Diversity and Abundance of Arsenic Biotransformation Genes in Paddy Soils from Southern China

Si-Yu Zhang,†‡ Fang-Jie Zhao,§∥ Guo-Xin Sun,† Jian-Qiang Su, Xiao-Ru Yang, Hu Li,‡ and Yong-Guan Zhu*†‡#

†State Key Lab of Urban and Regional Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, People’s Republic of China
‡University of Chinese Academy of Sciences, Beijing 100049, People’s Republic of China
§Jiangsu Collaborative Innovation Center for Solid Organic Waste Resource Utilization, College of Resources and Environmental Sciences, Nanjing Agricultural University, Nanjing 210095, People’s Republic of China
∥Sustainable Soil and Grassland Systems Department, Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, U.K.
#Key Laboratory of Urban Environment and Health, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen 361021, People’s Republic of China

Supporting Information

ABSTRACT: Microbe-mediated arsenic (As) biotransformation in paddy soils determines the fate of As in soils and its availability to rice plants, yet little is known about the microbial communities involved in As biotransformation. Here, we revealed wide distribution, high diversity, and abundance of arsenite (As(III)) oxidase genes (aioA), respiratory arsenate (As(V)) reductase genes (arrA), As(V) reductase genes (arsC), and As(III) S-adenosylmethionine methyltransferase genes (arsM) in 13 paddy soils collected across Southern China. Sequences grouped with As biotransformation genes are mainly from rice rhizosphere bacteria, such as some Proteobacteria, Gemmatimonadales, and Firmicutes. A significant correlation of gene abundance between arsC and arsM suggests that the two genes coexist well in the microbial As resistance system. Redundancy analysis (RDA) indicated that soil pH, EC, total C, N, As, and Fe, C/N ratio, SO4−-S, NO3−-N, and NH4+-N were the key factors driving diverse microbial community compositions. This study for the first time provides an overall picture of microbial communities involved in As biotransformation in paddy soils, and considering the wide distribution of paddy fields in the world, it also provides insights into the critical role of paddy fields in the As biogeochemical cycle.

INTRODUCTION

China is the world’s largest rice producer, accounting for about 30% of the total world production (http://beta.irri.org/statistics), mostly in Southern China.1 As a highly toxic metalloid, arsenic (As) contamination in paddy fields has emerged as a serious health concern worldwide,2 especially considering that the anaerobic conditions in paddy soils are conducive to As mobilization,3,4 resulting in a markedly enhanced bioavailability of As to rice plants.5 It is now recognized that consumption of rice constitutes a large proportion of the dietary intake of As for the populations in China and other Asian countries.5–8 Microbes are the key drivers for As biotransformation in paddy soils, catalyzing arsenate (As(V)) reduction, arsenite (As(III)) oxidation and methylation.9 Microcosm studies have demonstrated that microbes capable of As reduction, oxidation, and methylation often coexist in paddy soils, and their relative abundance and activity determine the fate of As in the paddy environment and the bioavailability of As to plants.10–12 There are two known microbial pathways for As(V) reduction, the respiratory pathway mediated by arrA genes13 and the detoxification pathway mediated by arsC genes.14 As(III) oxidation is catalyzed by As(III) oxidases, which are encoded by aioA and aioB genes for the two subunits of the enzyme15 and have been identified in several heterotrophic and chemoautotrophic microorganisms.16 Moreover, As(III) can be methylated by microbes into various organic As species.17 Arsenic biomethylation is catalyzed by As(III) S-adenosylmethionine methyltransferase (ArsM), which is encoded by arsM genes.18

Received: January 3, 2015
Revised: March 2, 2015
Accepted: March 4, 2015
Published: March 4, 2015

DOI: 10.1021/acs.est.5b00028
Environ. Sci. Technol. 2015, 49, 4138−4146
Although As biotransformation mechanisms have been well studied in microbial pure cultures, the distribution, diversity and abundance of genes responsible for As metabolism in paddy soils have not been well characterized. Metagenomic approach or PCR-based methods have been used to investigate As-related genes in aquatic environments or highly As contaminated environments. In this study, PCR-based methods were used to study aioA, arrA, arsC, and arsM gene abundances and phylogenetic diversities in 13 paddy field samples collected across Southern China. The correlations between different gene abundances were analyzed to investigate the coexistence of As biotransformation genes in paddy soil bacterial genomes. Moreover, in order to obtain a better understanding of how paddy soil properties affect the abundance and diversity of microbial As genes, key factors driving diverse composition of microbial communities involved in As biotransformation in paddy fields were investigated.

