moller, L.B., 2000. Control of copper homeostasis in *Escherichia coli* by a P-type ATPase, CopA, and a MerR-like transcriptional activator. *CopR. Gene* 261, 289–298.
- Read, D.S., Gweon, H.S., Bowes, M.J., Newbold, L.K., Field, D., Bailey, M.J., Griffiths, R.I., 2015. Catchment-scale biogeography of riverine bacterioplankton. *ISME J.* 9, 516–526.
- Rensing, C., Grass, G., 2003. *Escherichia coli* mechanisms of copper homeostasis in a changing environment. *FEMS (Fed. Eur. Microbiol. Soc.) Microbiol. Rev.* 27, 197–213.
- Ruiz-González, C., Niño-García, J.P., Giorgio, P.A., 2015. Terrestrial origin of bacterial communities in complex boreal freshwater networks. *Ecol. Lett.* 18, 1198–1206.
- Souffreau, C., Van der Gucht, K., Gremberghe, I., Kosten, S., Lacerot, G., Lobao, L.M., Moraes Huszar, V.L., Roland, F., Jeppesen, E., Vyverman, W., 2015. Environmental rather than spatial factors structure bacterioplankton communities in shallow lakes along a > 6000 km latitudinal gradient in South America. *Environ. Microbiol.* 17, 2336–2351.
- Sun, W., Xiao, E., Dong, Y., Tang, S., Krumins, V., Ning, Z., Sun, M., Zhao, Y., Wu, S., Xiao, T., 2016. Profiling microbial community in a watershed heavily contaminated by an active antimony (Sb) mine in Southwest China. *Sci. Total Environ.* 550, 297–308.
- Sun, W., Xiao, E., Xiao, T., Krumins, V., Wang, Q., Häggblom, M., Dong, Y., Tang, S., Hu, M., Li, B., Xia, B., Liu, W., 2017. Response of soil microbial communities to elevated antimony and arsenic contamination indicates the relationship between the innate microbiota and contaminant fractions. *Environ. Sci. Technol.* 51, 9165–9175.
- Sun, W., Xiao, T., Sun, M., Dong, Y., Ning, Z., Xiao, E., Tang, S., Li, J., 2015. Diversity of the sediment microbial community in the Aha watershed (Southwest China) in response to acid mine drainage pollution gradients. *Appl. Environ. Microbiol.* 81, 4874–4884.
- Valdes, J., Quatrini, R., Hallberg, K., Dopson, M., Valenzuela, P.D.T., Holmes, D.S., 2009. Draft genome sequence of the extremely acidophilic bacterium *Acidithiobacillus caldus* ATCC 51756 reveals metabolic versatility in the genus *Acidithiobacillus*. *J. Bacteriol.* 191, 5877–5878.
- Wakelin, S.A., Anand, R.R., Reith, F., Gregg, A.L., Noble, R.R.P., Goldfarb, K.C., Andersen, G.L., DeSantis, T.Z., Piceno, Y.M., Brodie, E.L., 2012. Bacterial communities associated with a mineral weathering profile at a sulphidic mine tailings dump in arid Western Australia. *FEMS (Fed. Eur. Microbiol. Soc.) Microbiol. Ecol.* 79, 298–311.
- Wang, M., Song, H., Chen, W.-Q., Lu, C., Hu, Q., Ren, Z., Yang, Y., Xu, Y., Zhong, A., Ling, W., 2011. Cancer mortality in a Chinese population surrounding a multi-metal sulphide mine in Guangdong province: an ecologic study. *BMC Publ. Health* 11, 1.
- Wang, Q., Garrity, G.M., Tiedje, J.M., Cole, J.R., 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* 73, 5261–5267.
- Yang, J., Pan, X., Zhao, C., Mou, S., Achal, V., Al-Misned, F.A., Mortuza, M.G., Gadd, G.M., 2016. Bioimmobilization of heavy metals in acidic copper mine tailings soil. *Geomicrobiol. J.* 33, 261–266.
- Yang, Y., Li, Y., Sun, Q.-y., 2014. Archaeal and bacterial communities in acid mine drainage from metal-rich abandoned tailing ponds, Tongling, China. *Trans. Nonferrous Metals Soc. China* 24, 3332–3342.
- Zhao, H.R., Xia, B.C., Qin, J.Q., Zhang, J.Y., 2012. Hydrogeochemical and mineralogical characteristics related to heavy metal attenuation in a stream polluted by acid mine drainage: a case study in Dabaoshan Mine, China. *J. Environ. Sci.-China* 24, 979–989.
- Zhou, X., Xia, B.C., 2010. Defining and modeling the soil geochemical background of heavy metals from the Hengshi River watershed (southern China): integrating EDA, stochastic simulation and magnetic parameters. *J. Hazard Mater.* 180, 542–551.

Supplementary Material

Acid mine drainage affects the diversity and metal resistance gene profile of sediment bacterial community along a river

Xiaohui Zhang^{1, 2, 3}, Song Tang^{4, 5, 6*}, Mao Wang^{7*}, Weimin Sun⁸, Yuwei Xie^{1, 9}, Hui Peng¹⁰,
Aimin Zhong¹¹, Hongling Liu^{1, 2, 3*}, Xiaowei Zhang^{1, 2, 3}, Hongxia Yu^{1, 2, 3}, John P. Giesy^{1, 9},
Markus Hecker^{6, 9}

¹State Key Laboratory of Pollution Control and Resource Reuse, School of the Environment, Nanjing University, Nanjing, Jiangsu 210023, China

²Jiangsu Key Laboratory of Environmental Safety and Health Risk of Chemicals, Nanjing, Jiangsu 210023, China

³Research Center for Environmental Toxicology & Safety of Chemicals, Nanjing University, Nanjing, Jiangsu 210023, China

⁴National Institute of Environmental Health, Chinese Center for Disease Control and Prevention, Beijing 100021, China

⁵Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing, Jiangsu 211166, China

⁶School of Environment and Sustainability, University of Saskatchewan, Saskatoon, SK S7N 5C3, Canada

⁷School of Public Health, Sun Yat-sen University, Guangzhou, Guangdong 510080, China

⁸Guangdong Key Laboratory of Agricultural Environment Pollution Integrated Control,

Guangdong Institute of Eco-Environmental Science & Technology, Guangzhou, Guangdong
510650, China

⁹Toxicology Centre, University of Saskatchewan, Saskatoon, SK S7N 5B3, Canada

¹⁰Department of Chemistry and School of the Environment, University of Toronto, Toronto,
Ontario M5S 3H6, Canada

¹¹The Centre for Disease Control and Prevention of Wengyuan County, Shaoguan,
Guangdong 512600, China

***Correspondence:**

No.7 Panjiayuan Nanli, Chaoyang District, Beijing 100021, China. Tel: +86-10-50930185;

E-mail: tangsong@nieh.chinacdc.cn (Dr. Song Tang)

No.74 Zhongshan 2nd Road, Yuexiu District, Guangzhou, Guangdong 510080, China. Tel:

+86-20-87335546; E-mail: wangmao@mail.sysu.edu.cn (Dr. Mao Wang)

School of the Environment, Nanjing University, Nanjing, Jiangsu 210023, China. Tel:

+86-25-89680356. E-mail: hliu@nju.edu.cn (Dr. Hongling Liu)

