

Response of Soil Microbial Communities to Elevated Antimony and Arsenic Contamination Indicates the Relationship between the Innate Microbiota and Contaminant Fractions

Weimin Sun,^{†,¶} Enzong Xiao,^{‡,§,||,¶} Tangfu Xiao,^{*,‡,§}[®] Valdis Krumins,[⊥] Qi Wang,[†] Max Häggblom,[#] Yiran Dong,^V[®] Song Tang,^O[®] Min Hu,[†] Baoqin Li,[†] Bingqing Xia,[†] and Wei Liu

[†]Guangdong Key Laboratory of Agricultural Environment Pollution Integrated Control, Guangdong Institute of Eco-Environmental Science & Technology, Guangzhou 510650, China

[‡]Key Laboratory of Water Quality and Conservation in the Pearl River Delta, Ministry of Education, School of Environmental Science and Engineering, Guangzhou University, Guangzhou 510006, China

[§]State Key Laboratory of Environmental Geochemistry, Chinese Academy of Sciences, Guiyang 550081, China

^{II}University of Chinese Academy of Sciences, Beijing 100049, China

¹Department of Environmental Sciences, Rutgers University, New Brunswick, New Jersey 08901, United States

[#]Department of Biochemistry and Microbiology, Rutgers University, New Brunswick, New Jersey 08901, United States

 $^
abla$ Institute for Genomic Biology, University of Illinois, Urbana-Champaign, Urbana, Illinois 61801, United States

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

•Water Resources Protection Bureau of Pearl River Water Resources Commission, Guangzhou 510611, China

Supporting Information

ABSTRACT: Mining of sulfide ore deposits containing metalloids, such as antimony and arsenic, has introduced serious soil contamination around the world, posing severe threats to food safety and human health. Hence, it is important to understand the behavior and composition of the microbial communities that control the mobilization or sequestration of these metal(loid)s. Here, we selected two sites in Southwest China with different levels of Sb and As contamination to study interactions among various Sb and As fractions and the soil microbiota, with a focus on the microbial response to metalloid contamination. Comprehensive geochemical analyses and 16S rRNA gene amplicon sequencing demonstrated distinct soil taxonomic inventories depending on Sb and As



contamination levels. Stochastic gradient boosting indicated that citric acid extractable Sb(V) and As(V) contributed 5% and 15%, respectively, to influencing the community diversity. Random forest predicted that low concentrations of Sb(V) and As(V)could enhance the community diversity but generally, the Sb and As contamination impairs microbial diversity. Co-occurrence network analysis indicated a strong correlation between the indigenous microbial communities and various Sb and As fractions. A number of taxa were identified as core genera due to their elevated abundances and positive correlation with contaminant fractions (total Sb and As concentrations, bioavailable Sb and As extractable fractions, and Sb and As redox species). Shotgun metagenomics indicated that Sb and As biogeochemical redox reactions may exist in contaminated soils. All these observations suggest the potential for bioremediation of Sb- and As-contaminated soils.

INTRODUCTION

Antimony (Sb) and arsenic (As) are two toxic metalloids belonging to group 15 of the Periodic Table. Both are suspected carcinogens and U.S. Environmental Protection Agency priority pollutants.^{1,2} Sb is widely distributed in the lithosphere and frequently associated with As in sulfide-rich ores.³ Thus, cocontamination of Sb and As is commonly observed in mining areas.⁴ The environmental toxicities of Sb and As are strongly dependent on their redox species.⁵ For both, their reduced forms, that is, Sb(III) and As(III), are more toxic than their oxidized counterparts, (Sb(V) and As(V)).^{6–8} In general, background concentrations of Sb in soil are low (<1 $mg \cdot kg^{-1}$).⁵ However, elevated concentrations of Sb in both soil

January 16, 2017 Received: Revised: June 30, 2017 Accepted: July 12, 2017 Published: July 12, 2017

and water have been found due to anthropogenic enrichment or contamination.³ Mining and smelting activities are the major sources of Sb contamination.⁹ China accounts for approximately 80% of the world's Sb production and thus large quantities of Sb have been released in China,¹⁰ causing contamination of surrounding environments.¹¹ Elevated concentrations of Sb greatly exceeding the background soil concentrations (1 mg·kg⁻¹) have been observed at numerous mining sites in China.¹¹

Despite the toxicity of Sb and As, microbially mediated Sb and As redox reactions have been observed. $^{12-15}$ Such reactions can enable precipitation, detoxification, and immobilization of Sb and As compounds, leading to promising strategies for bioremediation. There are many reports of As redox reactions in As-rich environments,^{16–18} but Sb biotransformations in natural habitats have received much less attention. Several microorganisms and enrichment cultures are able to oxidize or reduce Sb.¹⁹⁻²¹ However, our current understanding of Sb biotransformation is based on limited studies of bacterial isolates in culture^{19,20} or studies of the process in sediment microcosms.²¹ Fewer studies have previously investigated variations in indigenous microbial communities from Sb- and As- contaminated habitats. Therefore, an understanding of the indigenous microbial communities in Sb- and As-contaminated habitats is needed.

The extensive mining and smelting activities at the Banpo Sb mine, Southwest China, have made this area one of the most severely Sb-contaminated areas in the world. Elevated concentrations of Sb and contamination of both Sb and As have been observed in surface waters, aquatic sediments, and tailing dumps with effects on their bacterial communities. Therefore, these sites provided excellent natural laboratories to study the microbial response to different concentrations of Sb and As. However, the innate terrestrial soil bacterial microbiota of these sites has not yet been examined. In this study, we obtained soil samples from two contaminated sites with various levels of contamination. The aims were to (i) examine the impact of Sb and As contamination on the soil microbial community composition and diversity, (ii) elucidate the correlation between soil microbial assemblages with environmental factors, especially those directly linked to Sb and As contamination, and (iii) explore the potential metabolic pathways of the soil microbial ecosystems.