**MATERIALS AND METHODS**

**Soil Sampling.** Soil samples were collected from 13 distinct paddy fields across Southern China, Anqing (AQ), Changshu (CS), Jiangmen 1 (JM 1), Jiangmen 2 (JM 2), Fuzhou (FZ), Guilin (GL), Guiyang (GY), Jiangzhou (JZ), Jiaxing (JX), Mianyang (MY), Yingtan (YT), Changde (CD), and Zhanjiang (ZJ). The 13 soil samples are from dominant rice production fields located in 11 provinces in China. Soil properties, including pH, EC, total C, total N, nitrate (NO$_3^-$-N), ammonium (NH$_4^+$-N), total C/N ration, total S, sulfate (SO$_4^{2-}$-S), total Fe, total As, and phosphate-extractable As concentrations were determined following standard methods of soil analysis. Details are shown in the Supporting Information.

**Nucleic Acid Extraction and PCR.** DNA was extracted from 0.5 g of samples using the FastDNA SPIN Kit for soil (MP Biomedicals) according to the manufacturer’s instructions. PCR amplifications of aioA, arrA, arsC, and arsM genes were performed with the primers AroAdeg1F/AroAdeg1R and AroAdeg2F/AroAdeg2R, ASIF/ASIR, amtl-42-f/amlt-376-r and smrc-42-f/smrc-376-r, and arsMF1/arsMR2, respectively. Details are shown in the Supporting Information.

**Quantitative PCR of aioA, arrA, arsC, and arsM Genes.** The abundance of aioA, arrA, arsC, and arsM genes from 13 samples was estimated using the primers described above by quantitative real-time polymerase chain reaction (qPCR) performed on an iQTM5 Thermocycler (Bio-Rad). Details are shown in the Supporting Information. Assay efficiencies were 95.5% for aioA genes, 102.6% for arrA genes, 90.3% for arsC genes, 100.7% for arsM genes, and 102.2% for 16S rRNA genes.

**Terminal Restriction Fragment Length Polymorphism (T-RFLP) Analysis.** PCR amplifications of aioA, arrA, arsC, and arsM genes were conducted according to the optimized methods with each of the forward primers labeled with 6-carboxyfluorescein (FAM). The labeled PCR products were gel-purified with the Wizard SV Gel and PCR Clean-Up System (Promega) and then digested by the restriction enzyme TaqI (Takara Bio Inc., Japan) at 65 °C for 4 h. Digestion products were purified and then determined by ABI PRISM 3130XL Genetic Analyzer (Applied Biosystems). Relative abundances of each individual terminal restriction fragments (T-RFs) were calculated based on its peak areas in relation to total peak areas. T-RFs that occurred in at least four replicates and with a percentage >2% were listed. These T-RFs were similar to in vivo digestion of the clone sequences in the phylogenetic tree.

**Cloning and Community Analysis.** According to the results from qPCR and T-RFLP analyses, seven samples (AQ, JM 2, FZ, MY, ZJ, GL, and CD) were chosen for cloning and community analysis. They were chosen based on the relative gene abundance and the different microbial community composition ranges which included both samples with the highest gene abundance or biodiversity and the lowest and also included samples with the moderate gene abundance and biodiversity for aioA, arrA, arsC, and arsM. Details are shown in the Supporting Information. The nucleotide sequences obtained in this study have been submitted into the NCBI GenBank databases under the following accession numbers: aioA genes from KP060099-KP060406, arsC genes from KP060407-KP060549, arsM genes from KP060550-KP060962, and arrA genes from KP060963-KP061152.