Materials and Methods

Geochemical analysis

For grain size analysis, 0.5 g of dry sediment was dispersed by 10 mL of 1 M $i(\text{NaPO}_3)_6$ for 24 hrs and measured by a Malvern Mastersizer 2000 (Malvern Instruments GmbH, Rigipsstr. Herrenberg, Germany). Remaining sediments were lyophilized and thoroughly pulverized in a ball milling machine before sifting through a 200 mesh nylon sieve. Hydrogen peroxidase (HP) activity was measured by a potassium permanganate-based titration method. (Goldblith and Proctor, 1950) To measure electrical conductivity (EC) and pH, 5 g of dry sediment were mixed with 25 mL deionized water. The mixture was then shaken for 5 mins, and then left to equilibrate for 4 hrs before centrifugation for 10 mins at 2000 x g. The EC and pH of the supernatants was then measured by a Thermo Scientific ORION 5-Star (Thermo Fisher Scientific Inc, Waltham, MA, USA). Total organic carbon (TOC) was estimated by loss on ignition of sediments at 550 °C. Concentrations of cadmium (Cd), copper (Cu), lead (Pb), zinc (Zn), arsenic (As), calcium (Ca), cobalt (Co), chromium (Cr), potassium (K), magnesium (Mg), manganese (Mn), molybdenum (Mo), nickel (Ni), strontium (Sr), and vanadium (V), were determined by 0.2 g of sediment that was digested with 12 mL reverse aqua regia (HNO_3 and HCl , 3:1 volume ratio) in a MILESTONE ETHOS UP microwave digestion system (Milestone Office, Shelton, CT, USA) and analyzed by inductively coupled plasma optical emission spectrometry (ICP-OES) (PerkinElmer Office, Akron, OH, United States). Concentrations of sulfate (SO_4^{2-}) in sediments were extracted with a water-soil ratio of 5:1 and then analyzed by Thermo ICS-5000 ion chromatograph (Thermo Fisher Scientific Inc, Waltham, MA, USA). Fe^{3+} and Fe^{2+} were extracted by use of 0.5 M HCl for 1 h and were

measured spectrophotometrically with 1, 10-Phenanthroline and acetic acid-ammonium acetate buffering solution, respectively. Fe^{2+} was quantified by colorimetry at 510 nm by use of a VARIAN Cary 50 Conc UV-visible spectrophotometer (VARIAN medical systems, North Charleston, SC, USA). Fe^{3+} was reduced to Fe^{2+} by hydroxylamine hydrochloride ($\text{NH}_2\text{OH}\cdot\text{HCl}$) and the total concentration of Fe then determined under the same conditions as for Fe^{2+} . Fe^{3+} content in each sediment was calculated by the difference between total Fe and Fe^{2+} .

The pH of water samples was measured *in situ* by use of a multi-parameter sonde for water quality (YSI Incorporated, Yellow Springs, OH, USA). Concentrations of Cd, Cu, Pb, Zn and As in water were measured by ICP-MS (PerkinElmer Office, Akron, OH, United States).

BCR sequential extractions protocol

Since As, Cd, Cu, Pb, and Zn had significantly greater concentrations in sediments from the Hengshi River than in those from the Reference River, four phases of these metals were extracted from sediments: exchangeable, reducible, oxidizable, and residual. The exchangeable phase contained metal ions adsorbed on surfaces of sediment particles and carbonates in the sediment. A portion of these metals are associated with sediment carbonate and susceptible to release into water when pH was lowered. The reducible phase represented material cemented or coatings on particles. These phases are unstable under anoxic conditions. The oxidizable phase consists of metals bound to various forms of organic matter. The residual fraction consisted of metals bound within crystal structures of minerals.

Exchangeable and reducible phases are considered chemically mobile since they can release metals into the surrounding environment as geochemical conditions change. (Mossop and Davidson, 2003; Rao et al., 2008) In brief, the exchangeable fraction (F1) was extracted with 0.11 M acetic acid. The reducible fraction (F2) was removed by use of 0.5 M hydroxylamine hydrochloride, with the pH adjusted to 1.5 ± 0.1 by use of 2 M HNO₃. The oxidizable fraction (F3) was extracted by use of 8.8 M H₂O₂ (pH 2~3) followed by 1 M ammonium acetate of which the pH was adjusted to 2 ± 0.1 by use of 2 M HNO₃. Extraction of the residual phase (F4) with reverse aqua regia (HNO₃ and HCl, 3:1 volume ratio) in a Milestone Ethos Up microwave digestion system. The certified reference material BCR-701 (Pueyo et al., 2001) was analyzed in parallel. Concentrations of various metals were determined by ICP-MS (PerkinElmer Office, Akron, OH, United States) with indium used as an internal standard.

16S sequencing and data processing

PCR amplification (20 µL) was conducted use of Platinum[®] Taq polymerase (Life Technologies, Carlsbad, CA, USA) in a SureCycler 8800 Thermal Cycler (Agilent Technologies, Santa Clara, CA, USA) with the following conditions: 28 cycles of denaturation at 94°C for 15 s, annealing at 54°C for 30 s and extension at 68°C for 30 s, with a final extension at 68°C for 7 min. Triplicate PCR products were pooled for each subsample and the size and specificity were monitored on 2% agarose gels. Mixed PCR products were purified by use of the MinElute Gel Extraction Kit (Qiagen, Hilden, Germany). Purified DNA was quantified by using Qubit[™] dsDNA HS Assay Kits (Thermo Fisher Scientific Inc, Waltham, MA, USA). All purified PCR products were pooled in an equimolar

fashion and linked to sequence adapters by use of Ion Xpress™ Plus Region Library Kit (Life Technologies, Carlsbad, CA, USA). A Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA) was used to assess sizes and concentrations of amplicons. Finally, samples were adjusted to a concentration of 26 pM before being attached to surfaces of Ion Sphere particles (ISPs) with Ion OneTouch™ 200 Template Kit v2 DL (Life Technologies, Santa Clara, CA, USA) and run on an Ion Torrent Personal Genome Instrument (Life Technologies, Santa Clara, CA, USA).

All sequence reads were automatically filtered using the Ion server v.3.6.2 with default parameters. Low quality reads and sequence adaptors were trimmed and then the sequence reads were analyzed with the Quantitative Insights Into Microbial Ecology (QIIME) toolkit (Caporaso et al., 2010) and UPARSE pipeline (Edgar, 2013). “Noisy” sequences and chimeric sequences were discarded and sequence reads with: (i) a sequence shorter than 150 bp or longer than 190 bp; and/or (ii) with one or more ambiguous “N”; and/or (iii) with eight or more homopolymers were also removed. Operational taxonomic units (OTUs) clustering at 97% sequence similarity were obtained by pooling all sequence reads and using an UPARSE algorithm. Here, any cluster containing less than four reads was removed. Subsequently, a representative sequence was chosen from each OTU and classified by operating a Ribosomal Database Project (RDP) Classifier (Wang et al., 2007) with a confidence greater than 80% against the Greengene database (DeSantis et al., 2006). Archaeal, plastid, mitochondrial and chloroplast OTUs were discarded. Alpha diversity (Chao 1) was estimated using the package of “phyloseq” (McMurdie and Holmes, 2013) in R software v.3.3.1. For beta diversity (un-weighted UniFrac metric), to avoid biases generated by

differences in sequencing depth, all samples were normalized to the minimum number of total reads (13,740) among all 16S rRNA sequencing libraries.

Metagenomic data filter condition

Raw sequence reads that contained a quality score less than 38 for more than 40 bases, more than 10 bp ambiguous “N”, and an overlap with the adapter in excess of 15 bp were discarded to procure “clean” data for subsequent data analysis. Clean data were assembled de novo and analyzed by SOAPdenovo software.(Luo et al., 2012) Three K-mer values (49, 55 and 59) were chosen to assemble each sample into scaffolds and the scaffold with the largest N50 was selected as the final result. Subsequently, the final scaffolds were cut at “N” and then scaffigs, sequences without “N”, were obtained. Clean data from all samples were aligned with scaftigs to acquire unassembled paired-end reads using SoapAligner and 55 was selected as the K-mer value to assemble these mixed unassembled sequences. To obtain mixed scaftigs, scaffolds assembled from mixed sequences were cut at “N”.

Metagenomic shotgun sequencing generated approximately 144 million reads for use in bioinformatics analysis. After quality control using the NGS QC Toolkit, 143,810,551 (21 Gb, 99.79 %) high quality reads were obtained from four sequencing libraries. Subsequently, all high-quality reads were assembled, and a self-writing script was used to filter the assembled sequences shorter than 350 bp, which resulted in a total of 318,370,989 bases, with N50 values of 1198 bp, 1220 bp, 966 bp and 1257 bp in Zones 1, 2, 3 and 4, respectively. For gene prediction, 406,849 Open Reading Frame (ORF) were identified using MetaGeneMark.