MATERIALS AND METHODS

Site Location and Soil Sampling. The study sites are located in eastern Dushan County, Guizhou Province, Southwest China (Supporting Information (SI) Figure S1). One sampling site (SM) is located within a smelting plant that has processed Sb-bearing sulfide minerals (Sb₂S₃) from an adjacent mine since the 1960s. Another sampling site (AS) is located in the woods 800 m from the smelting plant, which is less affected by the smelting activities and therefore represented a lesscontaminated site than SM. Fourteen sandy soil samples were sampled within a 50 m radius of the smelting plant, and 12 loamy soil samples were collected from site AS, which has an area of ~150 m². All soil samples were obtained from the surface (0–15 cm) using a steel push corer. For each sampling point, a total of 3 cores of ~30 g were collected and mixed to obtain a soil composite.

Measurement of Sb and As Associations/Fractions. The citric acid extractable Sb and As redox species (Sb(III)-C and As(III)-C, Sb(V)-C and As(V)-C) measurement was

described previously.^{26,27} Briefly, 0.2 g of soil was mixed with 10 mL of 100 mM citric acid (pH 2.08) for 1h, followed by 4 h of equilibration at room temperature under a N₂ atmosphere in a glovebox to prevent oxidation. The supernatant was filtered through a 0.45 μ m membrane after centrifuging at 3500 rpm for 30 min, and analyzed to determine the Sb(III) and As(III) concentrations using HG-AFS (AFS-920, Jitian, Beijing). The supernatant was also used to determine the total citric acid extractable Sb and As as described previously.^{22,28} The concentrations of Sb(V) and As(V) were determined from the difference between the total Sb and As and Sb(III) and As(III), respectively. A modified five-stage sequential extraction scheme (see Table S1 for extractants) for measuring various Sb and As phases was adapted from previous methodologies proposed by Wenzel et al.²⁹ and Gault et al.³⁰ For detailed information, please refer to the Supporting Information.

Illumina MiSeq Sequencing of 16S rRNA Genes. Total genomic DNA was extracted from 0.5 g soil using the FastDNA spin kit (MP bio, Santa Ana, CA) according to the manufacturer's protocol. 16S rRNA genes were amplified using primer pair 515*f*/907R, targeting the V4–V5 hyper variable regions.^{22,31}16S rRNA gene amplicon sequencing was performed on the Illumina MiSeq platform at Novogene (Beijing, China). FLASH was used to merge paired-end reads³² and QIIME (Quantitative Insights Into Microbial Ecology) (V1.7.0) was used to filter the raw reads and analyze these reads.³³ Chimeric sequences were removed by comparing with the Gold database using UCHIME.³⁴ Operational taxonomic units (OTUs) (97% similarity) were clustered by UPARSE. The phylogenetic taxonomy were assigned according to the RDP classifier (Version 2.2) and the Green Genes Database.^{35,36}

Sequencing of Metagenomics DNA Libraries. Shotgun metagenomic libraries from five soil samples were constructed on an Illumina Hiseq 4000 platform (paired-end 150 bp reads) at Novogene. A total of 42.2 Gb raw sequence data were generated (281,353,938 raw reads). These data were filtered (PRINSEQ 0.20.4) and quality control was performed (Trimmomatic 0.36) for subsequent statistical analysis and gene function prediction. Additional detail on the quality control filtering is provided in the SI. Reads were taxonomically and functionally annotated by similarity searching against NCBI-nr\KEGG database analyzed by Megan 6.³⁷ Raw sequences have been made available in the NCBI Sequence Read Archive (Accession # PRJNA354810 (16S rRNA) and PRJNA384316 (metagenomics)).

Numerical Analyses. We used Random Forest (RF) and Stochastic Gradient Boosting (SGB) analyses to estimate the contributions of environmental factors and quantify the strengths of connections. RF is an algorithm for classification³⁸ that uses an ensemble of trees. Each tree is growing while training using both bagging (bootstrap aggregation) and random variable selection. $^{39-41}~\rm RF$ does not require a preselection of variables because it has good predictive performance even when most variables are noisy.42 RF was used to predict the contributions of the individual environmental factors on the microbial diversity (i.e., Shannon index, the calculation for Shannon index can be found in the SI). The contributions were calculated using environmental factors as predictor variables and the diversity index as the dependent variable in an unbiased recursive partitioning-based RF model. RF was implemented in the R "randomForest" package.⁴¹ Model validation was performed using a 10-fold cross validation

Sb(III)-C Sbtot Sbsrp Sb(V)-C Sbexe 8000 1250 3000 100000 1000 750 6000 75000 2000 750 500 4000 -50000 500 1000 25000 250 2000 250 0 mg/kg SM AS sм AS sм AS sм AS sм AS 븑 ѕм Assrp As(III)-C As(V)-C Astot Asexe 🕇 AS 150 : 60 200 -600 15 100 150 40 400 10 100 -50 20 50 200 0 siv sм AS s'n AS sм AS siv AS AS

Figure 1. Boxplots showing the distribution of contaminant fractions in the two studied sites, SM and AS in Dushan County, Guizhou Province, Southwest China. Please see SI Table S1 for extractants.

method. In each run, we tested the agreement between measurements and predictions by calculating the average area under the curve using the receiver operating characteristic (ROC) approach.⁴³

SGB has become a powerful method in predictive data mining.⁴⁴ It is related to both boosting and bagging,^{45,46} and provides an accurate model by minimizing the gradient descent of the loss function in iterative tree construction.⁴⁵ Many small regression trees are built sequentially from the gradient of the loss function of the previous tree. A tree is constructed from a random subsample of the data set at each iteration (selected without replacement), producing an incremental improvement in the model.⁴⁷ SGB incorporates randomness in the training data subset as an integral part of the gradient boosting procedure. The "shrinkage" parameter $0 < v \leq 1$ controls the learning rate of the procedure.⁴⁵ These calculations were performed in the R "dismo" package.

A co-occurrence network was employed to visualize the correlation between the abundant taxa (top 70 most abundant OTUs) and contaminant fractions. A connection indicates a strong (|r| > 0.6) and significant (p < 0.05) Spearman's correlation. The size of each node is proportional to the number of connections; the thickness of each connection (i.e., edge) between two nodes is proportional to the absolute value of the Spearman's correlation coefficient, ranging from 0.6 to 1. The co-occurrence networks were visualized using the interactive platform Gephi.^{48,49} A number of parameters (e.g., number of nodes and edges, average path length, clustering and modularity coefficient) were calculated using R "igraph" package.⁵⁰ Detailed information for other statistical methods such as PCoA and UPGMA is provided in the SI.