**Statistical Analysis.** Spearman’s rank-order correlation was used to test the correlations between the relative gene abundances; this nonparametric method was used because the relative gene copy numbers did not follow a normal distribution. Heatmaps and clustering analyses were generated using heatmap tool (heatmap.2) in the gplots package within the statistical program R 3.1.2. The relative abundances of the OTUs higher than 1% percentage (the percentage of each OTU to total OTUs) were used to generate heatmaps. Complete linkage clustering of seven samples (AQ, JM 2, FZ, MY, ZJ, GL, and CD) was calculated by the composition and relative abundance of aioA, arrA, arsC, and arsM genes. Redundancy analysis (RDA) of environmental variables and the absolute gene copy numbers or T-RFLP based microbial communities were chosen according to the result of detrended correspondence analysis (DCA). Environmental variables were forward selected based on the significance test performed by the Monte Carlo permutation test (P < 0.05). These calculations were performed in R 3.1.2 with the vegan package.

**RESULTS**

**Soil Properties.** Soil properties of the 13 paddy soils are listed in Table S1 in the Supporting Information. Total As concentration varied from 11.7 to 25.4 mg kg$^{-1}$ (Figure 1). Phosphate-extractable As (0.4–3.6 mg kg$^{-1}$) represented 20%–30% of the total As, and both inorganic As species (As(III) and As(V)) and methylated As species (DMA and MMA) were detected in most of the soil samples (Supporting Information, Table S2). As(V) was the dominant As species, accounting for 92–99% of the phosphate-extractable As concentration, because soils were air-dried before extraction.

**Identification and Quantification of the Gene Abundances of As Biotransformation.** The aioA, arrA, arsC, and arsM genes were detected in all samples. To minimize variances caused by different background bacterial abundances, extraction and analytical efficiencies, the absolute gene copy numbers of aioA, arrA, arsC, and arsM (Supporting Information, Figure S1) were normalized to that of ambient 16S rRNA genes. The relative gene abundances varied for aioA (0.2–6.1 × 10$^{-2}$), arrA (0.05–8.7 × 10$^{-3}$), arsC (0.02–1.1 × 10$^{-2}$), and arsM (0.1–2.6 × 10$^{-3}$) (Supporting Information, Figure S2). Among them, aioA genes were the most abundant, while arsM genes were the least abundant. arsM and arsC showed a strong linear relationship (R$^2 = 0.90$) (Figure 2). Spearman’s rank-order correlation (rho) analysis also showed that arsC gene abundances correlated strongly with arsM (rho =...
The correlations between \textit{arrA} and \textit{arsM} or \textit{arsC} gene abundances were also significant ($\rho = 0.604$ and $0.714$, $P < 0.05$ and 0.01, respectively), but these correlations were strongly influenced by the ZJ samples, which appear to be outliers, in fact there was no significant linear relationship between \textit{arrA} and \textit{arsM} or \textit{arsC} without ZJ samples ($R^2 = 0.15$ and 0.21, respectively) (Figure 2). The high \textit{arrA} gene abundances in ZJ samples might be contributed by the microbes represented by T-RFs 34 and 309 bp, which were the dominant microbial communities involved in As(V) respiratory reduction (Supporting Information, Figure S3). There were no significant correlations between \textit{aioA} genes and the other three genes.

**Biodiversity and Community Composition of \textit{aioA} Genes.** T-RFLP analysis showed the variation of microbial community compositions involved in As(III) oxidation in the 13 paddy soils. A total of 10 T-RFs (53, 64, 80, 88, 109, 130, 143, 220, 347, and 377 bp) were detected in \textit{aioA} gene T-RFLP profiles in soils (Supporting Information, Figure S3). The T-RF 80 bp was dominant in all 13 paddy soils, accounting for 29–75% of total T-RFs, followed by T-RFs 130, 88, and 220 bp. The higher percentage of T-RF 80 bp in JM 2, T-RF 64 bp in MY, T-RF 220 bp in GL, and T-RF 53 bp in YT and CD contributed to their different microbial communities from the other samples, which was also shown in RDA based on \textit{aioA} T-RFLP profiles (Figure 3). Heatmap of \textit{aioA} genes showed that the microbes responsible for As(III) oxidation represented by OTU 1 and 2 were the dominant genera in the 7 selected paddy soils. Cluster analysis based on \textit{aioA} gene sequences also showed that sample JM 2 separated from FZ, AQ and ZJ, and CD separated from MY and GL (Figure 4). Phylogenetic analysis of the 646 \textit{aioA} sequences from the seven selected samples allowed the identification of 308 unique OTUs based on a 97% cutoff (Supporting Information, Figure S5). The microbes were mainly \textit{α-Proteobacteria} (14%) and \textit{β-Proteobacteria} (37%), including \textit{Rhizobiales} and \textit{Burkholderiales}, together with three undefined clusters along with amplified \textit{aioA} genes from rice rhizosphere, rice root, aquatic sediment, and As contaminated soil.\textsuperscript{10,24,30}