Statistical analysis

Random Forests (RF) analysis was employed to determine the importance of geochemical parameters and quantify the strengths of the interactions. Random Forests are an ensemble learning method which could be used for classification or regression by constructing mass decision trees.(Ho, 1995) Variance of decision trees generated by growing deep could be well decreased by bootstrap aggregating (bagging) procedure, and the model would perform well even with noisy variables.(James et al., 2013) In this study, RF model was used to predict the importance of each geochemical parameters on microbial diversity (i.e. Chao1 index) by using the “ggRandomForest” package in R. In the model, geochemical parameters were severed as predictor variables and Chao1 index as dependent variable. We used 1000 decision trees to build the RF. RF partial dependence plots were also obtained to give a graphical depiction of the marginal effect of a predictor on the response variable after partialling out the effects of the other predictors in the model. The importance of each geochemical parameter was determined by the difference of the accuracy after replaced the given column of training set data by out-of-bag data (geochemical parameter).

“Phyloseq” and “edgeR”(Robinson et al., 2010) packages were used for detection of significantly differential abundant OTUs (logarithmic fold change with the base of 2, log₂FC) in samples between the Hengshi and Reference rivers. Mantel and partial Mantel tests (9999 permutations) were used to investigate correlations between beta diversities (un-weighted UniFrac metric) and all measured geochemical variables (Euclidean distance) using the “Vegan” package. Spearman's Ranked Correlations were used to visualize correlations between relative abundances of each genus in communities of bacteria with each other by use

of the “Corrplot” and “Hmisc” packages. Network visualization and module detection of co-occurrence relationships of bacterial genera in the Hengshi River were conducted by use of the interactive platform of Gephi v 0.9.1 with Fruchterman Reingold placement algorithm.(Bastian et al., 2009) Topology of the resulting network was described by a set of measures, including average node connectivity, average path length, diameter, cumulative degree distribution, clustering coefficient and modularity. In the co-occurrence network, nodes represent measured variables and genera and edges correspond to robust and significant correlations (Spearman's correlation coefficients $|\rho|>0.7$ and $p<0.01$) between nodes. For the MRGs, relative abundances with a same gene name were summed together regardless of bacterial species. For the COGs, the statistically significant features of a given COG between two datasets were identified by pairwise comparisons by use of two-tailed Fisher's exact test, with confidence intervals at 99% significant and Benjamini-Hochberg correction ($p < 0.05$). (Parks and Beiko, 2010; Chen et al., 2015)

Results

Mantel and Partial Mantel tests were utilized to determine the geochemical variables affecting the community functional structure (beta diversity). The results showed that the gradients of water quality parameters including pH, Cu, Zn, Cd and Pd and variables in sediments including pH, As, Pb and sulfate were still significantly correlated with beta diversity after eliminating the effect of spatial distance (Table S2).

Results of Fisher's exact test showed that 1337 of 3311 COGs had significantly different expressions among the four zones (two-tailed Fisher's exact test) (Figure S5A), indicating variability among genome profiles. A COG with significantly greater or lesser expressions in a specific zone compared to all other zones was defined as an indicator COG, which attributed the most or the fewest transcripts for the associated function across the four zones. As a result, 366, 259, 156 and 305 COGs had significantly greater expressions in Zones 1, 2, 3 and 4, respectively, when compared to all other zones. Correspondingly, 207, 212, 199 and 466 COGs exhibited lesser expressions in Zones 1, 2, 3 and 4 (Figure S5B), respectively, and all those COGs were identified as indicator genes of their respective zones. All COGs were sorted into three groups based on functions: information storage and processing, cellular processes and signaling and metabolism. Indicator COGs from the metabolism group dominated for all groups.

Table S1. Geochemical variables (mean \pm S.D.) measured in four zones of the Hengshi River and the Reference River. The results of Geochemical variables for each sampling sites are given in Table S3.

Variables	Zone 1	Zone 2	Zone 3	Zone 4	Reference
Aquatic pH	2.68 (\pm 0.037)	3.19 (\pm 0.2)	4.07 (\pm 0.46)	5.32 (\pm 0.45)	6.62 (\pm 0.36)
Aquatic Cu (mg/L)	5.63 (\pm 1.02)	1.38 (\pm 0.52)	0.57 (\pm 0.17)	0.1 (\pm 0.053)	0.0035 (\pm 0.0013)
Aquatic Zn (mg/L)	71.92 (\pm 12.24)	11.54 (\pm 7.84)	6.98 (\pm 2.24)	2.13 (\pm 0.88)	0.013 (\pm 0.0042)
Aquatic As (μg/L)	9.23	N.A.	N.A.	N.A.	N.A.
Aquatic Cd (mg/L)	0.3 (\pm 0.054)	0.051 (\pm 0.03)	0.027 (\pm 0.0075)	0.0097 (\pm 0.0036)	0.00009 (\pm 0.00002)
Aquatic Pb (mg/L)	0.54 (\pm 0.043)	0.16 (\pm 0.059)	0.15 (\pm 0.19)	0.016 (\pm 0.012)	0.0078 (\pm 0.0034)
Sediment pH	3.43 (\pm 0.34)	3.36 (\pm 0.53)	5.02 (\pm 0.58)	6.22 (\pm 1.09)	6.62 (\pm 0.52)
Sediment Conductivity (ms/cm)	0.66 (\pm 0.31)	0.88 (\pm 0.51)	0.29 (\pm 0.1)	0.28 (\pm 0.11)	0.071 (\pm 0.057)
Sediment TOC(kg/kg)	0.02 (\pm 0.0082)	0.043 (\pm 0.024)	0.04 (\pm 0.015)	0.028 (\pm 0.015)	0.016 (\pm 0.008)
Sediment Hydrogen Peroxidase ml (0.1mol/L KMnO₄)/(h·g)	2.42 (\pm 0.64)	2.53 (\pm 0.34)	3.09 (\pm 2.50)	1.9 (\pm 0.26)	2.1 (\pm 0.16)
Sediment Cu (g/kg)	0.45 (\pm 0.031)	0.54 (\pm 0.18)	0.95 (\pm 0.43)	1.01 (\pm 0.52)	0.03 (\pm 0.008)
Sediment Zn (g/kg)	0.36 (\pm 0.042)	0.32 (\pm 0.11)	0.32 (\pm 0.10)	0.39 (\pm 0.13)	0.096 (\pm 0.013)
Sediment As(g/kg)	1.03 (\pm 0.22)	0.43 (\pm 0.23)	0.47 (\pm 0.18)	0.27 (\pm 0.15)	0.035 (\pm 0.016)
Sediment Cd(mg/kg)	2.8 (\pm 0.49)	3.18 (\pm 1.38)	2.5 (\pm 0.8)	3.85 (\pm 2.16)	0.98 (\pm 0.13)
Sediment Pb (g/kg)	2.03 (\pm 0.58)	0.86 (\pm 0.44)	0.88 (\pm 0.46)	0.65 (\pm 0.48)	0.027 (\pm 0.0035)
Sediment Cu-Exchangeable (g/kg)	0.049 (\pm 0.014)	0.066 (\pm 0.037)	0.19 (\pm 0.11)	0.2 (\pm 0.064)	0.0074 (\pm 0.0049)
Sediment Cu-Reducible (g/kg)	0.035 (\pm 0.0018)	0.011 (\pm 0.0057)	0.14 (\pm 0.11)	0.079 (\pm 0.071)	0.00072 (\pm 0.00024)
Sediment Cu-Oxidisable (g/kg)	0.075 (\pm 0.11)	0.11 (\pm 0.092)	0.24 (\pm 0.12)	0.36 (\pm 0.35)	0.0065 (\pm 0.0031)
Sediment Cu-Residual (g/kg)	0.32 (\pm 0.051)	0.38 (\pm 0.1)	0.33 (\pm 0.065)	0.25 (\pm 0.1)	0.024 (\pm 0.0019)
Sediment Zn-Exchangeable (g/kg)	0.023 (\pm 0.0065)	0.038 (\pm 0.035)	0.059 (\pm 0.042)	0.093 (\pm 0.52)	0.0055 (\pm 0.0042)
Sediment Zn-Reducible (g/kg)	0.015 (\pm 0.00059)	0.017 (\pm 0.0044)	0.028 (\pm 0.013)	0.027 (\pm 0.0063)	0.0093 (\pm 0.0013)
Sediment Zn-Oxidisable (g/kg)	0.018 (\pm 0.0097)	0.061 (\pm 0.052)	0.042 (\pm 0.024)	0.051 (\pm 0.03)	0.0012 (\pm 0.00056)
Sediment Zn-Residual (g/kg)	0.29 (\pm 0.087)	0.19 (\pm 0.057)	0.18 (\pm 0.031)	0.17 (\pm 0.066)	0.039 (\pm 0.015)
Sediment As-Exchangeable (g/kg)	0.0003 (\pm 0.00030)	0.00025 (\pm 0.00027)	0.00031 (\pm 0.00024)	0.00059 (\pm 0.00052)	0.00072 (\pm 0.00038)
Sediment As-Reducible (g/kg)	0.0017 (\pm 0.00024)	0.0012 (\pm 0.00024)	0.00096 (\pm 0.00014)	0.002 (\pm 0.001)	0.002 (\pm 0.00063)