RESULTS

Distribution of Sb and As Contamination in Two Sampling Sites. All soil samples from the two sampling sites, SM and AS had elevated concentrations in total Sb (Sb_{tot}) and As (As_{tot}) (Tables S2 and S3), ranging from 21 to 330 000 times the previously reported background soil Sb concentrations in China.¹¹ Sb_{tot} was significantly different at the two sites (p < 0.05, hereafter) (Figure 1). The concentrations of Sb_{tot} in SM averaged 30 500 \pm 34 269 mg·kg⁻¹, whereas those from site AS averaged 92.5 \pm 170.5 mg·kg⁻¹. Some samples from SM exhibited extremely high Sb_{tot} concentrations (>50 000 mg·kg⁻¹), with the maximum Sb_{tot} exceeding 100 000 mg·kg⁻¹. These extremely high Sb concentrations may indicate direct spills of product at the smelting site. There was a significant difference in As_{tot} between SM and AS as well: As_{tot} in AS averaged 122 \pm 43 mg·kg⁻¹ while As_{tot} in SM was 288 \pm 180 mg·kg⁻¹ (Figure 1 and SI Tables S2 and S3).

Additional environmental parameters for the contaminated site are shown in (SI Figure S2, Tables S4 and S5). The pH was significantly higher in SM (6.9 ± 0.9) than AS (5.1 ± 0.5) (p < 0.05). Nitrate was significantly higher in AS, whereas C/N was significantly higher in SM. No significant difference in sulfate concentrations was observed between sites, but SM05-SM08 exhibited much higher sulfate concentrations than other samples.

Composition of Contamination Fractions. Among the five extractable Sb and As fractions, the easily exchangeable fraction (M_{exe} , M stands for either Sb or As, hereafter) and the specifically sorbed surface-bound fraction $(M_{\rm srp})$, which are considered bioaccessible, ^{51,52} together accounted for less than 13% of Sb_{tot} or As_{tot} in all samples, with higher percentages seen in the more contaminated site (SM) than in the moderately contaminated site (AS) (p < 0.05) (Figure 1). The other three extractable fractions, which are considered nonbioaccessible, 51,52 accounted for the majority of Sb_{tot} and As_{tot} (SI Table S6). There were significant positive correlations between various Sb and As fractions(p < 0.05) such as As_{tot} and Sb_{tot} (r= 0.679, p = 0.002), As_{exe} and Sb_{exe} (r = 0.943, p < 0.001), and As_{srp} and Sb_{srp} (r = 0.913, p < 0.001). This confirms previous observations of the co-occurrence of Sb and As contamination in mining areas.^{22,24} Citric acid extractable Sb (Sb(III)-C and Sb(V)-C) and As (As(III)-C and As(V)-C) redox fractions only accounted for a small portion of $M_{\rm tot}$. These Sb and As redox fractions were previously regarded as bioaccessible fractions.⁵³ In the current study, we selected the bioaccessible Sb and As fractions including Mexe, Msrp, M(III)-C, M(V)-C, and M_{tot} as the contaminant fractions to explore Sb/As-microbe interactions.



Figure 2. Relative influence (%) of environmental variables on the Shannon index as evaluated by the SGB model. M(III)-C and M(V)-C represents the citric acid extractable M(III) and M(V) (M stands for Sb and As).

Microbial Community Across Two Sampling Sites. A total of 1 450 536 valid reads (average read length 373 bp) were used for subsequent downstream analyses. The valid reads clustered into 6474 OTUs, but many of these OTUs could not be identified at lower taxonomic levels (e.g., family and genus). Phylum-level analysis indicated that nine bacterial phyla made up the vast majority of total 16S rRNA sequences, accounting for 95% of total reads (SI Figure S3). Three hundred and 80 distinct genera were found among the 26 soil samples. Achromobacter (12%), Arthrobacter (3.9%), and Methylobacterium (2.9%) were the top 3 most abundant genera (SI Figure S4). Microbial communities differed by site, as shown by the UPGMA tree and PCoA (SI Figure S5), suggesting that innate environmental parameters may shape the indigenous microbial communities. t tests further demonstrated such differences among individual phylotypes. For example, phyla such as Acidobacteria, Chloroflexi, Gemmatimonadetes, and Verrucomicrobia were significantly more abundant in the less contaminated site, whereas Actinobacteria and Cyanobacteria were significantly enriched in the more contaminated site (SI Figure S6). At the genus level, Anthrobacter, Janthinobacterium, and Pseudomonas, demonstrated significant higher relative abundances in SM than AS (Figure S6).

Correlation between Environmental Factors and Microbial Communities. We hypothesized that the observed differences in the microbial communities at the two sites may be shaped by innate environmental parameters. For this purpose, SGB and RF models were used to interpret the relative importance of environmental variables to the community diversity as expressed by Shannon index. As(V)-C, nitrate, C/N, total N, and Sb(V)-C were among the top five factors influencing the Shannon index by SGB (Figure 2). In contrast, SGB revealed weaker effects for other contaminant fractions, such as As_{exe} and Sb_{srp} . Partial dependency plots of As(V)-C, Sb(V)-C, Sb_{exe} , and Sb_{srp} demonstrated that the increase of their concentrations from 0 to certain threshold values (~5 mg·kg⁻¹ for Sb(V)-C, ~0.3 mg·kg⁻¹ for As(V)-C and ~120 mg·kg⁻¹ for Sb_{exe} and Sb_{srp}) were likely to increase the Shannon index (Figure 3). However, concentrations above threshold did not change the diversity index as indicated by the horizontal line.

The co-occurrence network further demonstrated correlations between individual phylotypes and contaminant fractions (Figure 4). This network consists of 80 nodes including 70 most abundant OTUs (accounting for >50% of total valid reads) and contaminant fractions. A modular network was obtained with two major modules consisting of most nodes. Interestingly, all contaminant fractions and 13 OTUs cooccurred in module II. Nine OTUs clustered in module II could be taxonomically identified at genus level. Among these identified genera, many contain members relating to Sb and As redox reactions such as *Arthrobacter, Janthinobacterium, Thiobacillus,* and *Pseudomonas*.