![Figure 1. Location and As concentration in 13 paddy soils in Southern China and As concentration in soil samples. The base map used is from the National Fundamental Geographic Information System of China. The orange columns represent As concentration in soil samples.](image1)

![Figure 2. Relationships of relative gene abundances of \textit{aioA}, \textit{arrA}, \textit{arsC}, and \textit{arsM} from the 13 paddy soils in Southern China. Lines are fitted according to linear fitting in OriginPro 8 program. To minimize variances caused by different background bacterial abundances, extraction and analytical efficiencies, the absolute gene copy numbers of \textit{aioA}, \textit{arrA}, \textit{arsC}, and \textit{arsM} (Supporting Information, Figure S1) were normalized to that of ambient 16S rRNA genes. For the relationships between \textit{arrA} and \textit{arsM} or \textit{arsC} gene abundances, ZJ samples were shown as outliers in the dashed circle.](image2)
Biodiversity and Community Composition of arrA Genes. Seven T-RFs (34, 58, 117, 123, 192, 309, and 379 bp) were detected in arrA gene T-RFLP profiles (Supporting Information, Figure S3). T-RF 34 bp had the highest relative abundance (42−91%) in all the 13 paddy soils, followed by T-RFs 58 and 123 bp except the AQ and FZ samples, for which the dominant T-RFs were 123 (49%) and 379 bp (52%), respectively. In sample GL, the percentage of T-RF 34 bp was the highest, which contributed most to its difference from the other samples. The differences of microbial community compositions for sample AQ, FZ, and GL were also shown in RDA based on arrA T-RFLP profiles (Figure 3). Heatmap of arrA genes revealed that the microbes associated with respiratory As(V) reduction represented by OTU 1-5 were the dominant genera in the 7 selected paddy soils. Cluster analysis based on arrA sequences showed that sample AQ was the most different from the other paddy soils (Figure 4). The phylogenetic tree of arrA sequences was constructed with 190 defined OTUs (identity >97%) after analysis of 613 arrA clone sequences (Supporting Information, Figure S6). The arrA sequences could be grouped into seven clusters, and most of the sequences were aligned to the arrA genes of uncultured bacteria identified in paddy soil, sediments, and bioreactors. Some of the arrA sequences were related to Geobacter uranireducens.

Biodiversity and Community Composition of arsC Genes. Five T-RFs (38, 53, 92, 146, and 203 bp) were detected in arsC gene T-RFLP profiles (Supporting Information, Figure S3). T-RFs of 38, 53, and 146 bp were detected in all the 13 paddy fields, and 53 bp was the dominant T-RF, accounting for 37−66% of the total T-RFs. T-RF 203 and 92 bp were detected only in samples CS, GL, JZ and AQ, YT, CD, respectively. The higher percentage of T-RF 146 bp in samples JM 1, JM 2, and GL contributed to their distinct microbial community compositions to the other samples and coincided with the result of RDA based on arsC T-RFLP profiles (Figure 3). Heatmap of arsC genes showed the dominant genera involved in As(V) reduction in these paddy soils, and cluster analysis based on arsC sequences also revealed that samples GL and JM 2 separately from the other paddy soils (Figure 4). The phylogenetic tree of arsC sequences was constructed with 143 defined OTUs (identity >97%) after analysis of 615 arsC clone sequences (Supporting Information, Figure S7) and revealed a dominance of bacteria belonging to the Proteobacteria phylum (α-Proteobacteria and γ-Proteobacteria with 43% and 8%, respectively), including Hoeflea, Sinorhizobium, Mesorhizobium, Polymorphum, and Enterobacter, and also three unknown clusters, mainly microbes associated with the rice rhizosphere, rice root, and As contaminated soils.

