Bacterial community analysis of stabilized soils in proximity to an exhausted mine

Jae Eun Park[1], Byung-Tae Lee[2], Byung-Yong Kim[3], Ahjeong Son[1†]

[1]Department of Environmental Science and Engineering, Ewha Womans University, Seoul 03760, Republic of Korea
[2]Department of Environmental Science and Engineering, Gwangju Institute of Science and Technology (GIST), Gwangju 61004, Republic of Korea
[3]Bioinformatcis Research Institute, Chun Lab, Inc., Seoul 06725, Republic of Korea

ABSTRACT

Soil stabilization is a soil remediation technique that reduces the mobility of heavy metals in soils. Although it is a well-established technique, it is nonetheless essential to perform a follow-up chemical assessment via a leaching test to evaluate the immobilization of heavy metals in the soil matrix. Unfortunately, a standard chemical assessment is not sufficient for evaluation of the biological functional state of stabilized soils slated for agricultural use. Therefore, it is useful to employ a pyrosequencing-based microbial community analysis for the purpose. In this study, a recently stabilized site in the proximity of an exhausted mine was analyzed for bacterial diversity, richness, and relative abundance as well as the effect of environmental factors. Based on the Shannon and Chao1 indices and rarefaction curves, the results showed that the stabilized layer exhibited lower bacterial diversity than control soils. The prevalence of dominant bacterial populations was examined in a hierarchical manner. Relatively high abundances of *Proteobacteria* and *Methylobacter tundripaludum* were observed in the stabilized soil. In particular, there was substantial abundance of the *Methylobacter* genus, which is known for its association with heavy metal contamination. The study demonstrated the efficacy of (micro)biological assessment for aiding in the understanding and post-management of stabilized soils.

Keywords: Bacterial community analysis, Heavy metal, *Methylobacter*, Pyrosequencing, Soil stabilization

1. Introduction

Once heavy metals move into soil, they can remain in the soil for a long period of time-approaching thousands of years [1-5]. Heavy metals can enter soils from several sources, including waste from mines and landfills [2, 6]. The mining industry has been a major source of heavy metal release to the environment [7-9]. Therefore, heavy metal contamination from mining activities has drawn increased attention over the last few decades.

Soil stabilization is a soil remediation technique that reduces the solubility and mobility of heavy metals by the addition of stabilizing agents (or stabilizers) to contaminated soils [8, 10]. The main stabilizers include lime and steel slag [8, 10, 11]. The stabilizers decrease the leaching and bioavailability of trace elements by various sorption processes. Although the stabilization process is often influenced by several environmental factors, including pH and organic matter [10, 12], it is simple and effective.

Furthermore, soil stabilization technology is well established for the mitigation of most heavy metals in soils.

However, it is also known that the soil stabilization method only temporarily immobilizes heavy metals in the soil matrix [10, 12]. Thus, a follow-up analysis is essential to monitor the mobility of heavy metals after soil stabilization. Currently, chemical assessment via a leaching test is a common practice for managing stabilized soils. However, the chemical assessment is not sufficient for evaluation of the biological functional state of stabilized soils. Assessment of the biological functional state of soils is particularly important for soils remediated by soil stabilization that are to be used for agricultural purposes.

To evaluate the biological functional state of soils after soil stabilization, microbial community analysis was performed. Among all possible bio-indicators, microorganisms or their genomes can be among the most suitable candidates for indicating the terrestrial ecosystem dynamics of heavy metals. Soil micro-

organisms provide a more reliable and faster result than other organisms due to their large quantities and fast turn-over. In particular, the prevalence of uncommon (i.e., selectively grown) microorganism species can be useful for identifying site abnormalities.

The objectives of this study were to evaluate the biological functional state of stabilized soils and to identify the prevalence of selectively grown microorganisms at the site. To achieve these objectives, soil bacterial community analysis was performed using pyrosequencing techniques. Bacterial diversity and richness were investigated using pyrosequencing data and diversity indices. The relative abundances and profiling of the microbial populations were reported for taxonomic levels ranging from phyla to species, and the effect of environmental factors (including heavy metal concentrations) were examined with multivariate statistical tools. The prevalence of dominant bacterial populations at the stabilized site were also analyzed.

2. Materials and Methods

2.1. Study Sites and Soil Sampling

Soil samples were collected from paddy fields near an exhausted mine (Seosung mine) located in Seosan, Chungcheongnam-do, Korea. The Seosung mine had provided Zn, Ag, Pb, and Au from the 1930s to the late 1970s to Korea. Until 2007 and prior to soil reclamation, soils in the vicinity of the sampling site were heavily contaminated with Cd and Pb [8, 13] from the tailings of the Seosung mine. In 2008, a major soil reclamation

was implemented for the Seosung mine along with 1,344 other mines in Korea [13, 14]. The stabilization was implemented at a depth of 40-60 cm with 5% lime stone and 2% steel slag. The stabilized layer was covered with soil excavated from a non-contaminated site in Korea. Soil from each stabilized layer (40-60 cm) and the cover layer (20-40 cm) was sampled in 2014. Stabilized and cover-layer soils were denoted as SS and CS, respectively. A non-contaminated (NC) sample was collected at the same time from the surface of a paddy approximately 6 km away from the stabilized site. The NC sample was taken from a pristine surface soil with no history of heavy metal contamination. The CS and NC samples were both used as negative controls in the experiment. The CS negative control was originally a foreign soil that was used to cover the stabilized soil layer and subsequently had direct contact with the stabilized layer below it for over 5 y.

Information on the sampling sites is presented in Fig. 1. Fig. 1(a) shows a geological map of the sampling site in Seosan, Korea. Fig. 1(b) and 1(c) show satellite pictures of the sampling site (36°89' N, 126°41' E) and the control site (36°84' N, 126°43' E).