Shotgun Metagenomics Analysis. Five shotgun metagenomics libraries were constructed from relatively lightly contaminated samples (AS07 and AS10) to extremely contaminated samples (SM06). In order to elucidate the metabolic potential of the microbial community with a focus on those relating to As metabolism, a total of 242 475 010 highquality reads were generated after filtering the five libraries. All assembled sequences in the five metagenome libraries were searched against the KEGG database. Gene families belonging to the Metabolism category within KEGG were dominant in the five libraries (SI Figure S7). More importantly, genes encoding for a number of As-related enzymes were identified (Figure 5). Genes related to the arsenic-resistance regulatory protein (arsR) and arsenate mycothiol transferase (arsC) were relatively more abundant than other As-related genes. Sequences for arsenite oxidase (*aoxA* and *aoxB*) were classified mostly to the Alphaproteobacteria, while the majority (>50%) of arsC sequences were attributed to Actinobacteria (Figure 5B). Other As-related genes were frequently classified to the Alphaproteobacteria and Actinobacteria as well. A number of As-related genes were positively correlated with many contaminant fractions, suggesting an adjustment of the microbial community in response to contamination (Figure 6A).

DISCUSSION

Our aims were to characterize the microbial community structure at Sb- and As-contaminated soils and to investigate the interactions between the microorganisms and the geochemical profiles, especially the contaminant fractions. This study was motivated by observations of dynamic microbial communities in various habitats from our previous studies.^{22,24,25} Here, we performed an in-depth analysis with the



Figure 3. Partial dependence plots for the RF model of the Shannon index impacted by members of contaminant fractions (the blue line is the partial dependence data line, the green line is the local polynomial regression fitting trend line, and the red band represents the 95% confidence interval). The partial plots reveal the dependencies of Shannon index on contaminant fractions. M(III)-C and M(V)-C represents the citric acid extractable M(III) and M(V) (M stands for Sb and As).

application of ensemble models (i.e., SGB and RF) to quantitatively investigate such interactions. In addition, shotgun

metagenomics was used to decipher the metabolic potentials of the soil microbial communities and the microbial response to



Figure 4. Co-occurrence network showing the correlation between bacterial taxa (OTUs) and contaminant fractions as nodes colored by modularity class. Numbers in the brackets indicate the OTU ID. Edges are only shown for strong (|r| > 0.6) and significant (P < 0.05) Spearman correlations. The size of each node is proportional to the number of connections (i.e., degree); the thickness of each connection between two nodes (i.e., edge) is proportional to the absolute value of Spearman's correlation coefficients, ranging from 0.6 to 1. M(III)-C and M(V)-C represent the citric acid extractable M(III) and M(V) (where M stands for Sb or As).

Sb and As contamination. The presence of various As-related genes and their positive correlations with contaminant fractions indicates the potential for biotransformation and resistance of As, and perhaps Sb, by the soil microbial communities.

Contaminant Fractions Affect Microbial Diversity. Elevated Sb and As concentrations were detected in all samples. Such co-occurrence patterns have been reported previously^{9,23,24} and may boost in situ Sb/As-microbe interactions. SGB indicated that As(V)-C and Sb(V)-C were important contributors influencing bacterial diversity. RF further predicted that low concentrations of M(V)-C had positive correlations with diversity. However, this effect was not obvious at higher concentrations of M(V)-C. Six of the ten measured contaminant fractions demonstrated negative effects on microbial diversity, suggesting that the contamination generally impairs the microbial diversity (Figure 3). The RF model predicted thresholds of various contaminant fractions (e.g., ~5 mg·kg⁻¹ for Sb(V)-C, ~ 0.3 mg·kg⁻¹ for As(V)-C) for microbial diversity. Above these thresholds the contaminant fraction had no additional effect on diversity. These predicted threshold values should be validated in future microcosm or mesocosm experiments with a focus on the concentrations below the horizontal part of the curve.

Co-Occurrence Patterns Suggest Sb/As-Microbe Interactions. The impact of contaminant fractions on diversity suggested Sb/As-microbe interactions. Hence, a further investigation of the effect of contaminant fractions on microbial communities and the biotic interactions was evaluated by cooccurrence network analysis (Figure 4). We found many bacterial OTUs that were positively correlated with the contaminant fractions, suggesting that contaminant fractions may enrich some specific taxa. Further, a modular structure was evident from the co-occurrence network. Each module contained highly interconnected nodes that are less connected to nodes in other modules and thus nodes within the module may share similar functions.^{49,54,55} We defined the most densely

Article



Figure 5. Percentage of reads annotated to Arsenic-related genes for five shotgun metagenomics sequencing libraries (A). Taxonomy of reads annotated to As-related genes for five shotgun metagenomics sequencing libraries (B). ARSC1, *arsC*; arsenate reductase; ASNA1, arsenite-transporting ATPase; ACR3, arsenite transporter, ACR3 family; ARSC2, *arsC*; arsenate reductase; *arsB*, arsenical pump membrane protein; AS3MT, arsenite methyltransferase; *aoxA*, arsenite oxidase small subunit; *aoxB*, arsenite oxidase large subunit; *arsH*, arsenical resistance protein ArsH; ACR2, arsenical-resistance protein 2; *arsC*, arsenate-mycothiol transferase; ArsR, family transcriptional regulator; ACR3, arsenite transporter, ACR3 family; ASNA1, arsenite-transporting ATPase.



Figure 6. Heatmap of Spearman's rank correlations coefficients between the contaminant fractions and (A) the annotated As-related genes as indicated in the bracket; (B) abundant genera (Red genera were identified as "core" genera). M(III)-C and M(V)-C represents the citric acid extractable M(III) and M(V) (M stands for Sb or As). The correlation coefficients are indicated by hue. *, p < 0.05;**, p < 0.01.