Biodiversity and Community Composition of arsM Genes. According to the T-RFLP analysis of arsM genes, 10 T-RFs (32, 59, 70, 83, 137, 158, 203, 224, 268, and 312 bp) were detected in the soils (Supporting Information, Figure S3). T-RF 83 bp was dominant in samples AQ, CS, JM 2, JZ, JX, MY, and
ZJ, with the relative abundances ranging from 34% to 46%. In the other six soils (JM 1, FZ, GL, GY, YT, and CD), T-RF 59 bp showed the highest abundance, accounting for 34–60%. The appearance of T-RF 70 bp in samples GL, JM 1, and GY and the highest abundance of T-RF 59 bp in samples FZ, YT, and CD contributed to their different microbial community compositions from the other samples, which were also revealed by RDA based on arsM T-RFLP profiles (Figure 3). The dominant genera involved in As methylation in these paddy soils were revealed in heatmap of arsM genes, and samples GL and JM 2 were shown as the most distinct from the other paddy soils in cluster analysis based on arsM sequences (Figure 4).

Analysis of the 656 arsM sequences from the seven selected samples allowed the identification of 413 unique OTUs based on a 97% cutoff (Supporting Information, Figure S8). These sequences belonged to Gemmatimonadales (16%), Firmicutes (9%), Actinobacteria (11%), α-Proteobacteria (22%), β-Proteobacteria (6%), δ-Proteobacteria (6%), and Archaea (6%) and to a lesser extent to two unknown clusters containing sequences from rice rhizosphere soil and rice root.11

RDA of Environmental Factors and As Biotransformation Microbial Community Compositions and Gene Abundances. According to RDA (Figure 3), the environmental factors pH ($P = 0.001$), SO$_4^{2-}$-S ($P = 0.02$), and total C/N ratio ($P = 0.02$) significantly explained the variation in the microbial community compositions based on aioA genes, and the first and second axis accounted for 34.8% and 21.7% of the total variance, respectively. The pH value explained the most variation (32.1%) of microbial community compositions involved in As(III) oxidation, followed by SO$_4^{2-}$-S (21.8%) and total C/N ratio (21.1%). For arsM based microbial communities, total Fe ($P = 0.02$), NO$_3^-$-N ($P = 0.02$), pH ($P = 0.04$), total As ($P = 0.04$), and NH$_4^+$-N ($P = 0.03$) significantly contributed to the variant compositions of microbial communities, and the first and second axis accounted for 31.3% and 23.6% of the total variance, respectively. Total Fe explained the most variation (22.9%) and followed by NO$_3^-$-N.
(19.7%), pH (19.4%), total As (18.6%), and NH₄⁺-N (16.3%). For *arsC* based microbial communities, EC (P = 0.002), SO₄²⁻-S (P = 0.03), total As (P = 0.03) and Fe (P = 0.05) explained a significant amount of the variability in microbial community compositions, and the first and second axis accounted for 42.7% and 19.8% of the total variance, respectively. EC explained the most (41.4%) variation and followed by SO₄²⁻-S (29.3%), total As (24.1%), and Fe (17.8%). For microbial communities based on *aioA* genes, the environmental factors NH₄⁺-N (P = 0.01), total Fe (P = 0.03), and N (P = 0.04) significantly correlated with the variation of microbial community compositions, and the first and second axis accounted for 32.6% and 15.0% of the total variance, respectively. Among them, NH₄⁺-N explained the most variation of microbial community compositions associated with respiratory As(V) reduction (20.4%) and followed by total Fe (20.3%) and N (20.3%). For the absolute gene abundances of *aioA*, *aioB*, *aioC*, *arsM*, and 16S rRNA, total N (P = 0.02), C (P = 0.02), and As (P = 0.03) significantly explained the variance in 13 paddy soils, and the first and second axis accounted for 65.4% and 0.1% of the total variance separately. Among them, total N explained the most variation (43.3%) and was followed by C (37.4%) and As (22.5%).