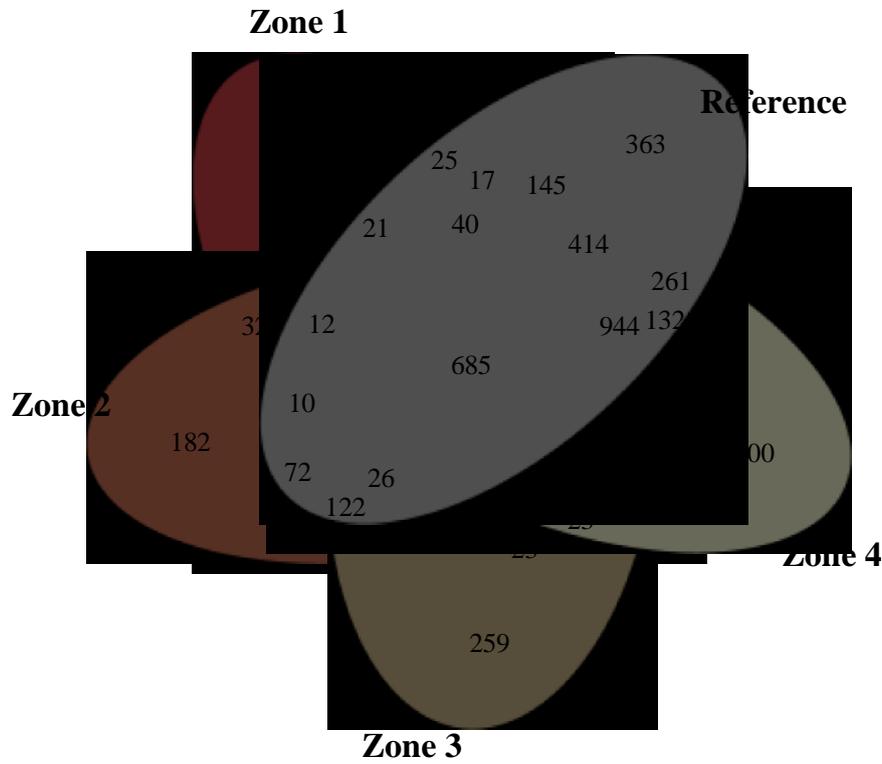
Sediment As-Oxidisable (g/kg)	0.0054 (±0.00062)	0.0067 (±0.0012)	0.0059 (±0.00072)	0.0066 (±0.00086)	0.0059 (±0.0003)
Sediment As-Residual (g/kg)	0.95 (±0.21)	0.389 (±0.22)	0.42 (±0.147)	0.26 (±0.15)	0.023 (±0.012)
Sediment Cd-Exchangeable (mg/kg)	0.86 (±0.45)	0.89 (±0.61)	0.63 (±0.31)	1.43 (±0.89)	0.13 (±0.032)
Sediment Cd-Reducible (mg/kg)	0.47 (±0.037)	0.44 (±0.072)	0.58 (±0.12)	0.71 (±0.31)	0.36 (±0.004)
Sediment Cd-Oxidisable (mg/kg)	0.36 (±0.15)	0.84 (±0.71)	0.40 (±0.21)	0.8 (±0.76)	0.013 (±0.01)
Sediment Cd-Residual (mg/kg)	0.94 (±0.13)	0.96 (±0.46)	1.27 (±0.076)	0.98 (±0.21)	0.41 (±0.12)
Sediment Pb-Exchangeable (g/kg)	0.087 (±0.053)	0.038 (±0.035)	0.049 (±0.021)	0.072 (±0.062)	0.0011 (±0.00072)
Sediment Pb-Reducible (g/kg)	0.33 (±0.19)	0.086 (±0.072)	0.17 (±0.079)	0.12 (±0.081)	0.0044 (±0.00091)
Sediment Pb-Oxidisable (g/kg)	0.045 (±0.0014)	0.021 (±0.013)	0.065 (±0.015)	0.046 (±0.03)	0.0029 (±0.0012)
Sediment Pb-Residual (g/kg)	1.18 (±0.21)	0.54 (±0.32)	0.47 (±0.27)	0.31 (±0.27)	0.018 (±0.0056)
Sediment Particle Size	13.57 (±4.37)	26.53 (±11.12)	21.2 (±5.61)	33.11 (±26.29)	28.023 (±3.86)
Sediment SO₄²⁻ (g/kg)	3.47 (±1.9)	3.65 (±2.2)	1.52 (±1.005)	0.86 (±0.29)	0.46 (±0.23)
Sediment Fe²⁺ (g/kg)	0.71 (±0.13)	1.11 (±0.45)	0.57 (±0.19)	1.69 (±1.45)	0.27 (±0.089)
Sediment Fe³⁺ (g/kg)	15.75 (±0.98)	14.93 (±1.92)	13.19 (±3.93)	7.17 (±1.96)	0.78 (±0.29)
Sediment T-Fe (g/kg)	16.45 (±0.9)	16.04 (±1.93)	13.76 (±4.08)	8.86 (±1.42)	1.05 (±0.36)

Table S2. Mantel and Partial (geographic distance-corrected) Mantel tests (9999 permutations) reveal the correlations between geochemical variables (Euclidean distance) and beta diversities of bacterial communities (unweighted UniFrac distance) in Hengshi River. Data in bold indicates significant correlations at $p < 0.05^*$, $p < 0.01^{**}$, and $p < 0.001^{***}$.

Items	Mantel Test		Partial Mantel Test	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
Distance	0.236	0.009**	-	-
Aq.pH	0.33	0.003**	0.253	0.012*
Aq.Cu	0.441	0.001***	0.395	0.003**
Aq.Zn	0.444	0.002**	0.415	0.002**
Aq.Cd	0.446	0.001***	0.415	0.002**
Aq.Pb	0.332	0.007**	0.293	0.012*
Sd.pH	0.295	0.005**	0.189	0.044*
Sd.EC	0.201	0.052	0.187	0.06
Sd.TOC	-0.042	0.61	-0.044	0.62
Sd.HP	0.02	0.422	0.029	0.409
Sd.Cu	0.114	0.155	0.038	0.341
Sd.Zn	0.054	0.277	0.044	0.301
Sd.As	0.309	0.005**	0.28	0.012*
Sd.Cd	0.214	0.064	0.191	0.08
Sd.Pb	0.232	0.023*	0.212	0.037*
Sd.Sulfate	0.245	0.026*	0.218	0.045*
Sd.Ferrous	0.214	0.065	0.158	0.146
Sd.Ferric	0.076	0.182	-0.039	0.626
Sd.Total.Fe	0.005	0.414	-0.104	0.914
Sd.PS	-0.025	0.523	-0.039	0.571
Sd.Cu.Exchangeable	0.033	0.322	-0.046	0.654
Sd.Cu.Reducible	-0.137	0.883	-0.172	0.946
Sd.Cu.Oxidisable	0.192	0.089	0.138	0.16
Sd.Cu.Residual	-0.089	0.76	-0.109	0.809
Sd.Zn.Exchangeable	0.18	0.063	0.121	0.155
Sd.Zn.Reducible	-0.085	0.819	-0.148	0.967
Sd.Zn.Oxidisable	0.026	0.383	0.028	0.374
Sd.Zn.Residual	0.049	0.34	0.038	0.365
Sd.As.Exchangeable	0.162	0.09	0.136	0.128
Sd.As.Reducible	0.038	0.365	0.003	0.456
Sd.As.Oxidisable	-0.026	0.602	-0.013	0.53
Sd.As.Residual	0.324	0.004**	0.295	0.008**
Sd.Cd.Exchangeable	0.201	0.054	0.17	0.084
Sd.Cd.Reducible	0.127	0.186	0.051	0.336
Sd.Cd.Oxidisable	0.149	0.111	0.134	0.143
Sd.Cd.Residual	-0.04	0.608	-0.044	0.63
Sd.Pb.Exchangeable	0.306	0.008**	0.303	0.007**
Sd.Pb.Reducible	0.133	0.156	0.117	0.188
Sd.Pb.Oxidisable	-0.02	0.565	-0.076	0.829
Sd.Pb.Residual	0.308	0.004**	0.29	0.006**

