Environmental Science Processes & Impacts

Microbial Community Structure with Trends in Methylation Gene Diversity and Abundance in Mercury-Contaminated Rice Paddy Soils in Guizhou, China

Journal:	*Environmental Science: Processes & Impacts*
Manuscript ID	EM-ART-11-2017-000558.R2
Article Type:	Paper
Date Submitted by the Author:	21-Feb-2018

Complete List of Authors:	Vishnivetskaya, Tatiana; University of Tennessee Hu, Haiyan; State Key Laboratory of Environmental Geochemistry, Institute of Geochemistry, Chinese Academy of Sciences Van Nostrand, Joy; University of Oklahoma Wymore, Ann; Oak Ridge National Laboratory, Biosciences Division Xu, Xiaohang; State Key Laboratory of Environmental Geochemistry, Institute of Geochemistry, Chinese Academy of Sciences Qiu, Guangle; Institute of Geochemistry, Chinese Academy of Sciences, Feng, Xinbin; Institute of Geochemistry, Zhou, Jizhong; University of Oklahoma Brown, Steven; Oak Ridge National Laboratory, Brandt, Craig; Oak Ridge National Laboratory, Biosciences Division Podar, Mircea; Oak Ridge National Laboratory, Gu, Baohua; Oak Ridge National Laboratory, Environmental Sciences Division Elias, Dwayne; Oak Ridge National Laboratory, Biosciences Division

Environmental Significance

The neurotoxic effects of methylmercury (MeHg) are well-known and the genes responsible for its biological production were recently discovered. Given this recent development few reports exist that directly utilize this genomic information. The current study is unique in that it discerns high versus low Hg contamination, the subsequent differential impacts on the bacterial community, and suggests which members of the community may be responsible for MeHg production in such a globally important food crop as rice paddies.

Microbial Community Structure with Trends in Methylation Gene Diversity and Abundance in Mercury-Contaminated Rice Paddy Soils in Guizhou, China

Tatiana A. Vishnivetskaya,[1#§] Haiyan Hu,[2,3§] Joy D. Van Nostrand,[4] Ann M. Wymore,[1] Xiaohang Xu,[3] Guangle Qiu,[3] Xinbin Feng,[3] Jizhong Zhou,[4] Steven D. Brown,[1] Craig C. Brandt,[1] Mircea Podar,[1] Baohua Gu,[2*] and Dwayne A. Elias[1*]

[1]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6036

[2]Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6038

[3]State Key Laboratory of Environmental Geochemistry, Institute of Geochemistry, Chinese Academy of Sciences, Guiyang 550002, China

[4]Department of Microbiology and Plant Biology, University of Oklahoma, Norman, Oklahoma 73019

[§] These authors contributed equally.

*Corresponding authors: B.G. (gub1@ornl.gov) and D.E. (eliasda@ornl.gov)

[#]Current address: Center for Environmental Biotechnology, University of Tennessee, Knoxville, Tennessee 37996, USA

Key words: Mercury, methylmercury, Hg-methylation, *hgcAB*, rice paddy soil, pyrosequencing

Abstract

Paddy soils from mercury (Hg)-contaminated rice fields in Guizhou, China were studied with respect to total mercury (THg) and methylmercury (MeHg) concentrations as well as Bacterial and Archaeal community composition. Total Hg (0.25-990 μg g^{-1}) and MeHg (1.3-30.5 ng g^{-1}) varied between samples. Pyrosequencing (454 FLX) of the hypervariable v1-v3 regions of the 16S rRNA genes showed that *Proteobacteria*, *Actinobacteria*, *Chloroflexi*, *Acidobacteria*, *Euryarchaeota*, and *Crenarchaeota* were dominant in all samples. The *Bacterial* α-diversity was higher in samples with relatively low THg and MeHg and decreased with increasing THg and MeHg concentrations. In contrast, Archaeal α-diversity increased with increasing of MeHg concentrations but did not correlate with changes in THg concentrations. Overall, the methylation gene *hgcAB* copy number increased with both increasing THg and MeHg concentrations. The microbial communities at high THg and high MeHg appear to be adapted by species that are both Hg resistant and carry *hgcAB* genes for MeHg production. The relatively high abundance of both sulfate-reducing δ-*Proteobacteria* and methanogenic Archaea, as well as their positive correlations with increasing THg and MeHg concentrations, suggests that these microorganisms are the primary Hg-methylators in the rice paddy soils in Guizhou, China.

Introduction

Mercury (Hg) is a global pollutant that is widely present in both surface and subsurface terrestrial environments as well as in the atmosphere, due in part to anthropogenic and natural Hg emissions from volcanism [1]. Man-made activities including coal combustion, Hg mining and industrial processes contribute to global Hg pollution [2]. More than 25% of the world's Hg emissions originate in China [3, 4], mostly from coal combustion, Hg mining activities and the use of Hg in industrial processes. The province of Guizhou, in particular, is the largest Hg producer in China due to historical smelting activities and thus draws national and international attention [5]. Hg contamination in Guizhou is widespread in the atmosphere, soil and water. Southern territories including Guizhou, Yunnan, and Sichuan provinces, the Yangtze valley, and the Zhu Jiang delta are the predominant rice producing agricultural regions in China. Since these areas are an economically important part of the Chinese economy, accounting for 26% of all world rice production [6, 7], the elevated Hg concentrations are of concern to human health since rice plants are known for taking up Hg or methylmercury (MeHg) and storing them in rice grains [8].

In anoxic rice paddy soils, inorganic Hg can be converted to MeHg by anaerobic bacteria including sulfate- and iron-reducing bacteria as well as methanogenic Archaea, which are abundant in these environments [9, 10]. Rice plants growing in Hg-contaminated soil can accumulate MeHg to high levels, causing imminent health risks to people whose primary food is rice [11]. Indeed, for local inhabitants in Guizhou province, rice intake is the main pathway of MeHg exposure as opposed to fish and fish product consumption [12, 13]. Therefore, Hg pollution in rice paddy fields is considered a major health concern in China and many parts of the world where rice is the primary source of nutrition.

The present study was undertaken to determine microbial community responses to Hg pollution in rice-paddy soils from Guizhou province in China. We focused on relationships between microbial communities and the Hg-methylation genes to Hg pollution with the goal of identifying the most abundant Hg-methylator diversities in these soils.

Materials and Methods

Study Area and sample collection

The sampling campaign was conducted in July 2011 during the rice growing season from historical Hg mining sites in Tongren District, Guizhou province, China (Figure 1). Historical Hg smelting activities produced large numbers of calcines, mine tailings and wastes, which were piled along stream and river banks near the mine processing sites and retorts. Although large-scale mining activities were shut down in 2001, artisanal Hg smelting activities were revived locally in Gouxi (3G) and Jinjiachang (4J) sites (Table 1) due to rising Hg demand. All paddy fields had been growing rice for at least 3 consecutive years and were flooded when soil samples were collected. The rice paddy fields in these sites also were irrigated with Hg-contaminated river and stream waters. Detailed description of the sampling sites, soil THg and MeHg concentrations, and geochemical properties is given in Table 1 and Table S1. Soils are mostly clay loam, and their mineralogical composition (Table S2) consists of mostly quartz, feldspar, illite, and montmorillonite. These sampling sites include one active artificial Hg mining site (5M), one upstream (3G), and seven downstream sites of the mining area (8Y, 9Y, 14Y, 16Y, 7S, 4J, 6D) (Table 1). The control site, Huaxi (1H), is located southeast of Guiyang city with no known historical Hg mining activities. At each sampling site, three to five subsamples were collected at ~5 cm depth from different locations within a grid area of about 4 × 4 m. These subsamples were composited and then divided into two subsamples in polyethylene bags for geochemical analyses and three additional subsamples in 2-mL Eppendorf tubes for DNA extraction. All samples were then immediately frozen in liquid N_2 and transported on ice within 24 h to the laboratory for analysis at the Institute of Geochemistry in Guiyang, China. One set of samples in the polyethylene bag was freeze-dried, homogenized, passed a 150-mesh sieve, and used for chemical and mineralogical analyses, whereas the other set of soil samples was stored at –20°C until porewater geochemical and biological analyses. Soil samples for genomic DNA (gDNA) extractions were also stored at –20°C until analysis.