**DISCUSSION**

**High Abundance of *aioA* Genes in Paddy Soils.** The abundance of *aioA* genes was almost 10 times higher than the other genes (Supporting Information, Figure S2). A higher diversity of *aioA* compared with *arrA* and *arsC* genes was also revealed by Shannon and Simpson diversity indices (Supporting Information, Table S4), indicating a high potential for microbial As oxidation in paddy soils. The major groups of As oxidizing bacteria, i.e., the *Rhizobiaceae* and *Burkholderiaceae* in α-Proteobacteria and β-Proteobacteria (Supporting Information, Figure S5), were typically rhizospheric microbes. These microbes have been shown to contribute significantly to the oxidation of As(III) in soils under both aerobic and anaerobic conditions. Previous studies have shown that the abundance of As(III)-oxidizing microbes could enhance As(III) oxidation to As(V) in the rice rhizosphere, resulting in decreased As mobility and bioavailability because As(V) can be sequestered on Fe/Mn hydroxide/oxyhydroxide in rhizosphere soil and on rice roots. For microbes, As(III) oxidation is also one of the As detoxification mechanisms, because the product of oxidation, As(V), is less toxic than As(III). In addition, compared with most of the other samples, the lower *aioA* gene abundances in samples ZJ, AQ, CD, and FZ, which had been drained to different extent at sampling, indicated that the decrease of mobile As(III) under nonflooded period resulted in the reduction of *aioA* gene abundances.

Both *arrA* and *arsC* Genes Were Identified in Paddy Soils. Genes involved in both respiratory As(V) reduction and detoxifying As(V) reduction were identified in the paddy soils. The abundance of *arrA* genes was much lower than that of *arsC* (Supporting Information, Figure S2), possibly because *arrA* genes present only in anaerobic microbes, which restricts its abundance in the paddy soils. Under flooded conditions, because of reductive dissolution of the Fe(hydr)oxide and reduction of As(V) to As(III), arsenic is quickly released from the soil to the pore water. The As(V) respiring bacteria could reduce As(V) under the anaerobic conditions to gain energy. *arrA* sequences closely related to *Geobacter uranireducens* species have been identified in this study, which has also been confirmed by Héry et al. that reported the *arrA* amplified from West Bengali sediments showed high amino acid sequence identity to sequences of putative *arrA* genes in genomes of *Geobacter uranireducens* and *G. lovleyi*. Moreover, *arrA* related to *Geobacter* species have been frequently detected in As rich sediments. Unlike *arrA*, the detoxification reductase genes *arsC* is present in both aerobic and anaerobic microbes and has been identified in environmental samples and rice rhizosphere. In this study, the microbes responsible for detoxification As(V) reduction belonged to some typically rhizospheric microbes, such as *Rhizobiales* and *Pseudomonadales* (Supporting Information, Figure S7) and could contribute to As(V) reduction and mobility in paddy soils under both aerobic and anaerobic conditions.

**High Diversity of *arsM* Genes in Paddy Soils.** Although *arsM* gene abundance was the lowest when compared with the other genes (Supporting Information, Figure S2), the diversity (estimated by Shannon and Simpson diversity indices) and the OTU richness (estimated by Chao and ACE) of *arsM* genes were much higher than *aioA*, *aioB*, and *arsC* genes at a 97/90% similarity cutoff in almost all the samples (Supporting Information, Table S4), suggesting a highly diverse microbial communities involved in As methylation in paddy fields. These diverse microbes could contribute to the presence of monomethylarsenate (MMAs(V)) and dimethylarsenate (DMAs(V)) detected in the soils (Supporting Information, Table S2). The representative *arsM*-containing genera in *Proteobacteria*, *Gemmatimonadales*, and *Firmicutes* (Supporting Information, Figure S8) have been identified as the major microbial community in the rice rhizosphere, and their significant contribution to As methylation has also been reported. Various volatile As species (mono-, di-, and trimethylarsonic) can be generated during the As methylation process and have been detected in the atmosphere above paddy fields. The wide distribution of As(III) methyltransferases in the 13 paddy soils underlines the potential for As methylation resulting in not only the accumulation of methylated As species in rice but also volatilization of As into the atmosphere. Previous studies have estimated that natural origin represented 62% of the volatile As released to the atmosphere and As biovolatilization accounted for 58% of natural As emissions. Mestrot et al. reported that As emitted from Bangladesh and Spanish paddy soils to the atmosphere was 23–30% for 4.5 t/year, respectively, and demonstrated the sizable contribution of paddy fields to As biovolatilization.