Approximately 2 kg of soils were collected at each sample point using a hand auger. The soil samples were transferred to plastic-lined canvas bags and transported to the laboratory. The samples for physicochemical analysis were air-dried for ~10 d. A portion of the dried soil was sieved through 100-mm mesh and subjected to heavy metal analysis. The remainder of the dried sample was sieved successively through appropriate mesh sizes and a final 10-mm mesh, and then subjected to pH, cation exchange capacity (CEC), and loss of ignition (LOI)
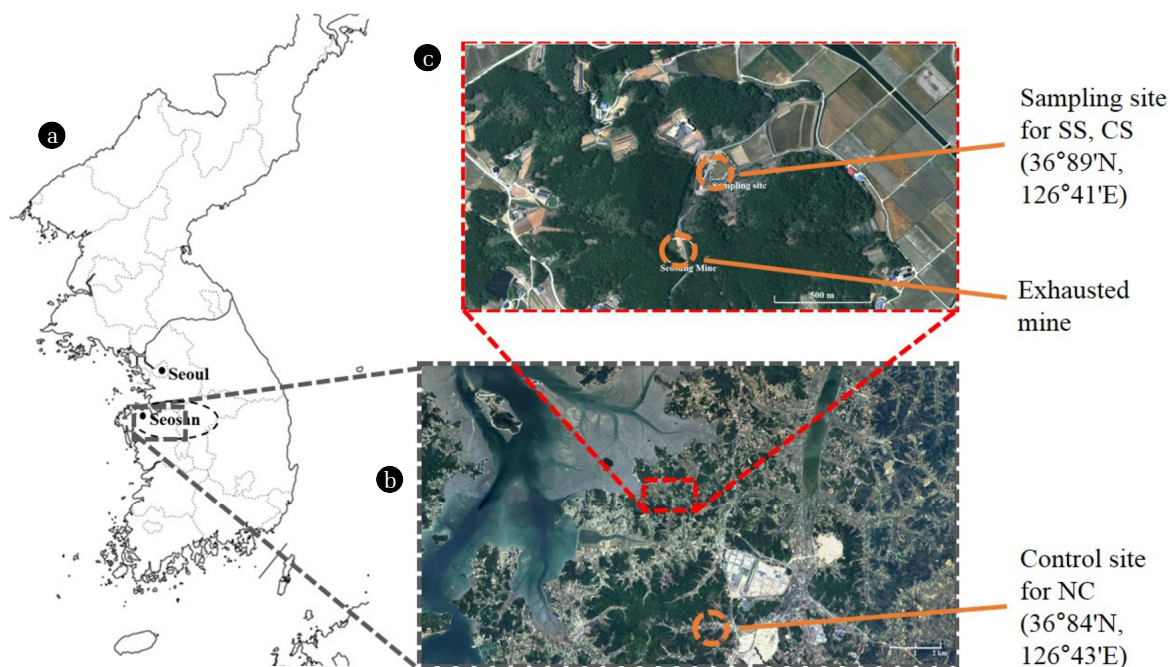

Fig. 1. (a) Geological map of the sampling site in Seosan, Korea. (b) and (c) Satellite pictures of the reclaimed site and the control site near Seosung mine. SS and CS were obtained from the reclaimed site and NC from the control site. The NC, CS, and SS refer to negative control, cover layer soil, and stabilized layer soil, respectively. The same description applies from Fig. 1 to Fig. 5.

measurements. The soil samples for DNA extraction were immediately stored after collection at -20℃ without air-drying.

2.2. Physicochemical Characterization of Soils

The physicochemical properties of the three soil samples, including moisture content, pH, CEC, and LOI, were examined. The pH was measured in a 1:5 soil-water slurry with an Orion 5 Star pH meter (Thermo Scientific, Barrington, IL, USA). The CEC was measured according to USEPA 9081 [15], and the LOI was measured using a thermogravimetric analyzer according to ASTM D7348-08 [16].

Heavy metal concentrations were measured in triplicate samples. Total concentrations of As, Cd, Cu, Ni, Pb, and Zn were determined by aqua regia digestion according to the Korean standard method for chemical analysis of soils [17]. For soil digestion, 1 mL of HNO_3 and 3 mL of HCl were added to 0.25 g of the soil. All the samples were analyzed by inductively coupled plasma-optical emission spectrometry (ICP-OES, Thermo Scientific, Franklin, MA, USA) at the Gwangju Institute of Science and Technology (Gwangju, Korea).

2.3. DNA Extraction

Genomic DNA (gDNA) was extracted from the soil samples using the commercial soil DNA extraction kit, PowerSoil DNA Isolation Kit (Mo Bio Laboratories, Solon, OH). Approximately 0.5 g (wet weight) of soil was used, and the extraction was performed on triplicate samples. DNA concentrations for all samples were obtained using the Nanodrop ND 2000 Spectrophotometer (Thermo Scientific, Barrington, IL, USA). The quantity and purity of the DNA were determined by OD_{280} and OD_{260} measurement.

2.4. 16S rRNA Pyrosequencing Analysis

Prior to the pyrosequencing analysis, PCR amplification of extracted gDNA (triplicate samples) was performed using primers targeting the V1 to V3 region of the 16S rRNA gene. The barcoded primers of 9F (5'-CCTATCCCCTGTGTGCCTTGGCAGTCTCA GACAGA GTTTGATCMTGGCTCAG-3'; underlined sequences indicating the target region of the primer) and 541R (5'-CCATCTCATCCCTGCGTGTCTCCGACTCAGXACATTACC GCG GCTGCTGG-3') [18] were used for PCR amplification. The PCR amplification was performed under the following thermal conditions: initial denaturation at 95℃ for 5 min, followed by 30 cycles of denaturation at 95℃ for 30 s, primer annealing at 55℃ for 30 s, and extension at 72℃ for 30 s, with a final elongation at 72℃ for 5 min. The PCR amplicons were confirmed by using 2% agarose gel electrophoresis and visualized with a Gel Doc system (BioRad, Hercules, CA, USA). Subsequently, the amplicons were purified with the QIAquick PCR purification kit (Qiagen, Valencia, CA, USA). The short fragments (*i.e.*, non-target products) were removed with Ampure beads (Agencourt Bioscience, MA, USA). The purity and the size of amplicons was assessed with a Bioanalyzer 2100 (Agilent, Palo Alto, CA, USA) using a DNA 7500 chip. Subsequently, the pyrosequencing was implemented with a Roche/454 GS Junior Sequencing system (Roche, Branford, CT, USA) at Chunlab Inc. (Seoul, Korea).

The results obtained using 454-pyrosequencing of the 16S rRNA gene were used to determine the richness and diversity of samples as well as to identify predominant bacterial species or populations of stabilized layer samples (triplicate).

2.5. Pyrosequencing Data Processing

The basic processing of the pyrosequencing data was conducted according to previous studies [18-20]. Sequencing reads from the samples were sorted by unique barcodes of each PCR amplicons. The sequences of the barcode, linker, and primers were removed from the original sequencing reads. Any reads containing two or more ambiguous nucleotides, a low quality score (< 25), or less than 300 bp were discarded. Potential chimeric sequences were detected by the Bellerophon method, which compares the BLASTN search results between forward half and reverse half sequences. After removing chimeric sequences, the taxonomic classification of each read was assigned based on comparison against the EzTaxon-e database [20].