Soil analyses

Soil pH was determined *in-situ* using a portable pH meter (Beckman Coulter, Inc., Fullerton, CA). Soil organic matter (SOM) was determined using the freeze-dried soil samples by the percent loss on ignition (LOI) for two hours at 550°C. Total mercury (THg) was analyzed via thermal decomposition, amalgamation and atomic absorption spectrophotometry using a Lumex

RA915+ Hg analyzer equipped with a PYRO 915+ pyrolysis attachment (Lumex, Model RA-915+/PYRO-915+, St. Petersburg, Russia). The detection limit was 0.2 ng Hg/g dry soil. The United States EPA method 1630 was used for total MeHg analysis. In brief, 0.3–0.4 g of soil sample was weighed and extracted using $CuSO_4$-methanol/solvent and methylene chloride. The extract was reacted with a sodium tetraethylborate ($NaBEt_4$) solution to form methylethyl Hg ($CH_3CH_3CH_2Hg$), which was separated from solution by purging with N_2 onto a Tenax trap. The trapped methylethyl Hg was thermally desorbed, separated from other Hg species by an isothermal gas chromatography (GC) column, decomposed to Hg(0) in a pyrolytic decomposition column (800^{o}C), and finally analyzed by cold-vapor atomic fluorescence spectrometry (CVAFS) (Brooks Rand model III, Brooks Rand Laboratories, Seattle, WA). The method detection limits for THg and MeHg were 0.02 and 0.002 ng Hg g^{-1} dry soil, respectively. Two certified reference soil materials, Montana soil (SRM-2710, National Institute of Standards and Technology) and Loamy Sand 1 (CRM024-050, Resource Technology Corporation), were also analyzed along with samples for THg. The recovery of THg in reference samples ranged from 91 to 110%. Similarly, the certified reference sediment (BCR-580, Institute for Reference Materials and Measurements) was used for MeHg analysis, and the recovery was 84 to 112%.

Soil samples (0.4 g) were also extracted with 0.05 M ammonium acetate (20 mL) for major exchangeable anion (e.g., chloride, nitrate, and sulfate) and metal ion analyses. The anion samples were filter sterilized via a 0.22 μm syringe filter into sterile 5 mL tubes for analysis. All samples were placed onto the Dionex ICS 5000+ Dual Pump, Dual Column system (ThermoFisher Scientific; Waltham, MA), as previously described [14]. Anions were analyzed using an AS11HC column with a KOH gradient of 0-60 mM, per the manufacturer's instructions. Calibration of each parameter was accomplished using 5-point calibration curves with the prepackaged standards from Dionex. The calibration curves were performed at the beginning of each run and included check standards after every 15 samples. Additionally, the extract was used for quantification of 13 metal ions via inductively coupled plasma – mass spectrometry (ICP-MS). Samples were prepared for analysis as previously described [15, 16].

Soil gDNA extraction

The frozen soil samples in Eppendorf tubes (triplicate) were thawed on ice, and the total community gDNA isolated from 0.5 g of each soil sample using the PowerSoil™ DNA Isolation

Kit (MoBio Laboratories, Carlsbad, CA, USA). The extracted gDNA from each soil was combined and frozen at –20°C and then freeze-dried in the laboratory of the Institute of Geochemistry, China, before they were shipped to the Oak Ridge National Laboratory (ORNL), Oak Ridge, Tennessee, for further purification and analyses. In brief, gDNA samples were dissolved in high-purity DNase-free water, and their concentrations and purities were assayed by measuring UV absorbance at 260 nm (A260) and the ratios of A260/A280 and A260/A230 using a spectrophotometer (NanoDrop2000^{TM}; Thermo Scientific Inc., West Palm Beach, FL). The ratios of A260/A280 and A260/A230 of all DNA samples were 1.8-2.0, suggesting acceptable DNA purity for pyrosequencing and Geochip analyses.

Pyrosequencing of the Bacterial and Archaeal 16S rRNA genes

The hypervariable v1-v3 region of the 16S rRNA gene was amplified using universal Bacterial (forward 5′-TYACCGCGGCTGCTGG-3′ and reverse 5′-AGAGTTTGATYMTGGCTCAG-3′) or Archaeal (forward 5′-CTCYSGTTGATCCYGCSRG-3′ and reverse 5′-GCTACRGVYSCTTTARRC-3′) primers. The 6 bp (Bacteria) and 7 bp (Archaea) unique tag sequences were fused with forward primers. PCR mixtures (50 µl) consisted of forward and reverse primers (1.5 µl each from 10 µM), 1 µl template DNA (10 to 80 ng μl^{-1}), and 0.5 µl (5 U μl^{-1}) high-fidelity Platinum® Taq DNA Polymerase (Invitrogen, Carlsbad, CA). Samples were denatured (94°C, 2 min) and amplified for 29 cycles at the following thermal regime for Bacteria (94°C for 15 sec, 55°C for 30 sec, and 68°C for 45 sec) and for Archaea (94°C for 15 sec, 53°C for 30 sec, and 68°C for 45 sec), followed by a final extension step (72°C, 3 min). The PCR reactions were visualized for correct size using 1.5% agarose gel. The PCR amplicons were purified using Agencourt AMPure solid-phase paramagnetic bead technology (Agencourt Bioscience Corporation, Beverly, MA). The purity, concentration, and size of the PCR amplicons were estimated using DNA 1000 chips and an Agilent model 2100 Bioanalyzer (Agilent Technologies, Inc., Waldbronn, Germany). Sequencing reactions were performed on a Life Sciences GS 454 FLX genome sequencer (Roche Diagnostics, Indianapolis, IN).

GeoChip analyses

Functional gene analysis was conducted on the GeoChip version 4 [17]. The GeoChip was manufactured by Roche NimbleGen (Madison, WI, USA) using their 12-plex format (12 arrays per slide, 135,000 probes per array). The array contained 82,074 functional gene probes for

141,995 CDS from 410 gene families. A total of 733 probes were specific to mercury cycling pathways, including genes for mercuric reductase (*merA*), organomercurial lyase (*merB*), mercuric resistance (*merR*), and mercuric transport (*merP*).

A total of ~1 µg DNA was labeled with Cy3 by random priming using the Klenow fragment of DNA polymerase. The gDNA was combined with 20 µl of 3 µg μl^{-1} random hexamer primers (Invitrogen Life Technologies, CA) and heated at 99.9°C for 5 min, cooled on ice, and then centrifuged. Then 20 µl of reaction buffer (2.5 µl dNTPs [5 mM (each) dATP, dGTP, and dCTP and 2.5 mM dTTP], 2 µl (40 U) of Klenow (Invitrogen Life Technologies, CA), and 1 µl of 100 nmol Cy3-dUTP dye (Amersham BioSciences, UK) were added and incubated for 3 h at 37°C. The labeling reaction was terminated by heating at 98°C for 3 min. The tubes were removed, placed on ice and the labeled gDNA was purified using a QIAquick PCR purification kit (Qiagen) and then dried using a SpeedVac (Thermo Scientific™).

Samples were hybridized as described previously [17]. Briefly, dried samples were rehydrated with NimbleGen sample tracking control to confirm sample identity and hybridization buffer (40% formamide, 25% SSC, 1% SDS, 2.38% Cy3-labelled alignment oligo (NimbleGen) and 2.8% Cy5-labelled universal standard target) [18]. An HX12 mixer (NimbleGen) was attached to the slide and samples were loaded onto the arrays and hybridized overnight (~16 h) with mixing on a hybridization station (MAUI, BioMicro Systems, Salt Lake City, UT, USA).

Microarray scanning, data normalization and analysis

The GeoChip microarrays were scanned with a MS200 Microarray Scanner (NimbleGen) at 100% of laser power and photomultiplier tube gain. Scanned images were gridded using NimbleScan software (Nimblegen) and raw data was uploaded onto the Microarray Data Manager (MGM) pipeline on the Institute for Environmental Genomics (IEG) website (http://ieg.ou.edu/microarray). The signal intensity of each spot was normalized by the mean of all spots and by the average universal standard signal [17, 19]. A floating signal-to-noise ratio (SNR) was used to determine positive probes, with SNR determined based on the value that would result in <5% of the Thermophile control probes were positive. Spots were also removed if the total signal was <1000.

Qualitative hgcAB PCR and quantitative hgcA qPCR

The recently described universal PCR primers for *hgcAB* were used to determine the presence and diversity of this gene pair. Amplicons were cut from the gels, cloned and sequenced as previously described [20]. Quantitative PCR (qPCR) was utilized for determining the abundance of the Hg-methylating gene *hgcA* [20, 21]. Briefly, the recently published degenerate primers that are specific for *hgcA* were used with the published protocol for the Hg-methylating clades of the δ-*Proteobacteria* and methanogenic Archaea [20]. Amplicon purity was checked via agarose gel for a single band.

Phylogenetic analyses

Raw 454 sequencing reads (~235 Mb) were denoised, filtered for quality and chimeric sequences, and demultiplexed using Qiime [22]. Each sample yielded from ~4,000 to ~20,000 Bacterial sequences and from ~6,000 to 17,000 Archaeal sequences. From sample 5M, only 12 Archaeal sequences were obtained, and therefore these sequences were not used in analyses. The sequences processed through the Qiime were uploaded to MG-RAST [23] under accession numbers 4550440.3, 4550442.3, 4550446.3, 4550448.3, 4550449.3, 4550451.3, 4550453.3, 4550455.3, 4550457.3, 4550459.3 (for Bacteria), and 4550439.3, 4550441.3, 4550445.3, 4550447.3, 4550450.3, 4550452.3, 4550454.3, 4550456.3, 4550458.3 (for Archaea). The 16S rRNA gene sequences were annotated in MG-RAST using the best-hit classification, RDP as annotation source, $1e^{-5}$ maximum e-value cutoff, 60% minimum identity cutoff, and 15 minimum alignment length cutoff [24]. The OTUs were called at 97% identity cutoff using Qiime. Sequences that could not be annotated to the genus level were grouped at higher taxonomical levels. The sequencing data were normalized using DESeq approach [25], which allows valid comparison across species and improves accuracy in the detection of differential abundance [26]. The α-diversity is shown as species richness, which is calculated as the antilog of the Shannon diversity [23].