**Strong Correlation of *arsC* and *arsM* Gene Abundance.** It is interesting to note that *arsC* gene abundance correlated strongly with *arsM* in the 13 paddy soils (Figure 2). *arsC* gene encodes the cytoplasmic ArsC reductase and catalyzes the cellular reduction of As(V) to As(III). The toxic As(III) can either be subsequently extruded by the efflux system or methylated to less toxic methylated As species (MMAs(V), DMAs(V), and volatile TMAs(III)) by ArsM and thus to complete the detoxification processes. Studies of microbes have shown that *arsM* gene is often in the gene cluster adjacent to other genes encoding As-resistance proteins. For example, *Rhodopseudomonas palustris* CGA009 or *Synechocystis* sp. PCC6803 both contain genes encoding ArsM, arsenical compound resistance 3 (ACR3), and *arsC* for As resistance. Therefore, the strong correlation between *arsC* and *arsM* gene abundances strongly implies that *arsM* gene coexists well with the *arsC* gene in the microbial As resistance system, especially for microbes in paddy fields. Moreover, there was an
indication of a similar clustering of soils from the 7 different geographical regions for microbes involved in As(V) reduction and As methylation based on arsC and arsM gene diversities (Figure 4). This result suggests that similar compositions of microbial communities are involved in As(V) reduction and As methylation in paddy soils.

Soil Properties Explaining Variations in Gene Abundances and Microbial Community Compositions Involved in As Biotransformation. Among the soil properties tested (Figure 3), total As concentration was found to be positively correlated with the abundances of aioA, arrA, arrC, and arsM genes. In contrast, 16S rRNA gene abundance had a negative correlation with As concentration but a positive correlation with total C and N concentrations, implying that As in the paddy fields possibly exerted a selection pressure on soil microbes resulting in decreased bacterial abundance. However, the presence of As in soils might result in higher abundance of microbes involved in As biotransformation, as Lami et al. reported that when amended with 200 μmol kg⁻¹ As(III), the total number of bacterial taxa decreased by 2-fold but some of the As(III) oxidizing bacterial groups increased up to 20-fold.47 Soil pH had a significant effect on the composition of microbial communities associated with both As(III) oxidation and As methylation, probably due to the narrow pH ranges for optimal growth of bacteria.48 The concentration of SO₄²⁻–S also explained 21.8% and 29.3% variation in the structure of microbial communities responsible for As(III) oxidation and detoxification As(V) reduction, respectively. A previous study with metaproteomic method showed the role of several dominant bacteria in highly As contaminated acid mine drainage recycling of both mineral and organic resources, such as As, Fe, S, urea, vitamins, nucleosides, and amino acids.27 In this case, microbes responsible for As(III) oxidation and As(V) reduction might share some microbial groups with that capable of sulfur oxidation, and it is reasonable to expect that SO₄²⁻–S concentration may have a significant influence on these microbial community composition. Total Fe concentration was another significant environmental factor responsible for the variation in the composition of microbial communities involved in respiratory As(V) reduction, detoxification As(V) reduction and As methylation in paddy soils. Many studies have demonstrated that iron oxhydroxides could sequester As(V), and the reduction of iron oxhydroxides could increase As bioavailability in paddy soils.49,50 Therefore, it could be Fe oxides that contribute to microbial community diversity. For microbes with arsC and arsM genes, the concentrations of both Fe and As were the dominant environmental factors driving the variation in bacterial community composition. While the concentration of total N, NH₄⁺–N, and total Fe were the dominant environmental factors responsible for diverse compositions of microbial communities involved in respiratory As(V) reduction. This could be explained as a previous study showed microbially mediated redox processes in paddy soils are strongly related to the microbial accessibility of C, N, and Fe.1 Moreover, the concentrations of NH₄⁺–N and NO₃⁻–N were also potential drivers of microbial community compositions involved in As methylation as shown in RDA in this study. In addition, the EC value was another significant environmental factor affecting microbial community structures, as it contributed most to the variation in the composition of microbial communities involved in As(V) reduction. It is possible that the effect of EC is related to its significant correlation with As concentration in the 13 soils (rho = 0.709, P < 0.01, Supporting Information, Table S5).