To further analyze the diversity and richness of soil bacterial communities, the CLcommunity program (Chunlab) was used. The richness and diversity indices were determined by the Chao1 [21] and Shannon diversity estimation [22]. Random subsampling was conducted to normalize the read size of samples in order to compare different read sizes among the samples. The overall phylogenetic distance between communities was estimated using Fast UniFrac. To compare operational taxonomic units (OTUs) between samples, shared OTUs (*i.e.*, the same OTUs that were obtained from different samples) were obtained with an XOR analysis. Rarefaction curves were also developed from the OTUs.

2.6. Statistical Analysis

Multivariate statistical analyses were conducted to determine the correlations between species abundance and environmental variables. A principal component analysis (PCA) was conducted with the relative abundance of each set of species data. A canonical correspondence analysis (CCA) was performed to determine the significant environmental variables (*e.g.*, moisture contents, pH, CEC, LOI, and heavy metal concentrations) associated with differences in the soil bacterial communities. The PCA and CCA are the most commonly used multivariate statistical tools for environmental studies [23-25]. The Monte Carlo permutation (*i.e.*, randomization) test [26] was used for CCA. Both PCA and CCA were performed using CANOCO 4.5.0 software (Microcomputer Power, Ithaca, USA). Community similarities were plotted by using ordination plots with scaling focused on inter-sample differences.

2.7. Phylogenetic Affiliation Analysis

To visually construct the phylogenetic affiliation of dominant species in the SS samples, a phylogenetic tree was constructed with 13 representative bacterial 16S rRNA gene sequences of *γ*-proteobacteria. Multiple-sequence alignment was performed on the sequencing data obtained from the pyrosequencing analysis by ClustalX [27] and BioEdit V7.0.0 [28]. Phylogenetic trees were constructed by neighbor-joining analysis using MEGA6 [29] with a bootstrap value of 1,000 replicates.

3. Results

3.1. Soil Properties and Heavy Metal Concentrations

Physicochemical properties of the soil samples (SS, CS, and NC) are presented in Table 1. As explained in the material and methods, SS refers to the target sample obtained from the stabilized layer, and CS and NC refer to the negative control soils. The pH of the three soils ranged from 6.2 to 8.7. Compared to the two negative control soils, the SS showed the highest CEC (cation exchange capacity) and also showed the lowest LOI, which refers to organic contents.

The concentrations of six heavy metals (As, Cd, Cu, Pb, Zn, and Ni) in three soil samples are also shown in Table 1. In contrast to the two negative control soils, the SS was heavily contaminated with all six heavy metals. In Korea, the maximum allowable concentration of corresponding heavy metals (As, Cd, Cu, Pb, Zn, and Ni) in soils is 25, 4, 150, 200, 300, and 100 mg/kg, respectively. Therefore, all of the heavy metal concentrations except Ni in the SS sample were higher than the maximum allowable concentration. In particular, the SS sample had ~50 and ~15 times higher Pb (10,123 ± 132 mg/kg) and Cd (60.1 ± 0.5 mg/kg), respectively, than the maximum allowable concentrations. In contrast, NC and CS had negligible concen-

trations of heavy metals compared to the maximum allowable concentrations. These results confirmed that the NC and CS soils were both suitable to serve as negative control soils.

3.2. Bacterial Diversity and Richness

Table 2 presents the indices of bacterial diversity and richness, which were calculated from triplicate samples. The OTUs, Shannon and Chao1 indices, and estimated coverage of each sample were determined using the sequences obtained from pyrosequencing. To precisely compare indices among SS, CS, and NC samples, the valid reads for the sequences were normalized by random selection. As a result, 4,187 fixed sequence reads were obtained for each of the triplicate samples of each soil type. The OTUs ranged from 1,185 to 2,039, with the largest value obtained from the NC. The Good's coverage index for each sample, calculated at a 97% similarity cut-off, ranged from 70% to 83%. For SS, the Good's index of 83% indicated that 17 additional phylotypes would be discovered for every 100 additional sequencing efforts. However, the two negative control samples (CS and NC) had lower Good's coverage values (i.e., 70-72%), indicating that the number of pyrosequencing reads for each sample may not be sufficient to represent the overall bacterial diversity.

Table 1. Physicochemical Properties of the Sample Soils. Heavy Metal Concentrations Were Obtained from Triplicate Samples

Reference ID	pH	CEC[1] (meq/100 g)	LOI[2] (%)	As (mg/kg)	Cd (mg/kg)	Cu (mg/kg)	Pb (mg/kg)	Zn (mg/kg)	Ni (mg/kg)	Ca (mg/kg)	Fe (mg/kg)
SS (Stabilized layer soil)	8.7	11.8	5.8	161.0 ± 7.4	60.1 ± 0.5	437.8 ± 9.6	10,122.7 ± 132.4	5,939.6 ± 27.0	34.7 ± 0.7	56,039.2 ± 1,138.7	86,720.0 ± 1,327.0
CS (Cover layer soil)	7.1	2.5	10.8	N.D.[3]	2.7 ± 0.4	66.0 ± 7.7	147.3 ± 4.1	119.6 ± 10.5	14.1 ± 2.6	6,870.0 ± 70.7	26,969.0 ± 3,243.0
NC (Negative control at control site)	6.2	5.5	15.0	N.D.	N.D.	23.1 ± 1.5	22.6 ± 0.4	55.2 ± 4.7	7.1 ± 1.0	1,375.4 ± 217.7	21,820.0 ± 1,813.0

[1]Cation exchange capacity
[2]Loss of ignition
[3]N.D. stands for non-detected.

Table 2. Summary of Pyrosequencing Results for Triplicate Samples

Reference ID	Normalized reads	OTUs[1]	Goods Lib. Coverage[2]	Shannon index	Chao1 index
SS-1	4,178	1,185	83	5.98	2,454
SS-2	4,178	1,248	82	6.10	2,594
SS-3	4,178	1,206	83	6.05	2,640
CS-1	4,178	1,875	72	7.04	3,936
CS-2	4,178	1,867	72	6.97	3,802
CS-3	4,178	1,906	70	6.92	4,488
NC-1	4,178	2,039	70	7.24	4,307
NC-2	4,178	1,990	71	7.21	4,057
NC-3	4,178	2,036	70	7.24	4,053

[1]OTU stands for operational taxonomic unit
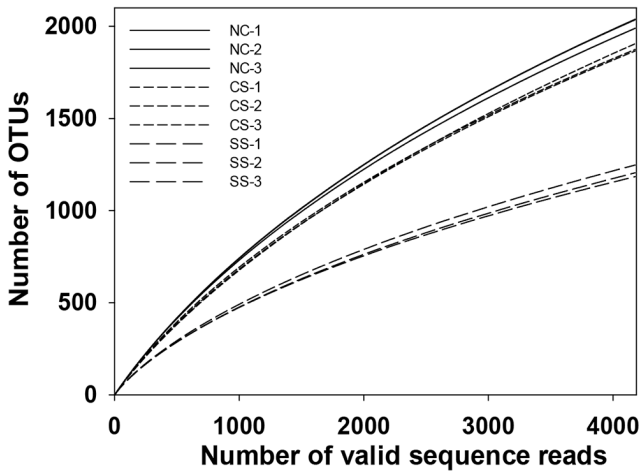[2]Calculated at a 97% sequences similarity cut-off (%)

Fig. 2. Rarefaction curves of triplicate samples of the three soil types. The long-dashed lines represent rarefaction curves of SS, short-dashed lines represent CS, and solid lines represent NC.