Figure S1. (A) Venn diagrams show the numbers of unique and shared operational taxonomic units (OTUs) among four different zones in Hengshi River. **(B)** Relative abundances of bacterial communities at phylum level observed at different sampling sites of the Hengshi River and the Reference River. Only phylum with relative abundance > 0.5% is shown.

(A)



(B)

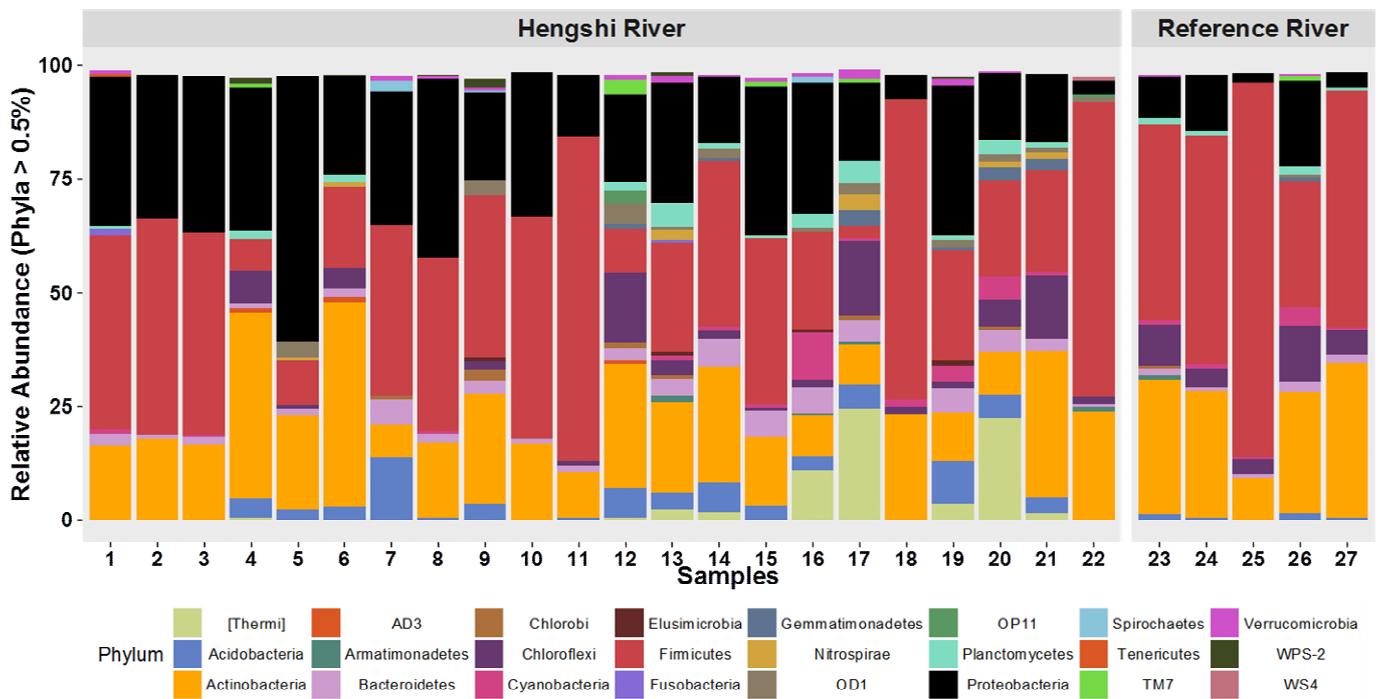


Figure S4. Heatmap of metagenomic sequencing taxa annotations at genus level. Relative abundances of each bacterial genus (top 35 genus) among four different zones of Hengshi River are normalized by Z-score.