Statistical analyses

Detrended correspondence analysis (DCA) was performed to analyze correlation of the abundances of 16S rRNA gene sequences annotated at genus level and the environmental geochemical data with CANOCO (version 4.5; Microcomputer Power, Ithaca, NY). The indirect gradient method and unimodal scaling model were used to identify patterns of 16S rRNA gene sequence variations among sites and correlations between 16S rRNA gene sequence distribution and geochemical measurements. The sequences annotated in MG-RAST using the RDP [27]. File

used as input to the CANOCO contained sequence-by-sequence classification results for each sample with normalized sequence count for each taxon in the hierarchy. The genera detected in less than 50% of samples at low (<10 sequence) abundance per sample were omitted from analyses. For Bacteria, correlations between geochemical data from 10 samples and 293 Bacterial genera detected in these samples were studied. For Archaea, correlations between geochemical and sequencing data from 9 samples (no sequencing data were obtained from 5M sample) were studied. Only 12 Archaeal genera were detected in these samples. Sequencing data were used as the response variables, and the predictor variables were the measured environmental and geochemical parameters.

Results and Discussion

Mercury and methylmercury in rice paddy soils

The paddy soils collected from 6 geographic locations (Figure S1) varied in concentrations of THg ranging from 0.25 to 990 μg g^{-1} (mean: 131 ng g^{-1}) and MeHg ranging from 1.3 to 30.5 ng g^{-1} (mean: 9.4 ng g^{-1}; Table 1). Based on cumulative THg and MeHg concentrations, the samples were placed into 3 groups (Table 1 and Figure 1A). Soils from Jinjiachang (4J), Meizixi (5M), Dashuixi (6D) and one sample from Yanwuping (8Y) had the highest THg (14.3-990 μg g^{-1}) and MeHg (8.9-30.5 ng g^{-1}) concentrations, respectively, and were classified as High/High. Soils from Huaxi (control site, 1H), Gouxi (3G), and 2 samples from Yanwuping (14Y, 16Y) had low THg (0.25-6.5 μg g^{-1}) and low MeHg (1.3-4.9 ng g^{-1}) concentrations with the lowest being the Huaxi control site, and were classified as Low/Low (Table 1). Finally, two samples from Yanwuping (9Y) and Sikeng (7S) showed high THg and low MeHg and were classified as High/Low. While there is an overall trend showing an increase in MeHg with increasing THg concentration (Figure 1A) and decreasing %MeHg with increasing THg concentrations (Figure 1B), these correlations were not significant in the rice paddy soils studied. This is however not surprising as a number of previous studies have reported a similar trend. The percentage of MeHg in THg concentrations varied between 0.002 and 0.59% (mean: 0.2%) and was negatively correlated with increasing THg concentrations (Figure 1B). Overall, the percentages of MeHg in samples grouped as Low/Low were higher than in High/Low and High/High groups except for sample 4J where MeHg reached 0.21% and was equal to Low/Low group (Table 1).

Local geochemical factors including soil organic matter (SOM) and ionic compositions play an important role in the bioavailability and methylation of inorganic Hg [28] as do other soil properties such as soil pH that may affect Hg speciation and sorption onto minerals and soil organic matter. At pH 5 and lower, sorption of Hg-SOM complexes on minerals is higher, whereas increasing pH decreases Hg-SOM adsorption mainly because it increases SOM solubility [29]. In the rice paddy soils studied, SOM varied from 2.18 to 4.15% while pH increased with increasing in THg and SOM (Table 1, Figure 1C). However, with respect to MeHg, there were no clear correlations (Figure 1D). These data suggest that the increased pH may have increased the bioavailable pool of Hg for Hg-methylation, but there appeared to be no clear correlation between THg or MeHg with nitrate or sulfate concentrations.

Within the 10 samples studied, four soils were from Yanwuping and represented a distance gradient of 2.0-14.5 km away from the Hg mine waste area. The trends from the larger sample set held for increasing MeHg with increasing THg and decreasing %MeHg with increasing THg (Figure 2A). However, while there were no clear trends or correlations in the larger sample set for THg and geochemical factors, such correlations did appear in the Yanwuping samples (Figure 2B). While %SOM did not change appreciably with increasing THg, both pH and sulfate did increase, suggesting that more Hg may be bioavailable for Hg-methylation since sulfate was available for the stimulation of sulfate-reducing bacteria (SRB). High sulfate has been shown to stimulate Hg-methylation overall, implicating Hg-methylating SRB [30, 31]. Interestingly, nitrate and sulfate appeared to have an inverse relationship where nitrate was prevalent in low THg and sulfate was also low, but as THg increased nitrate was lower and sulfate concentrations increased almost 10 times. The effect of nitrate in rice paddies is unclear since nitrate can inhibit sulfate-reduction in some SRB while others can reduce nitrate to ammonia during sulfate-reduction [32-34]. With respect to increasing distance from the source, THg decreased from 990 to 0.83 mg/L and this decrease was reflected in the MeHg concentrations (Figure 2C). However, the %MeHg increased from 0.002-0.595%, suggesting a tipping point where either THg fell below an inhibitory concentration or the resident Hg-methylators could only methylate a given amount of Hg. This was further complicated by the changes over distance with increasing nitrate and decreasing sulfate and pH (Figure 2D). In fact, figures 2B and 2D are essentially mirror images of one another since higher THg concentrations were close to the source and decreased with distance.

Microbial α-diversity in samples with different concentrations of THg and MeHg

For all samples studied, the number of 16S rRNA gene pyrosequences (Table S3) obtained with Bacterial and Archaeal primers varied between 3,467-18,278 and 4,218-13,313 sequences per sample, respectively. On average, less than 10% of the sequences were chimeric or annotated as Eukaryotic and were not used in analyses. The α-diversity of Bacterial species ranged from 55.4 to 94.4, while α-diversity of Archaeal species was lower from 2.7 to 13.3 (Table S3). Bacterial α-diversity decreased with increasing THg while Archaeal α-diversity remained stable (Table S3). With respect to increasing MeHg, Bacterial α-diversity decreased (Figure 3A) while Archaeal α-diversity markedly increased (Figure 3B). Higher Bacterial α-diversity at low THg and MeHg is similar to previous findings [35] where rice fields near the Wanshan Hg mining area showed increased Bacterial α-diversity at low THg < 4.8 μg/g and MeHg < 18.62 ng/g. The correlation between Archaeal α-diversity and increasing MeHg may indicate a role for methanogenic Archaea in transforming metals and metalloids into their volatile derivatives [36] or that they are more resistant to higher MeHg concentrations. Only *Methanobacteria* (abundance of 0.04-0.3%) correlated ($P<0.05$) with MeHg concentration, other methanogens, for example *Methanomicrobia* (abundance of 4.6-49.9%), did not show any correlations ($P>0.05$) with MeHg. This may extend to a tolerance of higher THg concentrations [37], facilitating the methylation of Hg and giving these organisms a competitive advantage. Increased Hg pollution is usually considered to be an environmental stress and results in decreased microbial diversity with consequent changes in the microbial community structure. This trend was also evident in the four Yanwuping samples (8Y, 9Y, 14Y, 16Y; Figures 3A, 3B). With respect distance from the contaminant source, Bacterial α-diversity increased and the Archaeal α-diversity remained relatively stable, with decreasing THg and MeHg (Table S3, Figure 3C). As noted above, the %MeHg increased with distance (Figure 3C), which is not likely due to more Hg-methylators being present but rather a reflection of the marked drop in THg with increasing distance from the source.

Microbial community structure

A total of 24 Bacterial phyla were detected in >50% of the samples. The phylum *Proteobacteria* (33.5-47.7%) was the most abundant, followed by *Actinobacteria* (3.8-18.5%), *Chloroflexi* (4.5-14.5%), and *Acidobacteria* (6.2-14%) and these 4 phyla were dominant in all samples. Among

the phylum *Proteobacteria*, the most abundant was β-*Proteobacteria* (4.4-19.8%), followed by α- (7.1-18.8%), δ- (8.7-13.6%), and γ -*Proteobacteria* (2.1-5.9%) (Figure 4A). The δ-*Proteobacteria* contain the well-known sulfate- and Fe(III)- reducing bacteria capable of Hg-methylation [21, 37] and within the four Yanwuping samples there was no appreciable change in these organisms with distance from the contamination source.

For the *Archaea*, only phyla of *Euryarchaeota* (42.4-79.2%) and *Crenarchaeota* (6.2-31.9%) were detected, with a dominance of *Euryarchaeota* in all samples. The majority of *Crenarchaeota* sequences were affiliated with the class *Thermoprotei* and most of them were unclassified *Thermoproteales* (6.2-31.9%) and unclassified *Desulfurococcales* (<1.7%). Archaea of the phylum *Euryarchaeota* were much diverse and belonged to 6 classes (Figure 4B), which included 15 families. Eleven families (73.4%) comprised the methanogenic Archaea, with the others from sulfate-reducing, acidophilic, and halophilic Archaea.