As an estimator of soil bacterial diversity, the Shannon index ranged from 5.98-7.24 among the 9 samples. The average Shannon index for SS was 6.04 ± 0.06, while the indices for CS and NC were 6.98 ± 0.06 and 7.23 ± 0.02, respectively. As an estimator of species richness or alpha-diversity, Chao1 explains the abundance of singleton species resulting in the greater species richness. Therefore, the value of Chao1 is proportional to the level of bacterial diversity. Chao1 ranged from 2,454-4,307 among the 9 samples. The SS Chao1 index was 2,562 ± 97, while the CS and NC indices were 4,075 ± 363 and 4,139 ± 146, respectively. Based on this comparison of the Shannon and Chao1 diversity indices, the bacterial diversity in SS is far lower than in the negative controls.

Rarefaction curves that plot the number of OTUs against the number of sequence reads for the 9 samples are shown in Fig. 2. The rarefaction analysis was consistent with the diversity indices in all samples. All the indices suggested that the soil from the stabilized layer (SS-1, SS-2, SS-3) had the lowest diversity, and the rarefaction analysis similarly indicated that the SS samples had lower numbers of OTUs per valid sequence read. In contrast, all of the CS and SS samples had higher numbers of OTUs per valid sequence read.

3.3. Soil Bacterial Community Analysis

Phyla level: Microbial phylum distributions in the samples are presented in Fig. 3(a). The analysis identified a total of 67 bacterial phyla, among which only 11 are depicted in Fig. 3(a). The remaining 56 phyla were removed because they represented less than 1% of the entire abundance. Three phyla represented 68-79% of the abundance: Proteobacteria, Chloroflexi, and Acidobacteria (depicted by a red-dashed square in Fig. 3(a)). Proteobacteria were the dominant phyla in all 3 soil types. The relative abundance of Proteobacteria in SS, CS, and NC was 57.0 ± 1.2%, 35.0 ± 1.6%, and 29.9 ± 1.0%, respectively. The relative abundances in SS, CS, and NC for Chloroflexi were 15.8 ± 0.5, 28.6 ± 1.2, and 28.1 ± 0.4, respectively, and for Acidobacteria were

5.5 ± 0.5, 5.4 ± 0.5, and 14.6 ± 1.0, respectively. Other abundant phyla were Actinobacteria (3-14%), Firmicutes (1-5%), and Bacteroidetes (1-5%).

The result of the PCA is presented in Fig. 3(b). The horizontal axis in Fig. 3(b) (PC1) explains 86.7% of the total variance in the data, and the vertical axis (PC2) represents an additional 12.4% of the variance. As indicated by the arrow on the right side of Fig. 3(b), PC1 was mainly affected by the Proteobacteria. As indicated by the dashed circle, a grouping occurred among the three SS samples and Proteobacteria, which can be explained by the relative proportion of Proteobacteria in the samples (Fig. 3(a)). The SS has the highest proportion (55%) of Proteobacteria compared to the other soil types (35% and 30% for CS and NS, respectively).
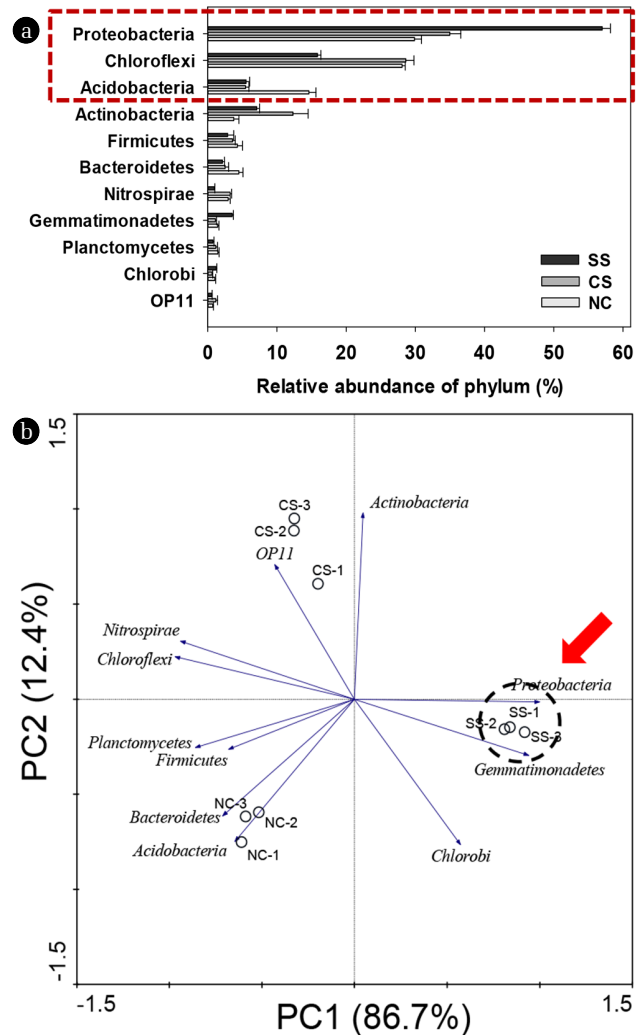

Fig. 3. Bacterial community analysis at the phylum level: (a) Taxonomic composition of phyla relative abundance. The bar and error bars represent average and standard deviation of triplicate samples. An arrow indicates the most abundant phylum, *Proteobacteria*. (b) Principal component analysis (PCA) plot of the taxonomic composition of the phyla in triplicate samples. The SS samples are grouped with the dominant phylum, *Proteobacteria* (dashed circle).

Class level: Microbial class distributions in the samples are presented in Fig. S1(a). Nineteen classes of over 1% relative abundance were selected and are shown in Fig. S1(a). Three classes represented 34-54% of the abundance: Anaerolineae, β-proteobacteria, and γ-proteobacteria (depicted by a red-dashed square in Fig. S1(a)). Other abundant classes were δ-proteobacteria (7.2-9.5%), α-proteobacteria (5.6-8.0%), and Actinobacteria (2.3-9.8%).