Implications. This study shows the ubiquitous distribution of microbes associated with As redox changes and methylation in paddy soils. These microbes are likely to play a key role in driving the biogeochemical cycle of As in the paddy environment and affecting As accumulation by rice plants.10–12 The distribution patterns of the dominant genera involved in As biotransformation and the cluster analysis based on aioA, arrA, arrC, and arsM diversities suggest a clustering of soils from different geographical regions, i.e., the separation of southeast samples (AQ, FZ, JM 2, and ZJ) from the southwest samples (MY, CD, and GL) based on aioA and arrA genes, and the separation of samples GL, ZJ, and JM 2 which locate at the most southern part in China from the other samples based on arsC and arsM genes (Figure 1), implying the different microbe-mediated As biotransformation abilities in variant paddy soils based on geographic region and might result in different As species and concentrations accumulated by rice plants. The separations of different geographical regions might result from different soil properties as revealed by RDA analysis based on aioA, arrA, arrC, and arsM T-RFLP profiles, which indicated the similar separation patterns for these samples (Figure 3). For aioA and arrA genes, the lower soil pH value and total C/N ratio in samples MY and GL, and the high concentration of total N in samples MY, GL, and CD contributed the most to the separation patterns. While for arsC and arsM genes, the different separation patterns might be due to the higher As concentrations in samples GL, ZJ, and JM 2. The environmental factors, such as pH, total C, N, and Fe, appear to drive the variation in gene abundance and microbial community associated with As biotransformation. These factors have also been implicated in the As biogeochemical cycle in paddy soils.5,14 Manipulating these environmental factors through cultivating rice aerobically or minimizing organic matter amendments in paddy soil can effectively reduce As accumulation in rice grain.5,46 Given that paddy soils make up the largest anthropogenic wetlands on earth, which are highly modified by anthropogenic activities,5 the identified phylogenetic and ecological diversity of microbes involved in As redox changes and methylation in paddy soils underlines their critical role in the global As biogeochemical cycle. Future in situ studies should be carried out to investigate how paddy conditions impact on As biotransformation gene abundance and microbial community composition by comparing adjacent paddy and nonpaddy soils.

ASSOCIATED CONTENT

Supporting Information

Gene abundances; community structure of aioA, arrA, arrC, and arsM genes; rarefaction curve of clone libraries; phylogeny of aioA, arrA, arrC, and arsM sequences; soil location and properties; primers sequences and PCR thermal cycling parameters; diversity and richness estimators (Shannon, Simpson, ACE and Chao1); Spearman’s rank-order correlation of soil properties. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*yPhone: 86-592-6190997; fax: 86-592-6190997; e-mail: ygzhu@rcees.ac.cn.

DOI: 10.1021/acs.est.5b00028
Environ. Sci. Technol. 2015, 49, 4138–4146
ACKNOWLEDGMENTS

This study is financially supported by Natural Science Foundation of China (Grant No. 41330553). We would like to thank Shu-Wei Wang in the Institute of Soil Science, Chinese Academy of Science for helping collecting paddy soil samples. We also thank Xin-Li An and Hui-Juan Xu in the Institute of Urban Environment, Chinese Academy of Sciences for helping with the DNA extraction and providing the National Fundamental Geographic Information System of China.

REFERENCES