While there are some differences between the microbial community structures of these geographically diverse samples, the similarities in the overall percent makeup of each community at the phylum level is strikingly similar (Figure 4A, B). Samples from Yanwuping showed some of the highest *Methanomicrobia* counts spanning the 14.5 Km from the source, which agreed with the relatively stable α-diversity over this distance (Figure 3B). Since all known Archaea Hg-methylators belong to this family [21, 37], the overall positive correlation suggests that they are more tolerant of higher THg and MeHg concentrations than non-methylating Archaea and as well as Bacteria.

Differences in microbial communities based on the concentrations of THg and MeHg

Since a number of species from Hg-methylating groups [37, 38] were detected in all samples, the microbial community structures from samples grouped based on THg and MeHg concentrations were compared using the STAMP software. The STAMP package is regularly used for statistical analyses of taxonomic or metabolic profiles for pairs or groups of samples [39]. Based on the STAMP analyses of group's taxonomic profiles, all samples contained species known to methylate Hg (eg. *Geobacter*, *Desulfovibrio* and *Syntrophus* spp.), and differences in their abundances were not significant ($P>0.05$) between samples. However, STAMP analyses did show abundance differences in High/High communities for *Verrucomicrobium* ($P=0.024$),

Methylobacter (P=0.029) and *Butyrivibrio* (P=0.044) in comparison with Low/Low samples (Figure S2A). Both genera have been observed in Hg-contaminated sites [40, 41] but show no evidence of Hg-methylation or MeHg-demethylation. Other bacteria abundant in High/High samples, for example *Butyrivibrio*, may cause limited Hg(II)-reduction or Hg(0) volatilization as was shown for the rumen bacterium *Butyrivibrio fibriosolvens* [42]. On the other hand, Low/Low communities possessed higher abundances of *Campylobacter* (P=0.026), *Acidithiobacillus* (P=0.032) and *Hyphomicrobium* (P=0.05). No Hg-cycling capability has been shown for these bacteria, but the *Hyphomicrobium denitrificans* genome showed the presence of mercuric (II) reductase (*merA*) [43].

Compared to the High/Low to High/High communities, *Micrococcus* (P=1.08e^{-3}) and *Halothiobacillus* (P=0.049) were more abundant in the former (Figure S2B), suggesting Hg tolerance but not methylation. Indeed, *Halothiobacillus hydrothermalis* from deep sea hydrothermal vents can resist up to 10 μM Hg(II), but does not possess any *mer* or *hgcAB* genes or orthologues and cannot reduce or oxidize Hg [44]. In contrast, comparing the High/Low and Low/Low, *Desulfobacterium* (P=0.02) and *Verrucomicrobium* (P=0.045) were more abundant in High/Low samples, while unclassified *Desulfuromonadales* (P=5.06e^{-3}), *Bradyrhizobium* (P=0.011), *Pseudonocardia* (P=0.02) were more abundant in Low/Low samples (Figure S2C). The discovery of more unclassified *Desulfuromonadales* in Low/Low samples may suggest that not all species methylate Hg or that the methylators are not prevalent in natural environments. This is consistent with *Desulfuromonadales* being a minor group in the Hg-methylating community of the Florida everglades [45] and *Bradyrhizobium canariense* being able to continue nitrogenase activity at $HgCl_2$ concentrations up to 200 μM [46].

Using the GeoChip v4 [17], a total of 30,851±1,052 probes were detected for each sample and a multivariate statistical analysis was used to determine if the microbial communities in the rice paddy soils showed similar groupings or a trend based on the concentrations of THg and MeHg (Figure S3). There were no clear groupings of the different samples, thus suggesting that THg or MeHg were not major factors contributing to the differences observed in the various microbial community structures. Communities from High/High samples were situated in the middle of the plot in close proximity to each other. Based on DCA results, 22.9% of the variation between bacterial communities could be explained by changes in the percentage of MeHg in the THg

concentrations while variations along the second axis (8.6%) could reflect differences in THg. Overall, there was no appreciable difference in *mer* gene counts between any samples nor for *merA* (mercuric reductase), *merB* (organomercurial lyase), *merP* (mercuric transport protein) or *merT* (mercuric transport protein, Table 2). The mercuric reductase (*merA*) comprised 79% of the *mer* gene counts while *merB*, known for MeHg demethylation [47], only accounted for 6% of the *mer* genes, suggesting that Hg(II) reduction to Hg(0) may be more prevalent than demethylation at these sites. This marked difference is odd given that these genes typically appear side-by-side in the operon and so this may suggest that organisms in these areas have operons where some genes have been removed. The mercuric transporters *merP* and *merT* were in similar abundance to *merB* with 11% and 3% of the gene counts. Regarding sulfate-reduction, the *apsAB* genes also did not have the expected 1:1 ratio but were present in a 3:1. Reasons for this are unknown but are far closer than the 13:1 observed with *merAB* (Table 2). The related *dsrAB* genes for sulfite-reduction appeared in a ~1:1 ratio (*dsrA* 59%, *dsrB* 41%) as would be expected and were present at twice the abundance of the *mer* genes (Table 2). The *mcrA* gene, indicative of methanogens was the least abundant at 68 counts per sample. Unfortunately, there is no explicit marker gene for *Geobacter* spp., however a total of 207 hits were detected (data not shown). These data as well as the increase in *dsrAB* relative abundance in High/High samples compared to High/Low ($p=0.095$ and 0.079, respectively) suggests that the sulfate-reducing bacteria, well-known for Hg-methylation, may be a group of active Hg-methylators at these sites, and that possible increased Hg/MeHg cycling is occurring.

Diversity of hgcAB and quantification of hgcA in rice paddy soils

The *hgcAB* genes are involved in Hg-methylation [21] and *hgcAB* was amplified from each sample, the amplicons cloned and sequenced for assessing Hg-methylator diversity while *hgcA* was quantified, using our recently reported protocol [20]. For *hgcAB*, five clones were selected for sequencing and the top hit from BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) is presented. In most cases, all five clones were sequenced successfully however in some cases as little as one sequence was obtained (Table S4). The vast majority of the sequences were from uncultured organisms with *hgcA* being originally detected from a different study. However, where the sequences did align with a known organism, they were a lone SRB, *Desulfobulbus propionicus*, or a methanogen *Methanoregula formicicum*, *Methanocella paludicola*, or one of several *Methanomicrobia* enrichment *hgcA* clones. This prevalence of unknown/uncultured organisms

with some signal from methanogens or *Desulfobulbus propionicus* was consistent along the distance gradient for the Yanwuping samples as well.

Overall the abundance of *hgcA* increased with increasing THg and MeHg concentrations (Figure 5), however, these relationships were variable depending upon the site. The *hgcA* signal was evenly split between the *Deltaproteobacteria* and the Archaea, so the cumulative abundance of *hgcA* is presented herein. While *hgcA* did increase with MeHg (Figure 5B), it appears that there may be a plateau in that correlation in the vicinity of 20 ng/g sediment. The decreasing bacterial diversity with increasing THg concentration (Figure 3A) and the concomitant increase in *hgcA* (Figure 5A) suggests that the Bacterial community may have been comprised of higher ratios of Hg tolerant bacteria. The positive linear correlation observed between *hgcA* gene abundance and MeHg content (Figure 5B) indicates that Hg-methylation is likely occurring in the rice paddy soils, but is also further evidence of the likelihood of increased cycling of Hg-methylation/MeHg-demethylation. With respect to the Yanwuping samples, *hgcA* counts followed the same trend of correlating with THg and MeHg concentrations. In this case, as THg and MeHg concentrations decreased with increasing distance from the source, *hgcA* counts decreased as well (Figure 5C). Quantitative PCR for *Firmicutes* was below detection in all cases.

Correlation between microbial communities and hgcA with variations in THg and MeHg.

Hg biotransformations are spread across Bacterial groups involving intertwined microbial processes including sulfate- [10] and iron- reduction [48], methanogenesis, acetogenesis, and syntrophy [49, 50]. This innate complexity of the *in-situ* Hg-biotransformation processes in natural complex microbial communities extends from the known Bacteria and Archaea to many unclassified and unassigned microorganisms. Hg-methylation has been most extensively demonstrated among *Proteobacteria* and *Firmicutes* [37, 38] to date.

DCA results showed positive correlations between 16S rRNA gene abundance with concentrations of THg, MeHg, and *hgcA* for a number of δ-*Proteobacterial methylators* such as *Geobacter, Desulfobulbus, Desulfovibrio, Desulfuromonas,* and *Syntrophus* as well as several non-methylators including *Desulfonema, Desulfopila, Desulforhopalus, Desulfomonile, Pelobacter, Desulfovirga, Syntrophobacter, Syntrophorhabdus, Smithella, Bilophila* and *Myxococcus* (Figure 6A). Interestingly, Hg-methylating genera such as *Desulfonatronospira*, *Desulfobacter*, and *Desulfococcus* were not detected in this study, perhaps due to the low salinity

in paddy soils of Guizhou province, China. Previously, these three genera have predominantly been found in brackish, marine environments and hypersaline soda lakes [51, 52]. Nine of the genera listed above, *Geobacter, Desulfonema, Desulfobulbus, Desulfuromonas, Desulfovibrio, Syntrophobacter, Syntrophus, Smithella, and Bilophila*, are capable of sulfate- or sulfite-reduction. The high diversity of these types of organisms in rice paddy soils supports the general opinion that such physiological types may be major contributors to MeHg pools in these soils. The most abundant genus was *Geobacter* ranging from 0.6% to 2.6% of the total Bacterial community, followed by *Desulfobulbus* (0.02-0.9%), *Syntrophobacter* (0.02-0.18%), and *Syntrophus* (0.005-0.12%), all known to contain Hg-methylating species [37].