The result of the PCA at the class level is presented in Fig. 4(b). The horizontal axis in Fig. S1(b) (PC1) explains 82.9% of the total variance in the data and the vertical axis (PC2) represents an additional 11.7% of the variance. As indicated by the arrow on the right side of Fig. S1(b), PC1 was largely affected by β-proteobacteria and γ-proteobacteria. As indicated by the dashed circle, a grouping occurred among the three SS samples and β-proteobacteria, γ-proteobacteria and Gemmatimonadetes, which can be explained by the relative proportion of β-proteobacteria and γ-proteobacteria in the samples (Fig. S1(a)). The SS has the highest proportion of β-proteobacteria and γ-proteobacteria (24.5% and 19.5%, respectively) compared to CS and NC (14.2% and 9.3% of β-proteobacteria, and 5.7% and 3.4 of γ-proteobacteria, respectively).

Order/Family level: The distribution of taxonomic order and family levels are presented in the supporting information (Fig. S2 and S3). The dominance and grouping patterns at these levels were similar to those at the phyla and class levels. As indicated by the dashed circle in the PCA plot (Fig. S2(b)), a grouping occurred among the three SS samples and Sphingomonadales, Chromatiales, *Thiobacillus*, Strolibacterium, Methylococcales, and Gemmatimonadales. In particular, Methylococcales and *Thiobacillus* were dominant in the SS samples compared to levels in the CS and NC samples (the dashed square in Fig. S2(a)). In detail, Methylococcales constituted 15.9% of total abundance in SS, but only 0.9% of the abundance in each of the other soil types. *Thiobacillus* constituted 11.3% of the abundance in SS, but only 2.5% in CS and 0.3% in NC. Methylococcales and *Thiobacillus* are the dominant orders observed that belong to the classes of γ-proteobacteria and β-proteobacteria, respectively. Note that both γ-proteobacteria and β-proteobacteria were the dominant classes and were well grouped with the SS samples, as described above. At the family level, the pattern was similar to that observed at the order level. Sub-classification of the two observed dominant orders (i.e., Methylococcales, *Thiobacillus*) were from the Methylomonas and *Thiobacillus* families, which were more dominant in the SS samples than in the CS and NC samples. A grouping in the PCA plot was also observed among these families and SS samples (Fig. S3).

Genus level: Microbial genus distributions in the samples are presented in Fig. 4. Only 28-52% of the sequences obtained from the soils were able to be classified to a genus level (at a bootstrap value ≥ 80%), indicating that the bacterial genera in the soils are largely unexplored. Twenty-five genera of over 1% relative abundance were selected, and their relative abundance is depicted in Fig. 4(a). In the PCA plot, the first PC explained 79.6% of the total variance and positively correlated with *Methylobacter, Thiobacillus, Acidiferrobacter, Anaeromyxobacter*, and other unidentified genera (Fig. 4(b)). The PCA plot also
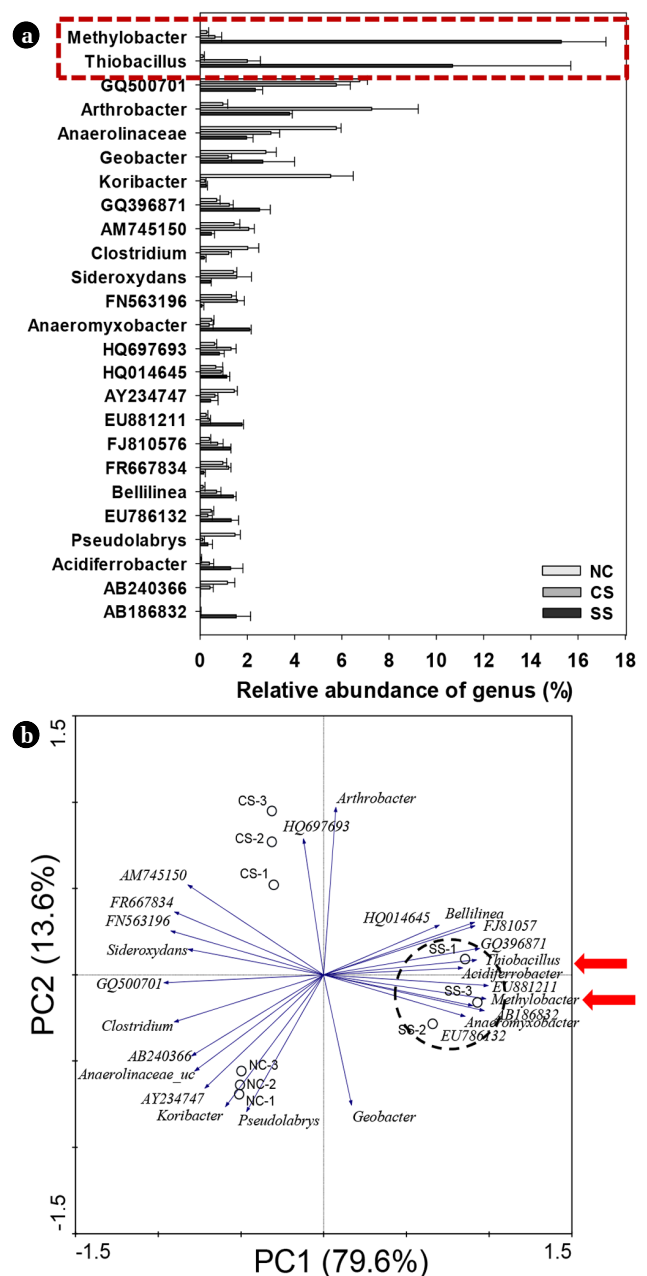

Fig. 4. Correlation between samples and the soil bacterial community at the genus level (a) Taxonomic composition of the bacterial communities. The bars represent averages, and error bars represent standard deviations of relative proportions. (b) PCA of relative abundances. The samples from stabilized soil are grouped with some genera (indicated by red arrows). The *Methylobacter* spp. and *Thiobacillus* spp. are correlated with stabilized layer soils (red arrow).

shows the grouping among the three SS samples and *Methylobacter* spp., *Thiobacillus* spp., and others (dashed circle in Fig. 4(b)). Similar to the patterns seen at the higher taxonomic levels, *Methylobacter* spp. and *Thiobacillus* spp. were more domi-

nant in SS samples (dashed square in Fig. 4(a)) than in the other soil types. The SS had the highest proportion of *Methylobacter* and *Thiobacillus* (15.3% and 10.7%, respectively, indicated by arrows in Fig. 4(b)) compared to two other samples (0.0% and 0.0% of *Methylobacter*, and 2.0% and 0.0% of *Thiobacillus* for CS and NC, respectively).

