FEMS Microbiology Ecology, 91, 2015

doi: 10.1093/femsec/fiu003
Advance Access Publication Date: 5 December 2014
Research Article

RESEARCH ARTICLE

# Microbial community composition and functions are resilient to metal pollution along two forest soil gradients

Hamed Azarbad[1,*], Maria Niklińska[1], Ryszard Laskowski[1], Nico M. van Straalen[2], Cornelis A. M. van Gestel[2], Jizhong Zhou[3,4,5], Zhili He[3], Chongqing Wen[3] and Wilfred F. M. Röling[6]

[1]Institute of Environmental Sciences, Jagiellonian University, Gronostajowa 7, 30-387 Krakow, Poland, [2]Department of Ecological Science, Faculty of Earth and Life Sciences, VU University, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands, [3]Institute for Environmental Genomics and Departments of Microbiology and Plant Biology, University of Oklahoma, Norman, OK 73072, USA, [4]Earth Science Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA, [5]State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing 100084, China and [6]Department of Molecular Cell Physiology, Faculty of Earth and Life Sciences, VU University, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands

*Corresponding author. Institute of Environmental Sciences, Jagiellonian University, Gronostajowa 7, 30-387 Krakow, Poland. Tel: +48 (12) 664 5447; E-mail: Hamed.Azarbad@uj.edu.pl

**One Sentence Summary:** Using two distinct gradients of long-term heavy metal pollution, we found that both microbial community composition as well as functional gene-based structure was only minorly affected by the metal pollution, despite strong differences in metal toxicity along the gradients.

**Editor.** Dr Cindy Nakatsu

## ABSTRACT

Despite the global importance of forests, it is virtually unknown how their soil microbial communities adapt at the phylogenetic and functional level to long-term metal pollution. Studying 12 sites located along two distinct gradients of metal pollution in Southern Poland revealed that functional potential and diversity (assessed using GeoChip 4.2) were highly similar across the gradients despite drastically diverging metal contamination levels. Metal pollution level did, however, significantly impact bacterial community structure (as shown by MiSeq Illumina sequencing of 16S rRNA genes), but not bacterial taxon richness and community composition. Metal pollution caused changes in the relative abundance of specific bacterial taxa, including *Acidobacteria*, *Actinobacteria*, *Bacteroidetes*, *Chloroflexi*, *Firmicutes*, *Planctomycetes* and *Proteobacteria*. Also, a group of metal-resistance genes showed significant correlations with metal concentrations in soil. Our study showed that microbial communities are resilient to metal pollution; despite differences in community structure, no clear impact of metal pollution levels on overall functional diversity was observed. While screens of phylogenetic marker genes, such as 16S rRNA genes, provide only limited insight into resilience mechanisms, analysis of specific functional genes, e.g. involved in metal resistance, appears to be a more promising strategy.

**Key words:** Illumina sequencing; GeoChip; metal pollution; soil microbial communities

**Received:** 12 June 2014; **Accepted:** 22 October 2014

## INTRODUCTION

Because of its crucial importance for nutrient turnover and global carbon balance, the decomposition of dead organic matter has been the subject of numerous studies on the effects of pollution on ecosystems. It is already known for many years that litter decomposition slows down in metal-polluted forests (Rühling and Tyler 1973), and Babich and Stotzky (1974) suggested that the direct reason is metal toxicity to soil microorganisms. Indeed, cadmium toxicity to microorganisms was confirmed later by Giesy (1978). This negative effect of metals on carbon mineralization rates was later found in a number of studies (e.g. Grodziński *et al.*, 1990; Laskowski, Maryański and Niklińska 1994; Azarbad *et al.*, 2013). Although metal toxicity to soil microorganisms and its ultimate effect on organic matter breakdown have been confirmed, it remained unclear whether it results mostly from an overall decrease in microbial diversity. A reduced biodiversity may also lead to a decreased functional diversity, because certain groups or soil microorganisms specialize in decomposing particular organic substrates.

More recently, such changes in the structure of microbial communities from different environments have been confirmed [with the term 'microbial community structure', we refer to the composition of a microbial community (which species are present) and the abundances of its members]. Gough and Stahl (2011) observed a general restructuring of soil microbial communities chronically exposed to metal pollution in lake sediments. Hemme *et al.* (2010) reported that prolonged exposure to high concentrations of heavy metals, nitric acid and organic solvents (~50 years) has resulted in a massive decrease in species diversity as well as a significant loss of metabolic diversity in groundwater communities. They showed that the majority of microbial populations may have become extinct after the introduction of contaminants, while certain community members, with key metabolic activities related to denitrification and metal resistance, survived to form the foundation of the new community. Recently, Kang *et al.* (2013), using a functional gene microarray (GeoChip 2.0), studied sediment microbial communities along a gradient of metal contamination in Lake DePue (IL, USA), which has been contaminated for more than 80 years by an adjacent Zn-smelting facility. In contrast to Hemme *et al.* (2010), they showed that the functional potentials of microbial communities were similar across all sites, although some individual gene categories showed differences. A subset of sulfate reduction genes (*dsr*) and several metal-resistance genes (*cadA*, *chrAB* and *czcABC*) were highly correlated with metal contamination.