Four *Firmicute* genera, *Anaerovorax* (<0.08%), *Clostridium* XIVa (<0.09%), *Fusibacter* (<0.05%) and *Saccharofermentans* (<0.12%), showed positive correlations with increasing THg, MeHg, and *hgcA* concentrations (Figure 6B) are not known to methylate Hg. Similarly, seven *Verrucomicrobia* genera positively correlated with increasing *hgcA* gene copy number, THg and MeHg concentrations, suggesting they may be able to transform Hg species or may benefit from association with Hg-methylating organisms. Among these, the most abundant was Subdivision 3 (1.4-4.5%), followed by *Luteolibacter* (0.17-2.67%), *Opitutus* (0.02-0.39%), Subdivision 5 (0.07-0.38%), and *Haloferula* (0.09-0.22%), *Prosthecobacter* (<0.11%), and *Alterococcus* (<0.08%) (Figure 6B). Relatively little is known about this phylum and although the occurrence of *Prosthecobacter* and *Alterococcus* showed correlations with THg, these organisms were not detected in all samples. The α-*Proteobacteria* and *Chloroflexi* did not exhibit any correlations with either THg or MeHg.

Among the Archaea, *Methanolinea* (0.02-2.5%), *Thermogymnomonas* (0.01-2.5%), and *Methanoculleus* (<0.12%) correlated with increases in *hgcA* copy number, THg, MeHg concentrations, and pH (Figure 6C). However, *Methanosaeta* (2.7-33.8%), *Methanosarcina* (0.9-12.1%), *Methanosphaera* (0.03-0.3%), *Methanospirillum* (<0.15%), and *Methanomicrococcus* (<0.07%) were located between THg, MeHg, *hgcA* and SOM on the DCA ordination plot (Figure 6C), suggesting that these Archaea contribute to decomposition of complex organic substances into CH_4 and CO_2 [53]. Methanogens of the orders *Methanobacteriales*, *Methanococcales*, and *Methanosarcinales* were suggested as the primary Hg-methylators in lake periphyton since the inhibition of methanogenesis totally inhibited methylation [54]. Sequences affiliated with the

Archaea *Methanolobus* and *Methanomethylovorans*, which have been positively tested for Hg-methylation [37], were only detected in two samples with high THg concentrations.

Conclusions

Several studies have pointed to the accumulation of MeHg in rice and the implications where rice as well as seafood are dietary staples [13, 55-58] with some evidence for methanogens as the primary Hg-methylators [59]. The high abundance of both sulfate-reducing bacteria (in soils with relatively high sulfate concentrations) and methanogens suggests that these microorganisms may be the primary Hg-methylators in the rice paddy soils in Guizhou province, China. The implication of sulfate-reducing bacteria is consistent with previous studies from wastewater treatment plants and estuaries [38] as well as from natural anoxic habitats such as salt marshes, estuarine, marine and freshwater sediments [10, 60-62] and *Sphagnum* moss mats [63]. The methanogenic Archaea appear to be the primary Hg-methylators in lake periphyton [54]. It is well-known that SRB and methanogens can live syntrophically [49, 64-67] and species within each group can methylate Hg(II). This is due to metabolic interdependence with the interspecies transfer of formate, H_2 and acetate, providing carbon and electrons to both species [49]. Similar synergistic growth may create favorable conditions for Hg-methylation in low-sulfate anoxic sediments where methylation might not be predicted to occur. These earlier findings support this study, which demonstrates that diverse sulfate- and sulfite-reducing bacteria and methanogens can coexist in natural environments and concurrently contribute to the formation of neurotoxic methylmercury.

Acknowledgments

We thank Zamin Yang for help with 454 FLX pyrosequencing and Xiangping Yin for geochemistry analyses. This research was sponsored in part by the Office of Biological and Environmental Research, Office of Science, US Department of Energy (DOE) as part of the Mercury Science Focus Area at Oak Ridge National Laboratory, which is managed by UT-Battelle LLC for the DOE under contract DE-AC05-00OR22725. HH, XX, GQ, and XF were supported by National Natural Science Foundation of China under grant 41303098.

Figure Legends

Figure 1. (A) The relationship between THg and MeHg concentrations in rice paddy soils, P = 0.199; (B) the percentage of MeHg in the THg decreased in samples with high THg concentration; (C) Soil organic matter and (D) pH as a function of the THg concentrations.

Figure 2. The relationship between THg and (A) MeHg and %MeHg/THg or (B) geochemical parameters in all rice paddy samples; the effect of distance in four Yanwuping samples on (C) THg, MeHg and %MeHg/THg as well as (D) geochemical parameters.

Figure 3: Relationships between Alpha diversity and MeHg in (A) Bacteria and (B) Archaea for all rice paddy samples, and (C) the effect of distance on THg, MeHg and Bacteria and Archaea Alpha diversity.

Figure 4. Microbial diversity detected in samples of rice paddy soils from Guizhou, China. (A) The relative abundances of dominant *Bacterial* phyla. Phylum *Proteobacteria* is presented at class level. Phyla of *Fibrobacteres*, SR1, *Fusobacteria*, *Chlamydiae*, OP11, OD1, TM7, *Lentisphaerae*, *Spirochaetes*, BRC1 were present at abundances below 0.7% and were combine together in to group "Other". (B) The relative abundances of dominant Archaea classes. Number of sequences identified at class level contained both identified and non-identified at genus level sequences. Only 12 sequences amplified with Archaea primers were obtained from sample 5M, 3 sequences were affiliated with class *Methanomicrobia*, and others with unclassified *Euryarchaeota*.

Figure 5: *hgcA* gene abundance correlated to (A) THg or (B) MeHg in all rice paddy soils studied and (C) as compared to THg and MeHg as a function of distance from the Hg source in the four Yanwuping samples.

Figure 6: Detrended correspondence analyses for the three primary Hg-methylating clades of (A) *Deltaproteobacteria*, (B) *Firmicutes* and (C) Archaea using 16S rRNA gene data.

References

1. J. E. Sonke, L.-E. Heimburger and A. Dommergue, *Comptes Rendus Geoscience*, 2013, **345**, 213–224.
2. UNEP, *UNEP Global Mercury Assessment 2013: Sources, Emissions, Releases and Environmental Transport*, United Nations Environment Programme Chemicals Branch, Geneva, Switzerland, 2013.
3. L. Zhang, S. Wang, L. Wang, Y. Wu, L. Duan, Q. Wu, F. Wang, M. Yang, H. Yang, J. Hao and X. Liu, *Environ Sci Technol*, 2015, **49**, 3185-3194.
4. N. Pirrone, S. Cinnirella, X. Feng, R. B. Finkelman, H. R. Friedli, J. Leaner, R. Mason, A. B. Mukherjee, G. B. Stracher, D. G. Streets and K. Telmer, *Atmospheric Chemistry and Physics*, 2010, **10**, 5951–5964.
5. X. Feng and G. Qiu, *Sci Total Environ*, 2008, **400**, 227-237.
6. R. L. Worden, A. M. Savada and R. E. Dolan, in *China: a country study*, Library of Congress. Federal Research Division, Washington D.C., 1987, ch. 286-292, p. 637.
7. J. J. Ewing and Z. Hongzhou, 2013.
8. H. Zhang, X. Feng, T. Larssen, G. Qiu and R. D. Vogt, *Environmental Health Perspectives*, 2010, **118**, 1183–1188.
9. G. C. Compeau and R. Bartha, *Appl Environ Microbiol*, 1985, **50**, 498-502.
10. C. C. Gilmour and E. A. Henry, *Environ Poll*, 1991, **71**, 131-169.
11. L. Zhang and M. H. Wong, *Environment International*, 2007, **33**, 108–121.
12. X. Feng, P. Li, G. Qiu, S. Wang, G. Li, L. Shang, B. Meng, H. Jiang, W. Bai, Z. Li and X. Fu, *Environ Sci Technol*, 2008, **42**, 326-332.
13. Y. D. Tong, L. B. Ou, L. Chen, H. H. Wang, C. Chen, X. J. Wang, W. Zhang and Q. G. Wang, *Environ Toxicol Chem*, 2015, **34**, 1161-1168.
14. A. J. King, S. P. Preheim, K. L. Bailey, M. S. Robeson, T. Roy Chowdhury, B. R. Crable, R. A. Hurt, T. Mehlhorn, K. A. Lowe, T. J. Phelps, A. V. Palumbo, C. C. Brandt, S. D. Brown, M. Podar, P. Zhang, W. A. Lancaster, F. Poole, D. B. Watson, M. W. Fields, J.-M. Chandonia, E. J. Alm, J. Zhou, M. W. W. Adams, T. C. Hazen, A. P. Arkin and D. A. Elias, *Environmental Science & Technology*, 2017, **51**, 2879-2889.
15. W. A. Lancaster, A. L. Menon, I. Scott, F. L. Poole, B. Vaccaro, M. P. Thorgersen, J. Geller, T. C. Hazen, R. A. Hurt, S. D. Brown, D. A. Elias and M. W. W. Adams, *Metall*, 2014, **6**, 1004-1013.
16. R. A. Hurt, M. S. Robeson, M. Shakya, J. G. Moberly, T. A. Vishnivetskaya, B. Gu and D. A. Elias, *PLoS ONE*, 2014, **9**, e102826.
17. Q. Tu, H. Yu, Z. He, Y. Deng, L. Wu, J. D. Van Nostrand, A. Zhou, J. Voordeckers, Y. J. Lee, Y. Qin, C. L. Hemme, Z. Shi, K. Xue, T. Yuan, A. Wang and J. Zhou, *Mol Ecol Resour*, 2014, **14**, 914-928.
18. Y. Liang, Z. He, L. Wu, Y. Deng, G. Li and J. Zhou, *Appl Environ Microbiol*, 2010, **76**, 1088-1094.
19. Z. Lu, Z. He, V. A. Parisi, S. Kang, Y. Deng, J. D. Van Nostrand, J. R. Masoner, I. M. Cozzarelli, J. M. Suflita and J. Zhou, *Environ Sci Technol*, 2012, **46**, 5824-5833.
20. G. A. Christensen, A. M. Wymore, A. J. King, M. Podar, J. R. A. Hurt, E. U. Santillan, A. Soren, C. C. Brandt, S. D. Brown, A. V. Palumbo, J. D. Wall, C. C. Gilmour and D. A. Elias, *Appl. Environ. Microbiol.*, 2016, **82**, 6068-6078.