Article

connected node in each module as a "hub", which we suggest may be used as an indicator to predict the potential behavior of other nodes in the same module. 56,57 Sb_{srp} was identified as the "hub" in module II, into which all the other contaminant fractions fell as well. These observations suggested that OTUs that co-occurred in module II may be associated with Sb and As redox reactions or exhibiting some resistance to these metalloids. Indeed, members of Pseudomonas and Thermomonas have been identified as As(III)-transforming bacteria.^{58,59} Planomicrobium have been detected in in shallow aquifers with high arsenic levels.⁶⁰ Consistently, Planomicrobium, Pseudomonas, Janthinobacterium, and Thermomonas exhibited positive correlations with many members of the contaminant fractions (Figure 6B). All these indications collectively suggested the effectiveness of the network analysis in identifying microbial responses to contamination.

Microbial Response to Contamination at the Genus Level. The severe Sb and As contamination may form the selection pressure to enrich those microbes adapted to the harsh environments. Therefore, exploration of the abundant taxa holds the potential to discover putative metabolic potentials.⁶¹ Finer detailed analysis at lower phylogenetic levels, such as genus, may provide better phylogenetic resolution potential for bioremediation.⁶² We selected five genera as "core" genera because of their high relative abundances across all the samples and significant positive correlations with most of the contaminant fractions (Figure 6B). Interestingly, all of the selected core genera have members with reported capabilities in As redox transformations and Sb resistance. Moreover, OTUs relating to four of the core genera, with the exception of Achromobacter, were classified in module II of the co-occurrence network, into which all contaminant fractions fell. In addition, three of these core genera were significantly enriched at the more contaminated site (SI Figure S6).

The correlation of these "core" genera with multiple contaminant fractions suggests their selection by Sb/As contamination. For example, Achromobacter, which is the most abundant genus across all samples, contains a species (i.e., Achromobacter arsenitoxydans) that possesses a genomic "arsenic island" including three As resistance operons and an As(III) oxidation operon.⁶³ The Achromobacter-related 16S rRNA sequences demonstrated 99% sequence similarity with A. arsenitoxydans SY8 over 373 bp. Another core genus, Janthinobacterium, has been detected in As- and Sb-rich mine drainage tailings⁴ and an adjacent Sb-contaminated river,² though its exact ecological role, specifically with respect to Sb, is unknown. Many species of Arthrobacter have been identified as As(III)-oxidizing bacteria.⁶⁴ Another core genus, Pseudomonas, is notable for its capability of degrading a number of hydrocarbons.^{65–67} However, Pseudomonas also contains many species that are able to oxidize As(III), such as Pseudomonas arsenitoxidans⁶⁸ and Pseudomonas putida.⁵⁹ Besides the five core genera, Mycoplana, Thermomonas, and Planomicrobium demonstrated positive correlations with many contaminant fractions. Members of these genera have seldom been correlated with As and Sb redox reactions, suggesting the need for additional study. Although these core genera have been associated with As redox reactions, direct evidence of Sb resistance or Sb metabolism is scarce due to the fact that relatively few microbial species have actually been examined for the ability to biotransform Sb.

Metabolic Functions Revealed by Metagenomics. The analysis of the 16S rRNA gene alone may not provide direct evidence of the metabolic potential. We performed shotgun metagenomics analysis to better understand the metabolic pathways in the contaminated soil and unravel the metagenomes of As- and Sb-contaminated habitats, with a focus on As-related genes. Genes encoding As metabolism were detected in the five metagenomics libraries developed from the representative samples over a wide range of contamination levels. Genes related to aoxA, aoxB and As(V) reductase (ARSC1 and ARSC2) were detected in all libraries, suggesting possible biogeochemical As redox reactions mediated by the microbial communities in these soils (Figure 5). Because As and Sb share structural similarities, microorganisms may use similar metabolic pathways to transform both As and Sb. Indeed, both As and Sb will induce the arsenic resistance operon, which contains *arsR*, arsenical pump membrane protein (arsB), and As(V) reductase (arsC).⁶⁹ Therefore, Sb(V) reduction may contribute to the enrichment of As(V) reductase in this study. It is fair to propose that Sb(III) oxidation may be catalyzed by the enzymes encoded by the same functional genes (e.g., *aoxA* and *aoxB*), whereas alternative pathways may also exist. An Agrobacterium tumefaciens mutant was reported to oxidize Sb(III) without detectable aoxAB expression, which suggests a different pathway than that used for As(III) oxidation.¹⁵ This finding was supported by a later study using different isolates,²⁰ and may explain the relatively low abundance of aoxA/B genes and no significant correlations between aoxA/B genes and Sb(III)-C and Sb(V)-C. Additionally, Sb(III) oxidation was shown to be partially inhibited in mutants with the knockout of aioA (encoding the large subunit of As(III) oxidase), suggesting that Sb(III) oxidation is partially mediated by aioA but that other unknown pathways must also exist.⁷⁰ The gene encoding for the arsenical pump membrane protein (arsB) was detected in the metagenome across all 5 libraries. Interestingly, Meng et al.⁷¹ reported that ArsB can catalyze the transport of Sb(III). Analysis of the genome of Comamonas sp. S44, an antimony-oxidizing bacterium, revealed the presence of genes encoding for ArsB.

In addition to the presence of As-related genes across all samples, we observed significant positive correlations among As-related genes and various contaminant fractions-not only of As fractions, but those of Sb as well. For instance, Sb(III)-C and Sb(V)-C are positively correlated with *arsC* and *arsH* while Sb_{exe} is positively correlated with As(III) methyltransferase (AS3MT), aoxA, and arsH, suggesting that Sb contamination may promote the proliferation of such genes (Figure 6A). The relatively high abundances of arsR suggest the enrichment of bacterial resistance in response to contamination. This resistance may be stimulated by the elevated Sb concentrations as indicated by the positive correlation with Sb_{tot} but not As_{tot}. Consistent with this observation, inactivation of arsR resulted in reduction of resistance to As(III) and Sb(III) in Staphylococcus xylosus plasmid pSX267.73 The metagenomics analysis suggests an alternative way to decipher the Sb-metabolizing capabilities within complicated communities, even though microbiological Sb interactions are currently poorly understood.

Although Sb and As cannot be removed from the environment, microorganisms may play an important role in their detoxification (e.g., As(III) and Sb(III) oxidation). The current study investigated the microbial response to As and Sb contamination and identified the potential functional members of the microbial community and key functional genes. These

observations can be instructive for cultivating active As/Sb metabolizing organisms that can influence the solubility and toxicity of these metalloids. Additionally, numerical analyses predicted the threshold values affecting microbial diversity, which may be informative for bench-scale microcosm and mesocosm studies, or used as an indicator for in situ bioremediation. These observations collectively hold the potential to provide fundamental knowledge for implementing an in situ bioremediation strategy.