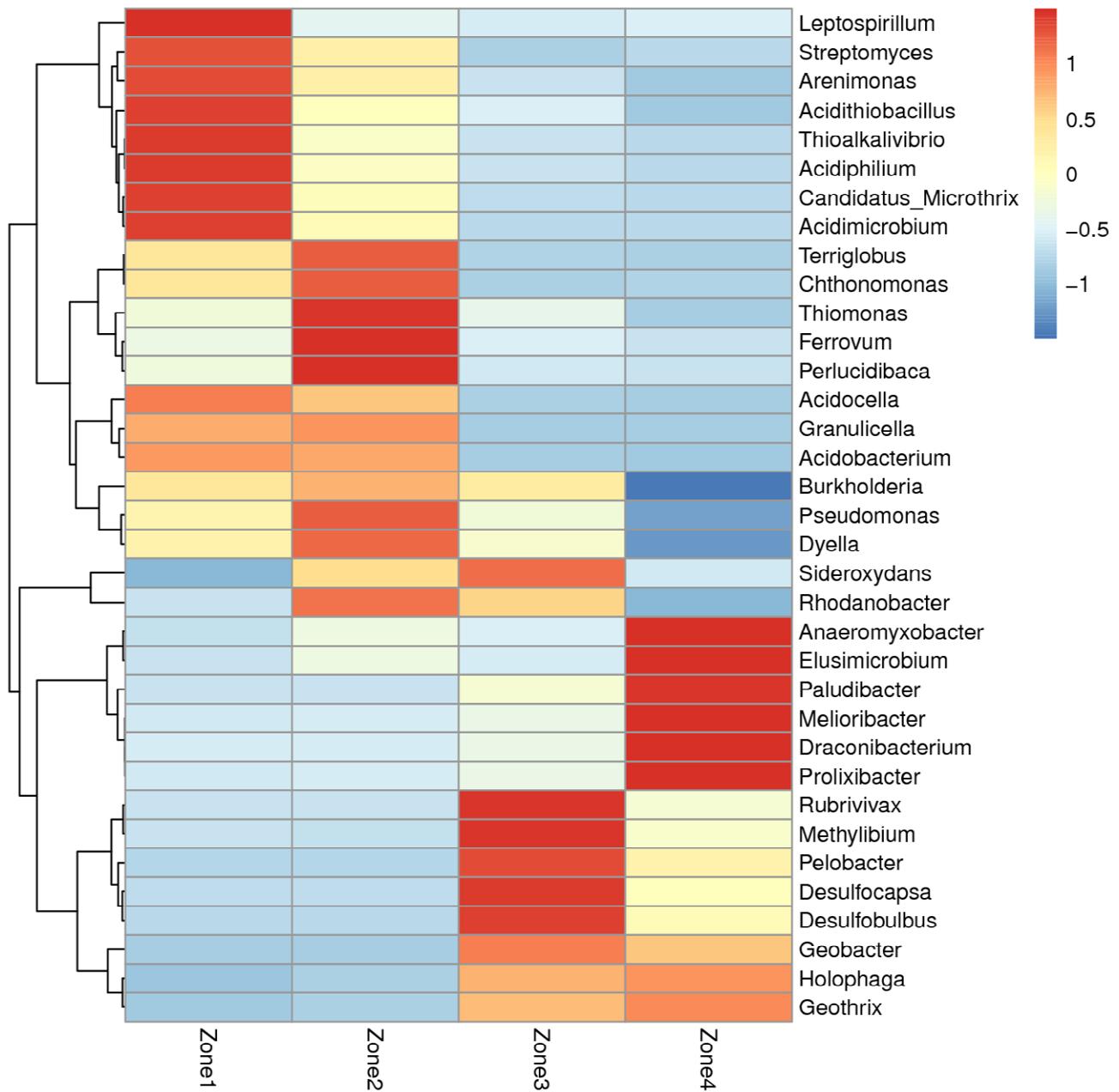
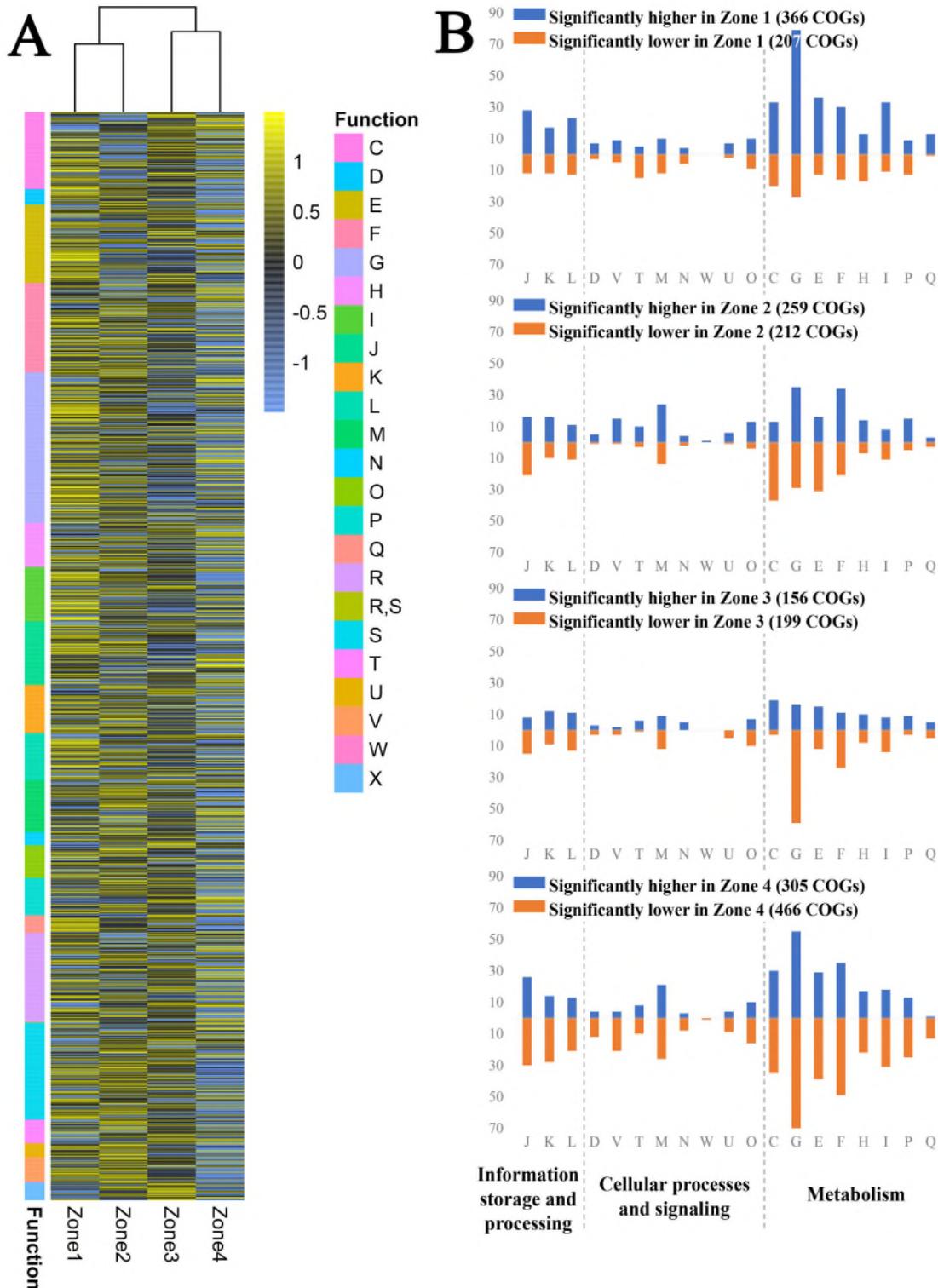


Figure S5. (A) Hierarchical cluster analysis of COGs with significant expression levels based on the relative abundance in each zone of Hengshi River. COG categories: C-Energy production and conversion; D-Cell cycle control, cell division, chromosome partitioning; E-Amino acid transport and metabolism; F-Nucleotide transport and metabolism; G-Carbohydrate transport and metabolism; H-Coenzyme transport and metabolism; I-Lipid transport and metabolism; J-Translation, ribosomal structure and biogenesis; K-Transcription; L-Replication, recombination and repair; M-Cell wall/membrane/envelope biogenesis; N-Cell motility; O-Posttranslational modification, protein turnover, chaperones; P-Inorganic ion transport and metabolism; Q-econdary metabolites biosynthesis, transport and catabolism; T-Signal transduction mechanisms; U-Intracellular trafficking, secretion, and vesicular transport; V-Defense mechanisms; W-Extracellular structures. **(B)** The number of COGs with significantly different (higher or lower) expression levels than those in other three zones.



References

- Bastian, M., Heymann, S., Jacomy, M., 2009. Gephi: An Open Source Software for Exploring and Manipulating Networks.
- Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Pena, A.G., Goodrich, J.K., Gordon, J.I., 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature methods* 7, 335-336.
- Chen, L.-x., Hu, M., Huang, L.-n., Hua, Z.-s., Kuang, J.-l., Li, S.-j., Shu, W.-s., 2015. Comparative metagenomic and metatranscriptomic analyses of microbial communities in acid mine drainage. *ISME Journal* 9, 1579-1592.
- DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., Huber, T., Dalevi, D., Hu, P., Andersen, G.L., 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and environmental microbiology* 72, 5069-5072.
- Edgar, R.C., 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature methods* 10, 996-998.
- Goldblith, S.A., Proctor, B.E., 1950. PHOTOMETRIC DETERMINATION OF CATALASE ACTIVITY. *Journal of Biological Chemistry* 187, 705-709.
- Ho, T.K., 1995. Random decision forests. *Document Analysis and Recognition, 1995., Proceedings of the Third International Conference on. IEEE*, pp. 278-282.
- James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. *An introduction to statistical learning.* Springer.
- Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., Liu, Y., 2012. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *GigaScience* 1, 1.
- McMurdie, P.J., Holmes, S., 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PloS one* 8, e61217.
- Mossop, K.F., Davidson, C.M., 2003. Comparison of original and modified BCR sequential extraction procedures for the fractionation of copper, iron, lead, manganese and zinc in soils and sediments. *Analytica Chimica Acta* 478, 111-118.
- Parks, D.H., Beiko, R.G., 2010. Identifying biologically relevant differences between metagenomic communities. *Bioinformatics* 26, 715-721.
- Pueyo, M., Rauret, G., Luck, D., Yli-Halla, M., Muntau, H., Quevauville, P., Lopez-Sanchez, J.F., 2001. Certification of the extractable contents of Cd, Cr, Cu, Ni, Pb and Zn in a freshwater sediment following a collaboratively tested and optimised three-step sequential extraction procedure. *Journal of Environmental Monitoring* 3, 243-250.
- Rao, C., Sahuquillo, A., Sanchez, J.L., 2008. A review of the different methods applied in environmental geochemistry for single and sequential extraction of trace elements in soils and related materials. *Water, Air, and Soil Pollution* 189, 291-333.
- Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139-140.
- Wang, Q., Garrity, G.M., Tiedje, J.M., Cole, J.R., 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and environmental microbiology* 73, 5261-5267.