21. J. M. Parks, A. Johs, M. Podar, R. Bridou, R. A. Hurt, Jr., S. D. Smith, S. J. Tomanicek, Y. Qian, S. D. Brown, C. C. Brandt, A. V. Palumbo, J. C. Smith, J. D. Wall, D. A. Elias and L. Liang, *Science*, 2013, **339**, 1332-1335.
22. J. G. Caporaso, J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman and E. K. Costello, *Nat Meth*, 2010, **7**, 335-336.
23. F. Meyer, D. Paarmann, M. D'Souza, R. Olson, E. Glass, M. Kubal, T. Paczian, A. Rodriguez, R. Stevens, A. Wilke, J. Wilkening and R. Edwards, *BMC Bioinformatics*, 2008, **9**, 386.
24. M. J. Claesson, O. O'Sullivan, Q. Wang, J. Nikkila, J. R. Marchesi, H. Smidt, W. M. de Vos, R. P. Ross and P. W. O'Toole, *PLoS One*, 2009, **4**, e6669.
25. S. Anders and W. Huber, *Genome Biol*, 2010, **11**, R106.
26. P. J. McMurdie and S. Holmes, *PLoS Comput Biol*, 2014, **10**, e1003531.
27. Q. Wang, G. M. Garrity, J. M. Tiedje and J. R. Cole, *Appl Environ Microbiol*, 2007, **73**, 5261-5267.
28. A. Andersson, in *The biogeochemistry of mercury in the environment*, ed. J. Nriagu, Biomedical Press, 1979.
29. Y. Yin, H. E. Allen, Y. Li, C. P. Huang and P. F. Sanders, *J Environ Qual*, 1996, **25**, 837–844.
30. R. P. Mason, C. L. Miller, C. Gilmour and A. Heyes, *Abstracts of Papers of the American Chemical Society*, 2002, **223**, U525-U525.
31. C. C. Gilmour, E. A. Henry and R. Mitchell, *Environ Sci Technol*, 1992, **26**, 2281-2287.
32. T. Dalsgaard and F. Bak, *Appl. Environ. Microbiol.*, 1994, **60**, 291-297.
33. C. C. Hwang, W. M. Wu, T. J. Gentry, J. Carley, G. A. Corbin, S. L. Carroll, D. B. Watson, P. M. Jardine, J. Z. Zhou, C. S. Criddle and M. W. Fields, *ISME J*, 2009, **3**, 47-64.
34. H. D. Kluber and R. Conrad, *FEMS Microbiol Ecol*, 1998, **25**, 301-318.
35. Y. R. Liu, J. J. Wang, Y. M. Zheng, L. M. Zhang and J. Z. He, *Microb Ecol*, 2014, **68**, 575-583.
36. J. Meyer, K. Michalke, T. Kouril and R. Hensel, *Syst Appl Microbiol*, 2008, **31**, 81-87.
37. C. C. Gilmour, M. Podar, A. L. Bullock, A. M. Graham, S. D. Brown, A. C. Somenahally, A. Johs, J. R. A. Hurt, K. L. Bailey and D. A. Elias, *Environ. Sci. Technol.*, 2013, **47**, 11810-11820.
38. M. Ranchou-Peyruse, M. Monperrus, R. Bridou, R. Duran, D. Amouroux, J. C. Salvado and R. Guyoneaud, *Geomicrobiol*, 2009, **26**, 1-8.
39. D. H. Parks, G. W. Tyson, P. Hugenholtz and R. G. Beiko, *Bioinformatics*, 2014, **30**, 3123-3124.
40. R. Yu, Doctor of Philosophy Ph.D. Dissertation, The State University of New Jersey, 2011.
41. T. A. Vishnivetskaya, J. J. Mosher, A. V. Palumbo, Z. K. Yang, M. Podar, S. D. Brown, S. C. Brooks, B. H. Gu, G. R. Southworth, M. M. Drake, C. C. Brandt and D. A. Elias, *Appl Environ Microbiol*, 2011, **77**, 302-311.
42. S. Kozak and C. W. Forsberg, *Appl Environ Microbiol*, 1979, **38**, 626-636.
43. R. Venkatramanan, O. Prakash, T. Woyke, P. Chain, L. A. Goodwin, D. Watson, S. Brooks, J. E. Kostka and S. J. Green, *Genome announcements*, 2013, **1**.
44. K. Cruz, M. Crespo-Medina, S. Borin, R. Cruz, C. Vetriani and T. Barkay, Halifax , Nova Scotia, Canada 2011.

45. H. S. Bae, F. E. Dierberg and A. Ogram, *Appl Environ Microbiol*, 2014, **80**, 6517-6526.
46. M. A. Quiñones, B. Ruiz-Díez, S. Fajardo, M. A. López-Berdonces, P. L. Higueras and M. Fernández-Pascual, *Plant Physiology and Biochemistry*, 2013, **73**, 168-175.
47. J. K. Schaefer, J. Yagi, J. R. Reinfelder, T. Cardona, K. M. Ellickson, S. Tel-Or and T. Barkay, *Environmental Science & Technology*, 2004, **38**, 4304-4311.
48. E. J. Fleming, E. E. Mack, P. G. Green and D. C. Nelson, *Appl Environ Microbiol*, 2006, **72**, 457-464.
49. K. Pak and R. Bartha, *Appl Environ Microbiol*, 1998, **64**, 1987-1990.
50. T. Barkay and I. Wagner-Dobler, *Adv Appl Microbiol*, 2005, **57**, 1-52.
51. D. Y. Sorokin, T. P. Tourova, A. M. Henstra, A. J. Stams, E. A. Galinski and G. Muyzer, *Microbiol*, 2008, **154**, 1444-1453.
52. L. L. Barton, in *Biotechnology handbooks*, Plenum Press, New York, 1995, vol. 8.
53. F. Ali Shah, Q. Mahmood, M. Maroof Shah, A. Pervez and S. Ahmad Asad, *The Scientific World Journal*, 2014, **2014**, 183752.
54. S. Hamelin, M. Amyot, T. Barkay, Y. Wang and D. Planas, *Environ Sci Technol*, 2011, **45**, 7693-7700.
55. M. Meng, B. Li, J. J. Shao, T. Wang, B. He, J. B. Shi, Z. H. Ye and G. B. Jiang, *Environ Pollut*, 2014, **184**, 179-186.
56. S. E. Rothenberg and X. B. Feng, *J. Geophys. Res.-Biogeosci.*, 2012, **117**, 16.
57. H. Zhang, X. Feng, T. Larssen, L. Shang and P. Li, *Environ. Sci. Technol.*, 2010, **44**, 4499-4504.
58. H. Zhang, X. Feng, T. Larssen, G. Qiu and R. D. Vogt, *Environmental Health Perspectives*, 2010, **118**, 1183-1188.
59. S. Sakai, H. Imachi, S. Hanada, A. Ohashi, H. Harada and Y. Kamagata, *International Journal of Systematic and Evolutionary Microbiology*, 2008, **58**, 929-936.
60. G. C. Compeau and R. Bartha, *Appl Environ Microbiol*, 1985, **50**, 498-502.
61. J. K. King, J. E. Kostka, M. E. Frischer and F. M. Saunders, *Appl Environ Microbiol*, 2000, **66**, 2430-2437.
62. J. M. Benoit, C. C. Gilmour and R. P. Mason, *Appl Environ Microbiol*, 2001, **67**, 51-58.
63. R. Q. Yu, I. Adatto, M. R. Montesdeoca, C. T. Driscoll, M. E. Hines and T. Barkay, *FEMS Microbiol. Ecol.*, 2010, **74**, 655-668.
64. M. P. Bryant, L. L. Campbell, C. A. Redy and M. R. Crabill, *Appl Environ Microbiol*, 1977, **33**, 1162-1169.
65. M. J. Mcinerney and M. P. Bryant, *Appl. Environ. Microbiol.*, 1981, **41**, 346-354.
66. K. Pak and R. Bartha, *Bull Environ Contamin Toxicol*, 1998, **61**, 690-694.
67. A. Traore, M.-L. Fardeau, C. E. Hatchikian, J. LeGall and J.-P. Belaich, *Appl Environ Microbiol*, 1983, **46**, 1152-1156.