ASSOCIATED CONTENT

S Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.est.7b00294.

Additional information as noted in the text (PDF)

AUTHOR INFORMATION

Corresponding Author

*Phone: +86-20-39366937; e-mail: tfxiao@gzhu.edu.cn. ORCID •

Tangfu Xiao: 0000-0001-8076-0322 Yiran Dong: 0000-0002-2092-1355 Song Tang: 0000-0003-3219-1422

Author Contributions

[¶]These authors contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This research was funded by the Public Welfare Foundation of the Ministry of Water Resources of China (201501011), the High-level Leading Talent Introduction Program of GDAS (2016GDASRC-0103), the Natural Science Foundation of China (41103080; 41473124; 41420104007), the Opening Fund of the State Key Laboratory of Environmental Geochemistry (SKLEG2016907), and SPICC Program (2016GDASPT-0105).

REFERENCES

(1) Filella, M.; Williams, P. A.; Belzile, N. Antimony in the environment: knowns and unknowns. *Environ. Chem.* **2009**, *6*, 95–105.

(2) Fu, Z.; Wu, F.; Mo, C.; Deng, Q.; Meng, W.; Giesy, J. P. Comparison of arsenic and antimony biogeochemical behavior in water, soil and tailings from Xikuangshan, China. *Sci. Total Environ.* **2016**, 539, 97–104.

(3) Mitsunobu, S.; Harada, T.; Takahashi, Y. Comparison of antimony behavior with that of arsenic under various soil redox conditions. *Environ. Sci. Technol.* **2006**, *40*, 7270–7276.

(4) Majzlan, J.; Lalinska, B.; Chovan, M.; Blass, U.; Brecht, B.; Gottlicher, J.; Steininger, R.; Hug, K.; Ziegler, S.; Gescher, J. A mineralogical, geochemical, and microbiogical assessment of the antimony- and arsenic-rich neutral mine drainage tailings near Pezinok, Slovakia. *Am. Mineral.* **2011**, *96*, 1–13.

(5) Filella, M.; Belzile, N.; Chen, Y. W. Antimony in the environment: a review focused on natural waters: I. Occurrence. *Earth-Sci. Rev.* **2002**, *57*, 125–176.

(6) Singh, R.; Singh, S.; Parihar, P.; Singh, V. P.; Prasad, S. M. Arsenic contamination, consequences and remediation techniques: a review. *Ecotoxicol. Environ. Saf.* **2015**, *112*, 247–270.

(7) Abedin, M. J.; Cresser, M. S.; Meharg, A. A.; Feldmann, J.; Cotter-Howells, J. Arsenic accumulation and metabolism in rice (Oryza sativa L.). *Environ. Sci. Technol.* **2002**, *36*, 962–968. (8) Jain, C.; Ali, I. Arsenic: occurrence, toxicity and speciation techniques. *Water Res.* **2000**, *34*, 4304–4312.

(9) Wilson, S. C.; Lockwood, P. V.; Ashley, P. M.; Tighe, M. The chemistry and behaviour of antimony in the soil environment with comparisons to arsenic: a critical review. *Environ. Pollut.* **2010**, *158*, 1169–1181.

(10) U.S Geological Survey. Antimony statistics and information. 2016, https://minerals.usgs.gov/minerals/pubs/commodity/antimony//index.html.

(11) He, M.; Wang, X.; Wu, F.; Fu, Z. Antimony pollution in China. *Sci. Total Environ.* **2012**, *421*, 41–50.

(12) Frankenberger, Jr, W. T. Environmental Chemistry of Arsenic; CRC Press: FL, 2001.

(13) Fowler, B. A. *Biological and Environmental Effects of Arsenic*, vol 6; Elsevier, Sole distributors for the USA. and Canada, Elsevier Science Pub. Co., 2013.

(14) Stolz, J. F.; Basu, P.; Santini, J. M.; Oremland, R. S. Arsenic and Selenium in Microbial Metabolism. *Annu. Rev. Microbiol.* **2006**, *60*, 107–130.

(15) Lehr, C. R.; Kashyap, D. R.; McDermott, T. R. New insights into microbial oxidation of antimony and arsenic. *Appl. Environ. Microbiol.* 2007, *73*, 2386–2389.

(16) Smedley, P.; Kinniburgh, D. A review of the source, behaviour and distribution of arsenic in natural waters. *Appl. Geochem.* **2002**, *17*, 517–568.

(17) Bissen, M.; Frimmel, F. H. Arsenic—a review. Part I: occurrence, toxicity, speciation, mobility. *Acta Hydrochim. Hydrobiol.* **2003**, *31*, 9–18.

(18) Mandal, B. K.; Suzuki, K. T. Arsenic round the world: a review. *Talanta* **2002**, *58*, 201–235.

(19) Abin, C. A.; Hollibaugh, J. T. Dissimilatory antimonate reduction and production of antimony trioxide microcrystals by a novel microorganism. *Environ. Sci. Technol.* **2013**, *48*, 681–688.

(20) Terry, L. R.; Kulp, T. R.; Wiatrowski, H.; Miller, L. G.; Oremland, R. S. Microbiological oxidation of antimony (III) with oxygen or nitrate by bacteria isolated from contaminated mine sediments. *Appl. Environ. Microbiol.* **2015**, *81*, 8478–8488.

(21) Kulp, T. R.; Miller, L. G.; Braiotta, F.; Webb, S. M.; Kocar, B. D.; Blum, J. S.; Oremland, R. S. Microbiological reduction of Sb (V) in anoxic freshwater sediments. *Environ. Sci. Technol.* **2013**, *48*, 218–226.

(22) Sun, W.; Xiao, E.; Dong, Y.; Tang, S.; Krumins, V.; Ning, Z.; Sun, M.; Zhao, Y.; Wu, S.; Xiao, T. Profiling microbial community in a watershed heavily contaminated by an active antimony (Sb) mine in Southwest China. *Sci. Total Environ.* **2016**, *550*, 297–308.