Species level: Microbial species distributions in the samples are presented in Fig. S4. Only 14-38% of the sequences were classified to a species level (at a bootstrap value $\geq$ 80%), indicating that the bacterial species in the soils are largely unexplored. Nineteen species with more than 1% relative abundance were selected, and their relative abundances are depicted in Fig. S4(a). In the PCA plot, the first PC explained 86.4% of the total variance and was positively correlated with *Methylobacter tundripaludum*, *Thiobacillus thioparus*, and other unidentified genera (Fig. S4(b)). The PCA plot also shows a loose grouping among the SS samples and *Methylobacter tundripaludum*, *Thiobacillus thioparus* and others (Fig. S4(b)). Similar to the pattern seen at the higher taxonomic levels, *Methylobacter tundripaludum* among *Methylobacter* spp., and *Thiobacillus thioparus* among *Thiobacillus* spp. were dominant in the SS samples (dashed square in Fig. S4(a)) compared to the other soil types. The SS has the highest proportion of *Methylobacter tundripaludum* and *Thiobacillus thioparus* (15.0% and 3.3%, respectively) compared to the two other soil types (0.3% and 0.5% of *Methylobacter tundripaludum,* and 0.0% and 0.4% of *Thiobacillus thioparus* for CS and NC, respectively). At the species level, *Methylobacter tundripaludum* had the highest relative abundance (627 sequences among total 4,178 sequences) in the microbial community (Fig. S4(a)).
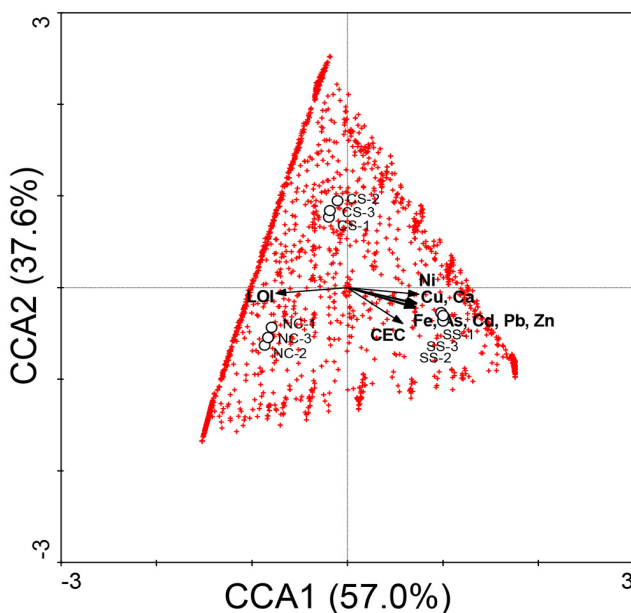
The result of the canonical correspondence analysis with Monte Carlo permutations is presented in Fig. 5. The first axis explained 57.0% of the total variance and the second axis explained 37.6% of the total variance. The first axis was positively correlated with heavy metals and CEC, but negatively correlated with LOI. The length of the environmental parameter arrows in the ordination plot indicates the strength of the relationship of that parameter to community composition. The result showed that the bacterial community structure was significantly affected by heavy metal concentration. Furthermore, the SS samples are positively correlated with heavy metals (the right side of Fig. 5). The CCA plot also distinguished the SS samples from the other soil types, and the heavy metals seem to be the major contributor to this differentiation.

Similarly, the SS samples are positively correlated with CEC and negatively correlated with LOI. CEC is a measure of how many cations (*e.g.*, heavy metals) are retained due to the negative charge of soil particles, and it is apparent that there is a positive correlation between heavy metals and CEC in these samples. On the other hand, the organic matter represented by LOI is

Fig. 5. Canonical correspondence analysis (CCA) based on relative abundances at the species level in conjunction with physicochemical properties of the samples. Red crosses indicate all species (sequences). Open circles indicate each of the triplicate samples of the three soil types. Black arrows indicate each physicochemical property of soils.

Fig. 6. Phylogenetic tree of 13 representative bacteria in the ɣ- proteobacteria class from the SS sample. *Methylobacter tundripaludum* is the dominant species (i.e., selectively grown) in the SS sample as indicated by the arrow.

often associated with hydrophobic contaminants such as DNAPLs, which may explain why the LOI was negatively correlated with the stabilized layer. However, these findings may be too preliminary for confident conclusions due to the low number of CEC and LOI measurements. The relationships seen in the CCA plot regarding CEC and LOI are probably derived from the higher CEC and the lower LOI of the SS samples compared to the other soil samples.

A phylogenetic analysis was also conducted to present the phylogenetic affiliation of the dominant species in the SS samples, which contains the substantial amount of heavy metals. The SS had the dominant species, *Methylobacter tundripaludum*, which belongs to the *γ-proteobacteria* class. The sequences of the 13 identified bacterial species of *γ-proteobacteria* in the SS were used to construct a phylogenetic tree (Fig. 6). The phylogenetic tree of the SS samples represents well the dominant bacterial populations of the heavy metal enriched environment, which are the methanotrophic microorganisms such as *Methylobacter tundripaludum*.

4. Discussion

The stabilized layer soil showed lower bacterial diversity and richness than that of the control soils. Since the stabilized layer soil contained higher concentrations of heavy metals, the lower diversity and richness can be attributed to the higher heavy metal concentrations. This is consistent with the findings by Chodak et al. [30], which reported a reduction in soil bacterial diversity due to heavy metal contamination. Sheik et al. [31] have also previously reported a substantial reduction in bacterial diversity in As and Cr contaminated soils. Therefore, the bacterial diversity index can function as a stability indicator and can be used to describe biological dynamics including stresses on soil bacterial communities [32].

To elucidate the detailed bacterial community profile of the stabilized soil, the pyrosequencing-based microbial analysis of the taxonomic hierarchy ranging from species to phyla. At the phyla level, a relatively higher percentage of Proteobacteria was observed in the stabilized layer soil compared to the cover layer soils. Furthermore, a relatively lower percentage of Acidobacteria and Actinobacteria were observed in stabilized layer soils than the cover layer soils. These observations are consistent with several prior studies in which heavy metal contamination of soils induced an increase in Proteobacteria and a decrease in Acidobacteria or Actinobacteria [19, 33]. Similarly, Sheik et al. [31] reported a prominent shift occurred towards Proteobacteria in As and Cr contaminated soils.