Sample ID	Sampling Time	Location	
		E	N
1	14-Oct-2014	113°43'47.19"	24°30'47.21"
2	14-Oct-2014	113°44'51.77"	24°30'38.92"
3	14-Oct-2014	113°45'04.28"	24°30'10.41"
4	14-Oct-2014	113°45'02.81"	24°30'09.78"
5	14-Oct-2014	113°45'04.75"	24°30'09.35"
6	14-Oct-2014	113°45'05.90"	24°30'08.86"
7	14-Oct-2014	113°45'25.05"	24°29'49.16"
8	14-Oct-2014	113°45'40.58"	24°29'34.40"
9	14-Oct-2014	113°46'24.46"	24°29'21.75"
10	14-Oct-2014	113°46'54.74"	24°29'12.19"
11	14-Oct-2014	113°47'44.79"	24°28'43.36"
12	14-Oct-2014	113°48'39.64"	24°28'37.88"
13	14-Oct-2014	113°48'51.97"	24°27'34.35"
14	14-Oct-2014	113°48'52.59"	24°27'17.54"
15	14-Oct-2014	113°48'45.17"	24°26'50.26"
16	14-Oct-2014	113°48'20.59"	24°26'18.86"
17	14-Oct-2014	113°48'37.12"	24°25'36.73"
18	14-Oct-2014	113°48'42.68"	24°25'05.78"
19	14-Oct-2014	113°48'45.15"	24°24'01.91"
20	14-Oct-2014	113°48'50.25"	24°22'60.00"
21	14-Oct-2014	113°48'29.10"	24°22'34.42"
22	14-Oct-2014	113°48'31.42"	24°21'56.71"
23	14-Oct-2014	114°06'47.16"	24°11'08.52"
24	14-Oct-2014	114°06'20.52"	24°11'04.92"
25	14-Oct-2014	114°06'12.60"	24°10'59.88"
26	14-Oct-2014	114°06'33.98"	24°10'59.16"
27	14-Oct-2014	114°06'33.73"	24°10'34.68"
Sample ID	Sampling Time	E	N
1	11-Oct-2015	113°43'47.19"	24°30'47.21"
7	11-Oct-2015	113°45'25.05"	24°29'49.16"
17	11-Oct-2015	113°48'37.12"	24°25'36.73"
22	11-Oct-2015	113°48'31.42"	24°21'56.71"

N.D.: Not Detected N.A.: Not Available C1: Exchangeable Fraction C2: Reducible Fraction

Aquatic Parameters						
pH	Cu (µg/L)	Zn (µg/L)	As (µg/L)	Cd (µg/L)	Pb (µg/L)	pH
2.64	6595.00	82730.00	9.23	352.40	569.10	2.98
2.68	6081.00	78230.00	N.D.	325.40	569.70	3.50
3.45	1344.00	2369.00	N.D.	20.20	94.27	3.80
2.73	4225.00	54810.00	N.D.	227.10	477.30	3.80
3.19	1110.00	9703.00	N.D.	42.53	158.00	2.70
3.32	1317.00	3863.00	N.D.	23.57	76.05	4.10
3.11	1955.00	17260.00	N.D.	73.72	254.10	3.80
2.93	1600.00	16020.00	N.D.	67.78	148.40	2.90
3.18	848.10	8765.00	N.D.	36.77	143.40	3.30
3.46	602.70	6641.00	N.D.	28.57	149.60	3.70
2.91	2287.00	27680.00	N.D.	112.00	246.90	2.60
3.41	776.40	9278.00	N.D.	38.06	141.60	5.50
3.47	633.10	7558.00	N.D.	30.25	67.67	3.80
4.51	398.40	5130.00	N.D.	20.77	34.24	5.50
4.25	314.30	3584.00	N.D.	15.05	47.47	5.30
4.27	546.00	6304.00	N.D.	26.27	48.67	4.90
4.51	764.90	10010.00	N.D.	31.29	565.10	5.10
5.11	197.70	3551.00	N.D.	15.08	7.72	5.60
4.64	41.00	963.20	N.D.	4.69	8.03	4.80
5.19	80.51	1477.00	N.D.	6.96	8.44	5.70
5.85	74.24	2368.00	N.D.	10.86	16.75	7.40
5.79	111.80	2288.00	N.D.	10.91	39.22	7.60
6.47	4.99	15.19	N.D.	0.12	12.40	6.40
6.51	4.98	19.72	N.D.	0.13	9.09	6.10
6.20	2.99	9.63	N.D.	0.06	4.78	6.40
6.63	1.73	7.79	N.D.	0.08	3.22	6.60
7.28	3.03	14.51	N.D.	0.08	9.67	7.60
pH	Cu (µg/L)	Zn (µg/L)	As (µg/L)	Cd (µg/L)	Pb (µg/L)	pH
N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	2.89
N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	3.92
N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	5.25
N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	7.10

C3: Oxidizable Fraction C4: Residual Fraction

Sediment Parameters				Se	
Conductivity (us/cm)	TOC(kg/kg)	Catalase	Sulfate (mg/kg)	Cu	Zn
1067.00	0.02	3.26	2523.17	485.20	376.83
332.00	0.01	1.70	6131.08	410.51	306.59
219.50	0.02	2.30	1768.49	602.44	250.71
588.00	0.03	2.30	1415.74	441.77	406.00
1334.00	0.02	N.D.	8420.65	249.51	239.02
317.00	0.03	2.15	1750.52	743.64	443.56
1030.00	0.03	3.29	5073.77	822.53	494.52
1532.00	0.09	2.64	4441.08	363.47	171.29
782.00	0.07	2.45	3418.76	574.93	378.11
358.00	0.03	2.44	1520.71	564.06	342.21
1471.00	0.05	2.45	3194.63	377.03	228.40
130.70	0.02	2.12	745.22	454.13	198.66
376.00	0.04	1.65	1802.53	321.84	179.66
266.80	0.04	2.05	963.68	1492.22	457.15
465.00	0.05	8.65	1739.17	1000.27	428.21
261.10	0.07	1.90	3470.78	1371.94	325.47
258.00	0.04	2.15	418.01	1053.55	329.29
440.00	0.04	1.75	1226.59	966.27	435.46
346.00	0.01	N.D.	1012.86	948.76	279.25
217.50	0.04	2.20	650.65	963.81	433.92
130.80	0.01	2.10	418.01	256.68	202.99
283.90	0.04	1.55	1001.51	1900.71	584.92
32.70	0.01	2.10	501.23	26.19	88.58
29.20	0.01	1.95	763.19	25.62	117.76
39.40	0.01	1.90	632.68	45.83	84.06
70.10	0.02	2.25	176.85	25.76	105.06
181.10	0.03	2.30	205.22	25.52	85.03
Conductivity (us/cm)	TOC(mg/kg)	Catalase	Sulfate (mg/kg)	Cu	Zn
990.00	N.A.	N.A.	2537.355778	441.7997235	350.6965242
895.00	N.A.	N.A.	720.6355211	542.4793615	438.7154838
301.00	N.A.	N.A.	1538.679781	980.6284081	317.9427735
244.00	N.A.	N.A.	1219.027804	1737.387134	281.074438

Dimensional Metal Total Concentration (mg/kg)

As	Cd	Pb	Fe(II)	Fe(III)	Cu-C1	Cu-C2
1287.33	3.25	2708.54	800.39	14371.45	35.62	36.43
749.15	2.12	1284.17	797.16	16283.59	67.59	35.22
133.02	2.51	320.53	517.44	16580.10	39.10	9.67
1062.90	3.03	2103.67	985.14	9991.60	44.16	32.13
245.51	2.12	538.17	1102.71	15619.51	49.39	4.07
476.83	3.79	684.81	1201.55	15970.93	93.35	21.21
361.68	6.20	877.03	2202.20	14482.56	141.96	7.20
202.62	1.33	475.06	733.85	15920.54	36.90	8.56
470.09	2.99	986.86	826.23	16028.42	74.53	17.14
882.40	3.87	1744.15	1082.69	15843.02	75.96	12.02
647.26	2.66	1274.17	706.72	15595.61	14.22	3.82
674.04	3.02	1359.47	625.97	9038.76	67.13	18.47
323.32	1.24	90.08	674.42	16289.41	41.15	15.10
621.45	3.58	1314.84	674.42	14972.87	350.55	226.00
623.47	3.08	1218.38	679.59	16109.17	256.61	123.52
257.86	1.87	541.17	153.75	6496.77	246.28	317.67
297.35	2.21	733.36	603.36	16241.60	206.54	132.79
382.17	3.49	1502.80	1037.47	9932.82	306.37	30.35
473.00	1.70	449.29	467.70	6412.47	182.13	157.12
132.16	3.54	353.33	598.84	7592.38	233.55	172.62
62.92	2.59	123.28	1943.15	7939.28	116.24	34.67
303.63	7.95	816.91	4403.10	3985.79	168.72	1.04
22.43	1.04	25.63	220.23	592.03	4.28	0.64
23.58	0.85	29.13	162.47	690.45	5.87	0.46
22.80	0.78	21.64	224.63	460.56	17.15	1.17
61.20	1.09	27.37	357.88	1302.97	5.82	0.61
45.86	1.12	32.01	394.70	844.55	4.10	0.71
As	Cd	Pb	Fe(II)	Fe(III)	Cu-C1	Cu-C2
1324.137275	3.130159988	2288.614275	854.005168	16089.79328	34.91419226	19.73349013
327.094388	6.016999332	907.9690119	2001.937984	15662.1447	68.79136271	43.78415548
244.8151872	2.452011132	396.7006537	1497.416021	15251.29199	178.3799021	236.6880904
260.2556065	7.656254551	461.4071985	1664.728682	11718.34625	211.3705351	394.9280073