(23) Xiao, E.; Krumins, V.; Dong, Y.; Xiao, T.; Ning, Z.; Xiao, Q.; Sun, W. Microbial diversity and community structure in an antimonyrich tailings dump. *Appl. Microbiol. Biotechnol.* **2016**, *100*, 1–13.

(24) Xiao, E.; Krumins, V.; Tang, S.; Xiao, T.; Ning, Z.; Lan, X.; Sun, W. Correlating microbial community profiles with geochemical conditions in a watershed heavily contaminated by an antimony tailing pond. *Environ. Pollut.* **2016**, *215*, 141–153.

(25) Sun, W.; Xiao, E.; Kalin, M.; Krumins, V.; Dong, Y.; Ning, Z.; Liu, T.; Sun, M.; Zhao, Y.; Wu, S. Remediation of antimony-rich mine waters: Assessment of antimony removal and shifts in the microbial community of an onsite field-scale bioreactor. *Environ. Pollut.* **2016**, *215*, 213–222.

(26) Fuentes, E.; Pinochet, H.; De Gregori, I.; Potin-Gautier, M. Redox speciation analysis of antimony in soil extracts by hydride generation atomic fluorescence spectrometry. *Spectrochim. Acta, Part B* **2003**, *58*, 1279–1289.

(27) Ge, Z.; Wei, C. Simultaneous analysis of SbIII, SbV and TMSb by high performance liquid chromatography-inductively coupled plasma-mass spectrometry detection: application to antimony speciation in soil samples. *J. Chromatogr. Sci.* **2013**, *51*, 391–399.

(28) Xiao, E.; Krumins, V.; Xiao, T.; Dong, Y.; Tang, S.; Ning, Z.; Huang, Z.; Sun, W. Depth-resolved microbial community analyses in two contrasting soil cores contaminated by antimony and arsenic. *Environ. Pollut.* **2016**, *221*, 244–255.

(29) Wenzel, W. W.; Kirchbaumer, N.; Prohaska, T.; Stingeder, G.; Lombi, E.; Adriano, D. C. Arsenic fractionation in soils using an improved sequential extraction procedure. *Anal. Chim. Acta* **2001**, *436*, 309–323.

(30) Gault, A. G.; Polya, D. A.; Charnock, J. M.; Islam, F. S.; Lloyd, J. R.; Chatterjee, D. Preliminary EXAFS studies of solid phase speciation of As in a West Bengali sediment. *Mineral. Mag.* **2003**, *67*, 1183–1191.

(31) Xiao, E.; Krumins, V.; Xiao, T.; Dong, Y.; Tang, S.; Ning, Z.; Huang, Z.; Sun, W. Depth-resolved microbial community analyses in two contrasting soil cores contaminated by antimony and arsenic. *Environ. Pollut.* **2017**, *221*, 244–255.

(32) Magoč, T.; Salzberg, S. L. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **2011**, *27*, 2957–2963.

(33) Caporaso, J. G.; Kuczynski, J.; Stombaugh, J.; Bittinger, K.; Bushman, F. D.; Costello, E. K.; Fierer, N.; Peña, A. G.; Goodrich, J. K.; Gordon, J. I. QIIME allows analysis of high-throughput community sequencing data. *Nat. Method* **2010**, *7*, 335–336.

(34) Haas, B. J.; Gevers, D.; Earl, A. M.; Feldgarden, M.; Ward, D. V.; Giannoukos, G.; Ciulla, D.; Tabbaa, D.; Highlander, S. K.; Sodergren, E. Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res.* **2011**, *21*, 494– 504.

(35) DeSantis, T. Z.; Hugenholtz, P.; Larsen, N.; Rojas, M.; Brodie, E. L.; Keller, K.; Huber, T.; Dalevi, D.; Hu, P.; Andersen, G. L. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **2006**, *72*, 5069–5072.

(36) Wang, Q.; Garrity, G. M.; Tiedje, J. M.; Cole, J. R. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **2007**, *73*, 5261–5267.

(37) Huson, D. H.; Auch, A. F.; Qi, J.; Schuster, S. C. MEGAN analysis of metagenomic data. *Genome Res.* 2007, 17, 377–386.

(38) Breiman, L. Random forests. Mach Learn 2001, 45, 5-32.

(39) Friedman, J, Hastie, T, Tibshirani, R. *The Elements of Statistical Learning*, Springer series in statistics, Springer: Berlin, 2001; Vol. 1.

(40) Breiman, L. Bagging predictors. Mach Learn. 1996, 24, 123–140.

(41) Wang, Q.; Xie, Z.; Li, F. Using ensemble models to identify and apportion heavy metal pollution sources in agricultural soils on a local scale. *Environ. Pollut.* **2015**, *206*, 227–235.

(42) Díaz, U. R.; De, A. S. A. Gene selection and classification of microarray data using random forest. *BMC Bioinf.* **2006**, *7*, 3.

(43) Fielding, A. H.; Bell, J. F. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ. Conserv.* **1997**, *24*, 38–49.

(44) John, L. Z. The elements of statistical learning: data mining, inference, and prediction. J. R Stat Soc. A 2010, 173, 693-694.

(45) Friedman, J. H. Stochastic Gradient Boosting. *Comput. Stat Data An* **1999**, *38*, 367–378.

(46) Friedman, J. H. Greedy function approximation: A gradient boosting machine. *Ann. Stat* **2001**, *29*, 1189–1232.

(47) Breiman, L. Using Adaptive Bagging to Debias Regressions; University of California: Berkeley, 1999.

(48) Newman, M. E. The structure and function of complex networks. *SIAM Rev.* 2003, 45, 167–256.

(49) Newman, M. E. Modularity and community structure in networks. *Proc. Natl. Acad. Sci. U. S. A.* 2006, *103*, 8577–8582.

(50) Csardi, G.; Nepusz, T. The i-graph software package for complex network research. *Int. J. Com Sys.* **2006**, *1695*, 1–9.

(51) Savonina, E. Y.; Fedotov, P. S.; Wennrich, R. Fractionation of Sb and As in soil and sludge samples using different continuous-flow extraction techniques. *Anal. Bioanal. Chem.* **2012**, 403, 1441–1449.