Table 1. Soil samples, sample description, and geochemical properties

Sample	Site	Description	pH	Chloride (mg/g)	Nitrate (mg/g)	Sulfate (mg/g)	SOM (%)	THg (μg/g)	MeHg (ng/g)	MeHg in THg (%)	Sample Grouping THg / MeHg
1H	Huaxi	Control site	6.74	0.03	0.02	0.41	4.15	0.25	1.30	0.52	Low / Low
8Y	Yanwuping	2 km downstream of Hg mine waste	7.09	0.05	0.02	2.51	2.94	990	20.75	0.002	High / High
9Y	Yanwuping	3 km downstream of Hg mine waste	7.29	0.05	0.02	0.69	3.08	38.5	3.81	0.01	High / Low
14Y	Yanwuping	10 km downstream of Hg mine waste	5.75	0.05	0.08	0.19	2.18	6.53	5.17	0.08	Low / Low
16Y	Yanwuping	14.5 km downstream of Hg mine waste	4.83	0.04	0.42	0.28	3.25	0.83	4.94	0.59	Low / Low
3G	Gouxi	0.2 km upstream of abandoned Hg mine	6.93	0.05	0.05	0.74	3.22	4.41	6.02	0.14	Low / Low
7S	Sikeng	0.1 km downstream of Hg mine waste	7.29	0.04	0.07	0.68	2.48	74.5	2.73	0.004	High / Low
4J	Jinjiachang	8 km downstream of Hg mine waste at Shibakeng	6.85	0.03	0.05	0.83	3.36	14.3	30.52	0.21	High / High
6D	Dashuixi	0.2 km downstream of Hg mine waste at Wukeng	7.63	0.05	0.02	0.77	3.34	68.35	9.97	0.01	High / High
5M	Meizixi	0.2 km from active Hg mine at Gouxi	6.91	0.06	0.01	0.79	3.39	112	8.96	0.01	High / High

Table 2: Comparison of gene counts in Rice Paddy Samples according to Geochip Analysis

	Sample ID														
gene	**1H**	**8Y**	**9Y**	**14Y**	**16Y**	**3G**	**7S**	**4J**	**6D**	**5M**	**Average gene count in Yanwuping**	**Standard Deviation**	**% of gene count**	**Average gene count all samples**	**Standard Deviation**
merA	248	243	255	248	222	237	249	246	238	244	242	14	79	243	9
merB	18	20	19	18	19	21	20	19	20	18	19	1	6	19	1
merC	3	3	3	4	4	2	2	3	3	2	4	1	1	3	1
merF	0	0	1	0	1	0	0	0	0	0	1	1	0	0	0
merG	2	1	2	2	1	2	1	2	1	2	2	1	0	2	1
merP	31	37	34	32	33	33	35	33	34	32	34	2	11	33	2
merT	6	8	8	8	7	8	8	8	8	7	8	1	3	8	1
total	308	312	322	312	287	303	315	311	304	305	308	15	100	308	9
apsA	111	104	109	102	101	101	117	112	107	111	104	4	77	108	5
apsB	32	31	33	30	27	25	32	32	32	27	30	3	23	30	3
total	143	135	142	132	128	126	149	144	139	138	134	6	100	138	7
dsrA	365	390	384	364	345	346	385	371	359	348	371	20	59	366	17
dsrB	249	255	264	262	238	248	266	258	252	250	255	12	41	254	9
total	614	645	648	626	583	594	651	629	611	598	626	30	100	620	24
mcrA	57	73	70	69	59	62	69	73	75	61	68	6	NA	67	6

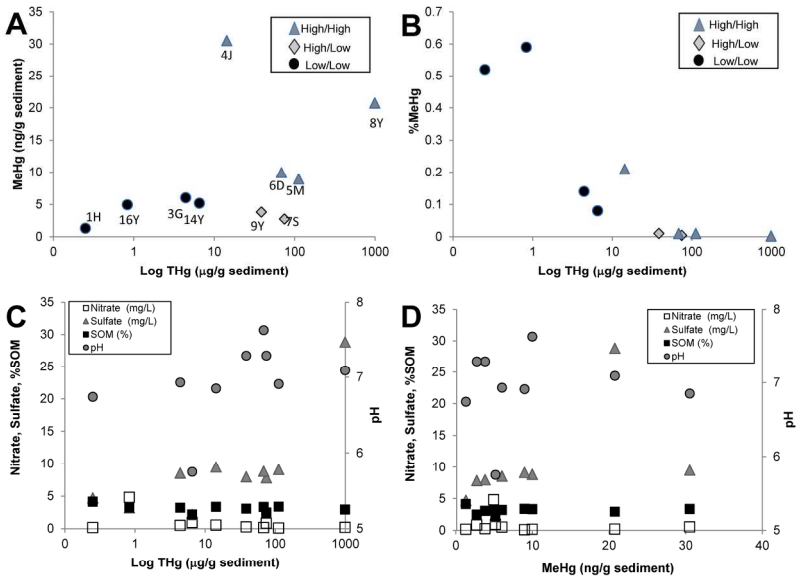


Figure 1. (A) The relationship between THg and MeHg concentrations in rice paddy soils, P = 0.199; (B) the percentage of MeHg in the THg decreased in samples with high THg concentration; (C) Soil organic matter and (D) pH as a function of the THg concentrations

215x166mm (300 x 300 DPI)

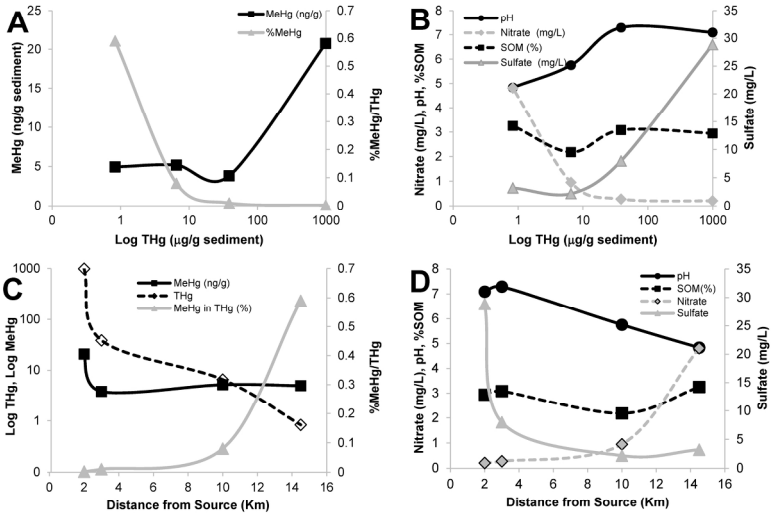


Figure 2. The relationship between THg and (A) MeHg and %MeHg/THg or (B) geochemical parameters in all rice paddy samples; the effect of distance in four Yanwuping samples on (C) THg, MeHg and %MeHg/THg as well as (D) geochemical parameters.

215x166mm (300 x 300 DPI)

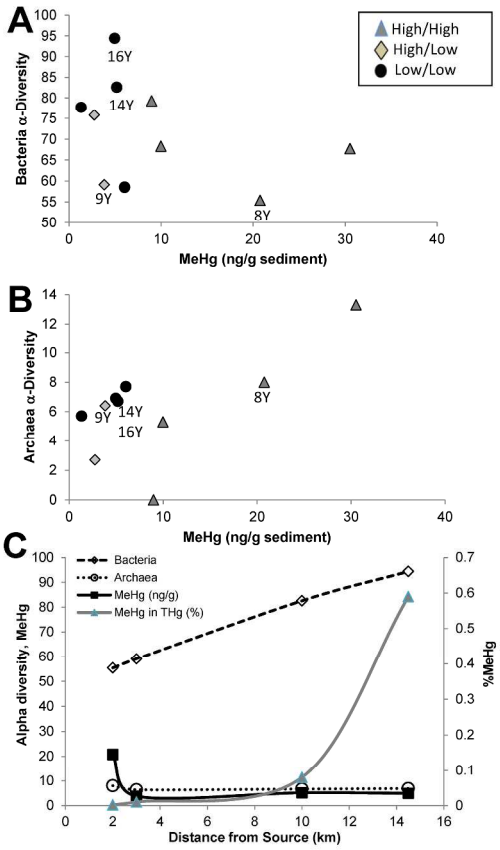


Figure 3: Relationships between Alpha diversity and MeHg in (A) Bacteria and (B) Archaea for all rice paddy samples, and (C) the effect of distance on THg, MeHg and Bacteria and Archaea Alpha diversity.