Cu-C3	Cu-C4	Zn-C1	Zn-C2	Zn-C3	Zn-C4	As-C1
89.04	332.68	32.25	15.14	20.81	273.04	0.74
63.30	253.63	19.18	14.29	5.40	188.76	0.07
170.96	426.16	0.99	15.32	67.13	148.18	0.13
73.75	377.31	17.81	15.73	28.81	399.65	0.15
44.48	176.20	66.88	12.60	8.15	124.19	0.15
251.30	371.68	26.13	14.80	162.82	209.78	0.89
249.62	419.33	118.26	26.02	125.64	210.49	0.25
24.37	319.37	25.26	20.76	23.94	105.87	0.04
38.89	565.62	27.81	15.13	30.12	290.89	0.10
97.56	372.12	30.79	18.24	52.40	233.74	0.45
17.68	362.18	6.76	11.69	17.29	196.15	0.03
116.83	252.02	3.00	11.22	20.06	177.71	0.73
55.20	246.56	5.35	11.51	5.79	129.19	0.02
410.69	381.57	104.87	41.34	77.84	224.96	0.28
263.44	321.78	104.25	37.81	43.36	208.97	0.48
343.26	362.91	76.31	40.50	63.06	162.73	0.21
233.83	422.50	60.31	28.44	40.53	182.78	0.10
356.76	248.11	76.39	23.12	54.38	237.24	1.04
246.27	285.29	45.36	25.39	29.15	197.20	0.12
97.21	248.59	102.79	38.15	42.40	184.88	0.24
62.34	66.31	51.86	19.61	22.87	42.29	0.16
1037.80	390.39	188.37	28.67	107.10	178.23	1.37
4.38	24.53	1.12	6.82	1.99	60.38	0.35
3.76	25.24	2.42	9.29	0.36	42.61	0.22
12.28	26.86	13.04	10.23	0.80	45.22	0.87
5.21	21.21	6.03	9.36	1.50	31.08	0.89
7.10	23.26	4.79	10.69	1.30	15.63	1.26
Cu-C3	Cu-C4	Zn-C1	Zn-C2	Zn-C3	Zn-C4	As-C1
41.12387071	263.2250672	48.5758533	24.14880042	24.76287559	199.9980319	0.756904312
30.71110736	356.5497769	55.11440235	32.99192395	75.49038806	228.9245019	0.403048436
94.19210045	406.8394244	71.69605329	25.55363164	30.18979671	178.9845503	0.416468834
653.102911	406.8175957	110.4807127	44.07567588	29.40035566	90.43289445	1.256412872

Metal Fractions (mg/kg)						
As-C2	As-C3	As-C4	Cd-C1	Cd-C2	Cd-C3	Cd-C4
2.05	6.27	1194.39	1.49	0.51	0.34	1.06
1.51	4.77	683.14	0.47	0.43	0.19	0.76
1.79	4.85	118.97	0.14	0.39	0.92	0.78
1.55	5.26	965.18	0.62	0.45	0.55	1.00
1.19	4.84	220.96	1.13	0.40	0.19	0.59
1.22	7.84	420.41	0.64	0.47	1.79	0.77
1.14	7.35	324.33	2.12	0.60	2.12	0.77
1.03	7.41	177.62	0.41	0.37	0.15	0.43
1.01	7.40	430.42	1.34	0.44	0.37	1.95
1.06	7.97	801.18	1.05	0.46	0.93	0.99
0.99	5.59	620.00	0.27	0.37	0.27	1.38
1.22	6.59	608.44	0.65	0.50	0.64	1.30
0.93	4.81	346.70	0.18	0.37	0.15	1.31
1.05	5.59	515.21	0.89	0.74	0.66	1.27
0.86	5.19	571.11	1.09	0.62	0.28	1.31
0.80	6.22	236.82	0.36	0.58	0.17	1.10
0.89	6.79	264.73	0.58	0.63	0.47	1.32
3.47	7.18	388.47	1.57	0.43	0.71	0.94
2.78	6.87	443.36	0.23	0.53	0.25	1.06
0.81	7.71	118.37	1.00	0.85	0.67	1.29
0.98	5.73	52.99	1.39	0.49	0.12	0.66
1.80	5.48	280.65	2.94	1.26	2.25	0.95
1.47	5.62	14.41	0.10	0.36	0.00	0.47
1.30	5.62	15.50	0.09	0.35	0.00	0.54
1.66	5.61	11.79	0.12	0.36	0.03	0.49
2.76	6.23	43.42	0.14	0.35	0.02	0.30
2.70	6.22	30.55	0.18	0.36	0.01	0.23
As-C2	As-C3	As-C4	Cd-C1	Cd-C2	Cd-C3	Cd-C4
3.406426157	15.05196942	993.0599139	1.635752817	0.368932158	0.469047845	1.17428065
1.499432562	4.645342651	231.8173428	2.848734251	0.906813297	1.792852801	1.004051135
1.425086994	4.407082666	174.62432	0.959205929	0.664858117	0.592040371	1.222377099
2.632318645	5.33283689	189.5314649	3.176782594	0.970658199	2.432629776	1.076535592

Pb-C1	Pb-C2	Pb-C3	Pb-C4	Particle Size
				D [4, 3]
108.02	581.18	45.67	1353.71	7.38
13.69	113.55	43.44	881.32	16.60
11.34	29.60	5.05	252.43	28.79
138.62	308.13	46.74	1316.92	16.72
11.06	17.48	12.40	328.01	N.A.
55.09	188.92	20.65	228.34	16.12
45.34	65.99	18.53	503.95	44.06
13.83	19.58	19.36	276.05	40.43
40.59	87.46	50.99	624.72	23.55
119.64	217.33	12.29	1165.64	12.09
8.37	58.73	30.10	909.10	20.70
72.99	192.06	48.69	748.58	19.78
17.04	56.82	43.33	31.75	18.23
71.63	254.64	69.92	754.00	15.53
61.55	269.60	71.87	622.12	33.19
35.02	127.77	86.61	259.30	19.84
33.21	98.53	71.71	393.89	20.63
151.04	239.55	73.40	813.33	14.14
32.23	187.88	51.82	162.90	84.75
24.98	73.76	16.39	152.50	21.06
9.38	20.61	6.57	51.39	28.73
142.71	75.78	82.33	390.87	16.86
1.80	5.51	2.49	24.11	30.80
2.15	4.54	1.83	24.58	31.72
0.83	3.96	1.88	17.65	30.90
0.26	2.83	2.92	10.42	22.57
0.62	4.91	5.23	13.41	24.14
Pb-C1	Pb-C2	Pb-C3	Pb-C4	Particle Size
246.9532399	614.8647189	13.64618892	1577.161289	N.A.
56.47633636	110.0154711	7.105665458	667.5245266	N.A.
25.00377847	87.85580334	5.62478133	227.685881	N.A.
93.59214126	97.51945021	8.076158885	194.513084	N.A.