(52) Buanuam, J.; Wennrich, R. Dynamic flow-through sequential extraction for assessment of fractional transformation and interelement associations of arsenic in stabilized soil and sludge. *J. Hazard. Mater.* **2010**, *184*, 849–854. (53) Fuentes, E.; Pinochet, H.; Potin-Gautier, M.; De, G. I. Fractionation and redox speciation of antimony in agricultural soils by hydride generation-atomic fluorescence spectrometry and stability of Sb(III) and Sb(V) during extraction with different extractant solutions. *J. Aoac Int.* **2004**, *87*, 60–67.

(54) Roger, G.; Amaral, L. A. Nunes, Functional cartography of complex metabolic networks. *Nature* **2005**, *433*, 895–900.

(55) Hartwell, L. H.; Hopfield, J. J.; Leibler, S.; Murray, A. W. From molecular to modular cell biology. *Nature* **1999**, *402*, 47–52.

(56) Li, B.; Yang, Y.; Ma, L.; Ju, F.; Guo, F.; Tiedje, J. M.; Zhang, T. Metagenomic and network analysis reveal wide distribution and cooccurrence of environmental antibiotic resistance genes. *ISME J.* **2015**, *9*, 2490–2502.

(57) Hu, H. W.; Wang, J. T.; Li, J.; Li, J. J.; Ma, Y. B.; Chen, D.; He, J. Z. Field-based evidence for copper contamination induced changes of antibiotic resistance in agricultural soils. *Environ. Microbiol.* **2016**, *18*, 3896–3909.

(58) Costa, P. S.; Scholte, L. L.; Reis, M. P.; Chaves, A. V.; Oliveira, P. L.; Itabayana, L. B.; Suhadolnik, M. L. S.; Barbosa, F. A.; Chartone-Souza, E.; Nascimento, A. M. Bacteria and genes involved in arsenic speciation in sediment impacted by long-term gold mining. *PLoS One* **2014**, *9*, e95655.

(59) Abdrashitova, S.; Mynbaeva, B.; Ilialetdinov, A. Arsenic oxidation by the heterotrophic bacteria Pseudomonas putida and Alcaligenes eutrophus. *Mikrobiologiia* **1980**, *50*, 41–45.

(60) Li, P.; Jiang, D.; Li, B.; Dai, X.; Wang, Y.; Jiang, Z.; Wang, Y. Comparative survey of bacterial and archaeal communities in high arsenic shallow aquifers using 454 pyrosequencing and traditional methods. *Ecotoxicology* **2014**, *23*, 1878–1889.

(61) Bond, P. L.; Smriga, S. P.; Banfield, J. F. Phylogeny of microorganisms populating a thick, subaerial, predominantly litho-trophic biofilm at an extreme acid mine drainage site. *Appl. Environ. Microbiol.* **2000**, *66*, 3842–3849.

(62) Ho, A.; Angel, R.; Veraart, A. J.; Daebeler, A.; Jia, Z.; Kim, S. Y.; Kerckhof, F.-M.; Boon, N.; Bodelier, P. L. Biotic Interactions in Microbial Communities as Modulators of Biogeochemical Processes: Methanotrophy as a Model System. *Front. Microbiol.* **2016**, *7*, 1285.

(63) Li, X.; Hu, Y.; Gong, J.; Lin, Y.; Johnstone, L.; Rensing, C.; Wang, G. Genome sequence of the highly efficient arsenite-oxidizing bacterium Achromobacter arsenitoxydans SY8. *J. Bacteriol.* **2012**, *194*, 1243–1244.

(64) Prasad, K. S.; Subramanian, V.; Paul, J. Purification and characterization of arsenite oxidase from Arthrobacter sp. *BioMetals* **2009**, *22*, 711–721.

(65) Ramos, J. L.; Duque, E.; Huertas, M. J.; HaïDour, A. Isolation and expansion of the catabolic potential of a Pseudomonas putida strain able to grow in the presence of high concentrations of aromatic hydrocarbons. J. Bacteriol. **1995**, 177, 3911–3916.

(66) Deziel, E.; Paquette, G.; Villemur, R.; Lepine, F.; Bisaillon, J. Biosurfactant production by a soil pseudomonas strain growing on polycyclic aromatic hydrocarbons. *Appl. Environ. Microbiol.* **1996**, *62*, 1908–1912.

(67) Whyte, L. G.; Bourbonniere, L.; Greer, C. W. Biodegradation of petroleum hydrocarbons by psychrotrophic Pseudomonas strains possessing both alkane (alk) and naphthalene (nah) catabolic pathways. *Appl. Environ. Microbiol.* **1997**, *63*, 3719–3723.

(68) Ilialetdinov, A.; Abdrashitova, S. Autotrophic arsenic oxidation by a Pseudomonas arsenitoxidans culture. *Mikrobiologiia* **1980**, *50*, 197–204.

(69) Ji, G.; Silver, S. Reduction of arsenate to arsenite by the ArsC protein of the arsenic resistance operon of Staphylococcus aureus plasmid pI258. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 9474–9478.

(70) Wang, Q.; Warelow, T. P.; Kang, Y. S.; Romano, C.; Osborne, T. H.; Lehr, C. R.; Bothner, B.; McDermott, T. R.; Santini, J. M.; Wang, G. (2015) Arsenite oxidase also functions as an antimonite oxidase. *Appl. Environ. Microbiol.* **2015**, *81*, 1959–1965.

(71) Meng, Y.-L.; Liu, Z.; Rosen, B. P. As (III) and Sb (III) uptake by GlpF and efflux by ArsB in Escherichia coli. *J. Biol. Chem.* **2004**, *279*, 18334–18341.

(72) Li, J.; Wang, Q.; Zhang, S.; Qin, D.; Wang, G. Phylogenetic and genome analyses of antimony-oxidizing bacteria isolated from antimony mined soil. *Int. Biodeterior. Biodegrad.* **2013**, *76*, 76–80.

(73) Rosenstein, R.; Peschel, A.; Wieland, B.; Götz, F. Expression and regulation of the antimonite, arsenite, and arsenate resistance operon of *Staphylococcus xylosus plasmid* pSX267. *J. Bacteriol.* **1992**, 174, 3676–3683.