The major classes of bacteria found in the SS samples were Anaerolineae, β-proteobacteria, γ-proteobacteria, and α-proteobacteria. Several studies have reported higher abundances of Anaerolineae [34], α-proteobacteria, and γ-proteobacteria [35] at heavy metal contaminated sites. Within the γ-proteobacteria and β-proteobacteria, Methylococcales and *Thiobacillus* were the dominant orders, respectively. This is consistent with the observations of Chen et al. [36] where Methylobacterium and

Thiobacillus were the dominant members in mine tailings. Previously, Methylococcales abundance was also reported in an organic fertilizer, and it may also indicate contamination with hydrocarbons [37, 38].

At the genus level, *Thiobacillus* and *Methylobacter* were dominant in the stabilized layer soil. *Thiobacillus* spp. is known as one of the chemolithoautotrophs that are acidophilic and aerobic. It is known to have an ability to solubilize heavy metals by producing sulfuric acids and heavy metal complexing agents [39]. In addition, the bacteria that are most active in bioleaching belong to the *Thiobacillus* genus [40]. Therefore, *Thiobacillus* spp. have frequently been reported as a dominant species in the mining environment [37, 41]. The heavy metal-tolerant genera including *Thiobacillus* may have an ability to survive in highly contaminated soils. Therefore, the heavy metal-tolerant genera may be dominant in soils with high concentrations of heavy metals.

Likewise, *Methylobacter* spp. has been shown to have a significantly higher metal resorption ability [42], and hence, they can survive in highly contaminated soils. Among the *Methylobacter* spp., *Methylobacter tundripaludum* was the prominent species (Fig. 6). *Methylobacter tundripaludum* was first identified in arctic wetland soil in Norway [43]. *Methylobacter tundripaludum* is known as one of the methanotrophic, aerobic methane-oxidizing bacteria (MOB). MOB are a unique and important group of bacteria that act as natural filters, controlling the release of methane from anoxic soils [44]. More importantly, *Methylobacter* are known to have a high tolerance for heavy metals such as Zn, Cd, and Ni [45].

As shown in prior studies, methanobactin in methanotrophic bacteria can bind to a variety of metals and may play an important role in solubilizing metals *in situ* [46, 47]. Methanobactin appears to be responsible for copper trafficking [47-49]. A large amount of methanobactin can accumulate when methanotrophs grow under copper-limited conditions [48, 49]. Therefore, *Methylobacter tundripaludum* has the potential to tolerate heavy metals and to bind a variety of metals. Even though this species and its capabilities are largely unexplored, heavy metal binding ability and/or metal tolerance may allow *Methylobacter tundripaludum* (and its protein, methanobactin) to become a suitable bio-indicator of heavy metal contamination and/or site stabilization.

To employ *Methylobacter tundripaludum* as a bio-indicator of heavy metal contamination or stabilized soils, further studies are necessary to isolate the members of *Methylobacter* spp. that may have the capacity for heavy metal binding and tolerance. This is to be followed by elucidating their mechanisms for metabolizing heavy metals as well as co-existence.

5. Conclusions

The sites in proximity to exhausted mine areas are usually stabilized and used for general agricultural purposes. However, the immobilization of heavy metals by the stabilization process is temporary. Therefore, leaching will occur eventually, and this can lead to catastrophic contamination of agricultural produce

and groundwater. Therefore, it is useful to implement long-term monitoring of stabilized sites via periodic microbial analysis as well as with leaching tests. Simple diversity indices such as the Shannon index can be an appropriate microbiological tool. In addition, we have identified *Methylobacter* spp. or *Methylobacter tundripaludum* as a potential bio-indicator for assessing heavy metal contamination. To further verify the correlation of heavy metals contamination with the abundances of *Methylobacter* spp, more site studies are necessary.

Acknowledgments

References

1. Lee JS, Chon HT, Kim KW. Human risk assessment of As, Cd, Cu and Zn in the abandoned metal mine site. *Environ. Geochem. Health* 2005;27:185-191.

2. Melaku S, Dams L, Moens L. Determination of trace elements in agricultural soil samples by inductively coupled plasma-mass spectrometry: Microwave acid digestion versus aqua regia extraction. *Anal. Chim. Acta* 2005;543:117-123.

3. Guo G, Zhou Q, Ma LQ. Availability and assessment of fixing additives for the in situ remediation of heavy metal contaminated soils: A review. *Environ. Monit. Assess.* 2006;116:513-528.

4. Lee JS, Lee SW, Chon HT, Kim KW. Evaluation of human exposure to arsenic due to rice ingestion in the vicinity of abandoned Myungbong Au-Ag mine site, Korea. *J. Geochem. Explor.* 2008;96:231-235.

5. Kim J, Koo SW, Kim JY, et al. Influence of acid mine drainage on microbial communities in stream and groundwater samples at Guryong Mine, South Korea. *Environ. Geol.* 2009;58:1567-1574.

6. Brookes PC, McGrath SP. Effects of metal toxicity on the size of the soil microbial biomass. *J. Soil Sci.* 1984;35:341-346.

7. Jung MC, Thornton I. Heavy metal contamination of soils and plants in the vicinity of a lead-zinc mine, Korea. *Appl. Geochem.* 1996;11:53-59.

8. Ok YS, Lim JE, Moon DH. Stabilization of Pb and Cd contaminated soils and soil quality improvements using waste oyster shells. *Environ. Geochem. Health* 2011;33:83-91.

9. Razo I, Carrizales L, Castro J, Díaz-Barriga F, Monroy M. Arsenic and heavy metal pollution of soil, water and sediments in a semi-arid climate mining area in Mexico. *Water Air Soil Pollut.* 2004;152:129-152.

10. Kumpiene J, Lagerkvist A, Maurice C. Stabilization of As, Cr, Cu, Pb, and Zn in soil using amendments: A review. *Waste Manage.* 2008;28:215-225.

11. Ko MS, Kim JY, Bang S, Lee JS, Ko JI, Kim KW. Stabilization of the As-contaminated soil from the metal mining areas in Korea. *Environ. Geochem. Health* 2012;34:143-149.

12. Mamindy-Pajany Y, Hurel C, Geret F, Romeo M, Marmier N. Comparison of mineral-based amendments for ex-situ stabilization of trace elements (As, Cd, Cu, Mo, Ni, Zn) in marine dredged sediments: A pilot-scale experiment. *J. Hazard. Mater.* 2013;252:213-219.