279x361mm (300 x 300 DPI)

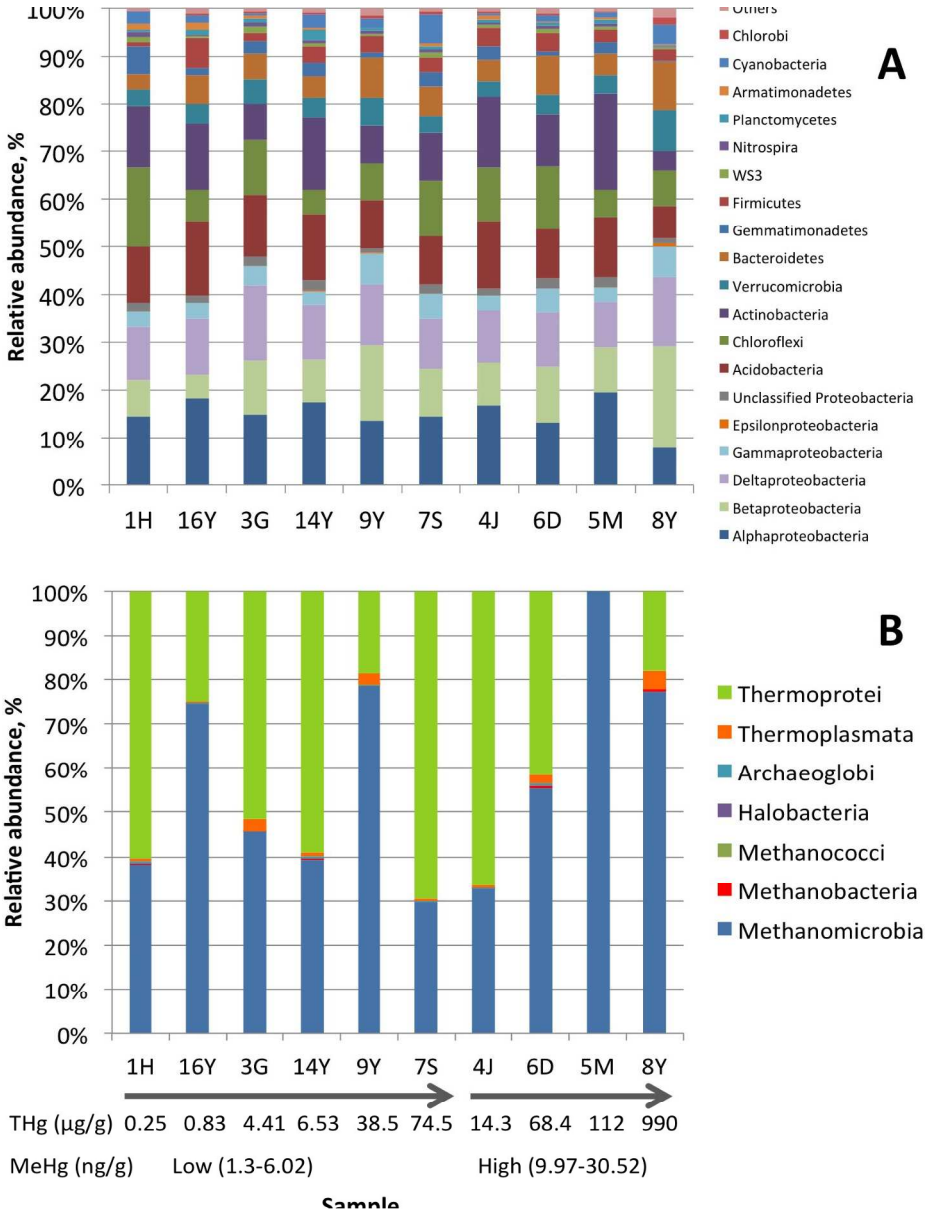


Figure 4. Microbial diversity detected in samples of rice paddy soils from Guizhou, China. (A) The relative abundances of dominant Bacterial phyla. Phylum *Proteobacteria* is presented at class level. Phyla of *Fibrobacteres*, SR1, *Fusobacteria*, *Chlamydiae*, OP11, OD1, TM7, *Lentisphaerae*, *Spirochaetes*, BRC1 were present at abundances below 0.7% and were combine together in to group "Other". (B) The relative abundances of dominant Archaea classes. Number of sequences identified at class level contained both identified and non-identified at genus level sequences. Only 12 sequences amplified with Archaea primers were obtained from sample 5M, 3 sequences were affiliated with class *Methanomicrobia*, and others with unclassified *Euryarchaeota*.

279x361mm (300 x 300 DPI)

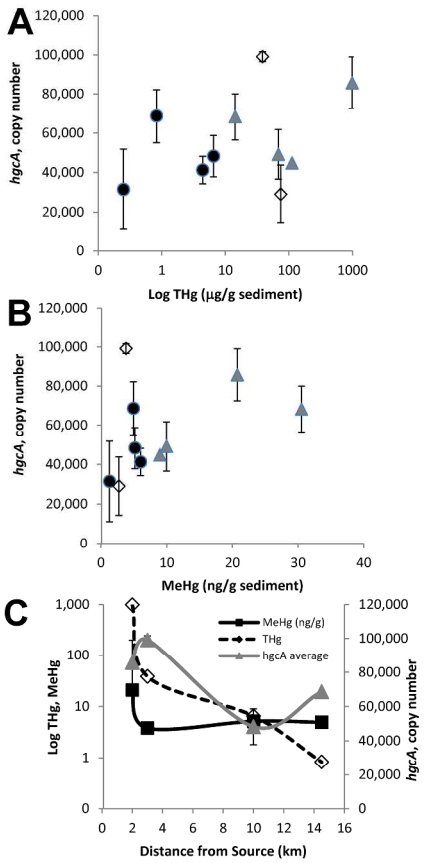


Figure 5: *hgcA* gene abundance correlated to (A) THg or (B) MeHg in all rice paddy soils studied and (C) as compared to THg and MeHg as a function of distance from the Hg source in the four Yanwuping samples.

279x361mm (300 x 300 DPI)

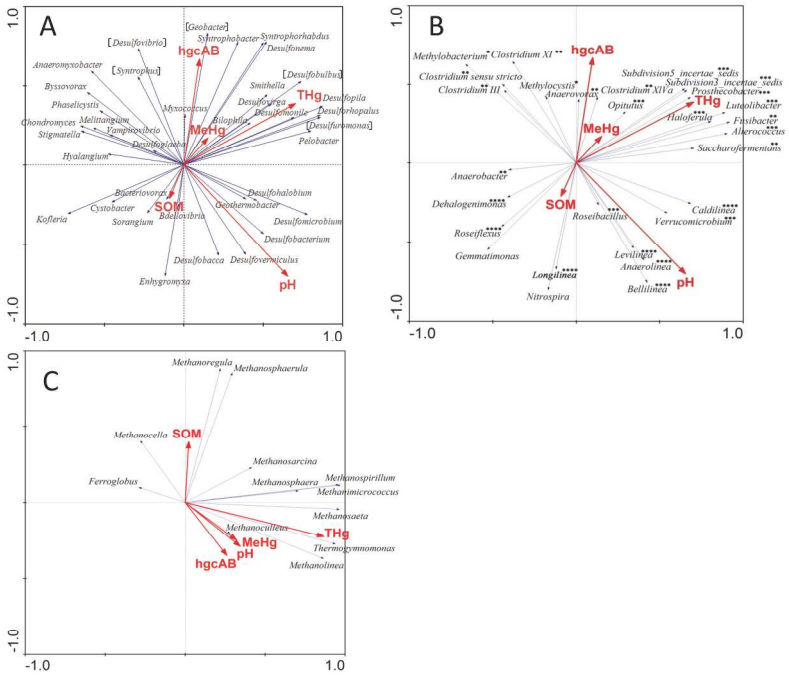
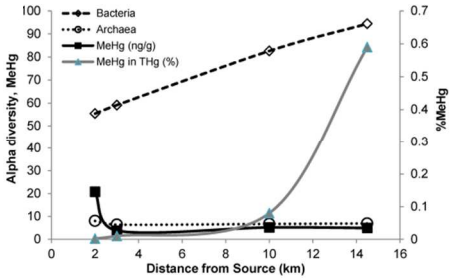


Figure 6: Detrended correspondence analyses for the three primary Hg-methylating clades of (A) *Deltaproteobacteria*, (B) *Firmicutes* and (C) Archaea using 16S rRNA gene data.

215x166mm (300 x 300 DPI)



Sulfate-reducing bacteria and methanogens are the primary Hg-methylators in Chinese rice paddies and methanogens are more abundant with high methylmercury.