13. MCIE. Mine pollution prevention plan (2007-2012). Release copy of MCIE, Korea. 2006.

14. MKE. Mine pollution prevention plan (2011-2016). Release copy of MKE, Korea, 2011.

15. U.S. EPA. SW-846 Test Method 9081: Cation-exchange capacity of soils (Sodium Acetate). 1986.

16. ASTM. ASTM D7348-08: Standard test methods for loss on ignition (LOI) of solid combustion residues. 2013.

17. KMOE. Standard method for soil pollution. Korea Ministry of Environment. 2002.

18. Chun J, Kim KY, Lee JH, Choi Y. The analysis of oral microbial communities of wild-type and toll-like receptor 2-deficient mice using a 454 GS FLX Titanium pyrosequencer. *BMC Microbiol.* 2010;10:101.

19. Hur M, Kim Y, Song HR, Kim JM, Choi YI, Yi H. Effect of genetically modified poplars on soil microbial communities during the phytoremediation of waste mine tailings. *Appl. Environ. Microbiol.* 2011;77:7611-7619.

20. Kim BS, Kim JN, Yoon SH, Chun J, Cerniglia CE. Impact of enrofloxacin on the human intestinal microbiota revealed by comparative molecular analysis. *Anaerobe* 2012;18:310-320.

21. Chao A. Nonparametric estimation of the number of classes in a community. *Scand. J. Stat.* 1984;11:265-270.

22. Shannon CE. The mathematical theory of communication. 1963, MD Comput, 14 (1997). p. 306-317.

23. Miranda J, Andrade E, Lopez-suarez A, Ledesma TR, Cahill A, Wakabayashi PH. A receptor model for atmospheric aerosols from a southwestern site in Mexico City. *Atmos. Environ.* 1996;30:3471-3479.

24. Diaz RV, Aldape F, Flores MJ. Identification of airborne particulate sources, of samples collected in Ticoman, Mexico, using PIXE and multivariate analysis. *Nucl. Instrum. Methods Phys. Res. Sect. B-Beam Interact. Mater. Atoms* 2002;189:249-253.

25. Gong M, Wu L, Bi XY, et al. Assessing heavy-metal contamination and sources by GIS-based approach and multivariate analysis of urban-rural topsoils in Wuhan, central China. *Environ. Geochem. Health* 2010;32:59-72.

26. Hudson RR, Boos DD, Kaplan NL. A statistical test for detecting population subdivision. *Mol. Biol. Evol.* 1992;9:138-151.

27. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X: Windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 1997;25:4876-4882.

28. Hall TA. Bioedit: A user-friendly biological sequence alignment editor and analysis. *Nucleic Acids Symp. Ser.* 1999;41:95-98.

29. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 2013;30:2725-2729.

30. Chodak M, Golebiewski M, Morawska-Ploskonka J, Kuduk K, Niklinska M. Diversity of microorganisms from forest soils differently polluted with heavy metals. *Appl. Soil Ecol.* 2013;64:7-14.

31. Sheik CS, Mitchell TW, Rizvi FZ, et al. Exposure of soil microbial

communities to chromium and arsenic alters their diversity and structure. *PLoS ONE* 2012;7:e40059.

32. Atlas SA, Kleinert HD, Camargo MJ, et al. Purification, sequencing and synthesis of natriuretic and vasoactive rat atrial peptide. *Nature* 1984;309:717-719.

33. Golebiewski M, Deja-Sikora E, Cichosz M, Tretyn A, Wrobel B. 16S rDNA pyrosequencing analysis of bacterial community in heavy metals polluted soils. *Microb. Ecol.* 2014;67:635-647.

34. Yin HQ, Niu JJ, Ren YH, et al. An integrated insight into the response of sedimentary microbial communities to heavy metal contamination. *Sci. Rep.* 2015;5:14266.

35. Zhu JY, Zhang JX, Li Q, et al. Phylogenetic analysis of bacterial community composition in sediment contaminated with multiple heavy metals from the Xiangjiang River in China. *Mar. Pollut. Bull.* 2013;70:134-139.

36. Chen LX, Li JT, Chen YT, et al. Shifts in microbial community composition and function in the acidification of a lead/zinc mine tailings. *Environ. Microbiol.* 2013;15:2431-2444.

37. Daquiado AR, Kuppusamy S, Kim SY, et al. Pyrosequencing analysis of bacterial community diversity in long-term fertilized paddy field soil. *Appl. Soil Ecol.* 2016;108:84-91.

38. Abbasian F, Palanisami T, Megharaj M, Naidu R, Lockington R, Ramadass K. Microbial diversity and hydrocarbon degrading gene capacity of a crude oil field soil as determined by metagenomics analysis. *Biotechnol. Prog.* 2016;32:638-648.

39. Nies DH. Microbial heavy-metal resistance. *Appl. Microbiol. Biotechnol.* 1999;51:730-750.

40. Bosecker K. Bioleaching: Metal solubilization by microorganisms. *FEMS Microbiol.* 1997;20:591-604.

41. Schippers A, Breuker A, Blazejak A, Bosecker K, Kock D, Wright TL. The biogeochemistry and microbiology of sulfidic mine waste and bioleaching dumps and heaps, and novel Fe(II)-oxidizing bacteria. *Hydrometallurgy* 2010;104:342-350.

42. Ondruschka J, Bley T. Biosorption of environmentally relevant heavy metals on selected biomaterials for wastewater treatment. *Eng. Life Sci.* 2003;3:215-218.

43. Wartiainen I, Hestnes AG, McDonald IR, Svenning MM. Methylobacter tundripaludum sp nov., a methane-oxidizing bacterium from Arctic wetland soil on the Svalbard islands, Norway (78 degrees N). *Int. J. Syst. Evol. Microbiol.* 2006;56:109-113.

44. Bussmann I, Rahalkar M, Schink B. Cultivation of methanotrophic bacteria in opposing gradients of methane and oxygen. *FEMS Microbiol. Ecol.* 2006;56:331-344.

45. Rajkumar M, Ae N, Freitas H. Endophytic bacteria and their potential to enhance heavy metal phytoextraction. *Chemosphere* 2009;77:153-160.

46. Choi DW, Antholine WE, Do YS, et al. Effect of methanobactin on the activity and electron paramagnetic resonance spectra of the membrane-associated methane monooxygenase in *Methylococcus capsulatus* Bath. *Microbiology* 2005;151:3417-3426.

47. Kim HJ, Graham DW, DiSpirito AA, et al. Methanobactin, a copper-acquisition compound from methane-oxidizing bacteria. *Science* 2004;305:1612-1615.

48. Balasubramanian R, Rosenzweig AC. Copper methanobactin: A molecule whose time has come. *Curr. Opin. Chem. Biol.* 2008;12:245-249.

49. Knapp CW, Fowle DA, Kulczycki E, Roberts JA, Graham DW. Methane monooxygenase gene expression mediated by methanobactin in the presence of mineral copper sources. *Proc. Natl. Acad. Sci. USA* 2007;104:12040-12045.