However, despite the global importance of forests, it is virtually unknown how their soil microbial communities adapt to long-term pollution at the functional level, such as the diversity in functional gene categories (e.g. metal-resistance genes), and how microbial community structure and functional potential relate to each other in metal-polluted forest soils. Knowledge on the linkage between function (traits) and phylogeny in long-term polluted forest soils would contribute to developing a framework for predicting the phylogenetic distribution of specific traits.

In this study, soil microbial communities in two long-term metal-polluted areas, Olkusz and Miasteczko Śląskie, Southern Poland, were investigated. Both areas have been seriously contaminated since the 1970s by two large zinc-and-lead smelters. In close proximity to the smelters, concentrations of some metals exceed their natural background levels by two orders of magnitude: the upper soil layers contain up to 4300 mg $kg^{-1}$ total Zn and 2900 mg $kg^{-1}$ total Pb, putting them among the most polluted soils in Europe (Nowicki 1993; Azarbad *et al.*, 2013). Our earlier study provided baseline knowledge on microbial community structure and general soil functioning in relation to soil characteristics, and demonstrated strong negative impacts of metals on soil respiration and microbial biomass (Azarbad *et al.*, 2013). However, the limited taxonomic resolution of the applied phospholipid fatty acid analyses did not allow us to identify the specific microbial groups that shift in abundance across the pollution gradients nor could we identify specific changes in key functional gene categories, connected with carbon (C), nitrogen (N) and sulfur (S) cycling and metal resistance.

The objective of this study was to reveal genetic details behind the differences in soil microbial communities observed earlier along the two metal pollution gradients. Studies along a gradient allow for better control over confounding non-pollution related factors, compared to more traditional approaches that contrast polluted vs unpolluted sites or comprise random sampling within an area exposed to a single pollution source (Hurlbert 1984). Previously, we observed that simultaneously studying these two gradients allowed for much stronger inference about effects of metal pollution (Azarbad *et al.*, 2013). We hypothesized that chronic metal pollution should be reflected in soil microbial communities, not only in changes of their taxonomic and functional diversity but also in metabolic potentials. In particular, we expected an increased frequency of metal-resistance genes with increasing concentrations of metals in soils. These goals were reached by analyzing microbial communities from the two gradients with Illumina MiSeq-based sequencing of 16S rRNA gene amplicons and GeoChip, a high-throughput functional gene array (He *et al.*, 2007).

## MATERIALS AND METHODS

### Site description and soil sampling

Two long-term polluted areas (Olkusz and Miasteczko Śląskie) were investigated (Table 1). The Olkusz area is the major zinc and lead industry region in Southern Poland. Mining in the area goes back to medieval times, and intensive modern industry started in 1967 when the 'Bolesław' zinc-and-lead smelter was established. Many years of mining and smelting have brought high metal concentrations in soils over a large area around the emission sources (Roberts, Scheinost and Sparks 2002). Miasteczko Śląskie is a small town located ca. 48 km NE from the 'Bolesław' smelter and home to another large zinc-and-lead smelter, constructed also in 1967. The intensive industry, especially in the 1970s when old-fashioned technological processes were used, resulted in heavy contamination of the environment with metals. The dominating tree species in the forests sampled in both areas are Scots pine (*Pinus sylvestris*) followed by Common beech (*Fagus sylvatica*) (Niklińska, Chodak and Laskowski 2005).

In the Olkusz area (O), a transect was established extending from 1.9 to 31.8 km away from the smelter, and in the Miasteczko Śląskie area (M) from 2.1 to 52.6 km away from the source of pollution (Table 1). Along each transect six sampling locations were chosen, all in *P. sylvestris* forest stands, and at each location a plot was laid out of approximately 100 $m^2$. In May 2011, from each plot 10 samples (each approximately 400–500 g wet mass) of the organic topsoil O layer (about 10 cm thick) were randomly taken with a 5-cm-diameter soil auger, and the top 10 cm of each sample was used for analysis, sieved (mesh size 2 mm), and mixed to obtain one representative sample per location. The soil samples were stored frozen (−20°C) until molecular analyses were performed.

**Table 1.** Geographical coordinates and distances from the smelters of the 12 locations along two pollution gradients in Olkusz (O1–O6) and Miasteczko Śląskie (M1–M6) southern Poland, with associated microbial and chemical characteristics of the top 10 cm soil organic layer.

| Site | Latitude (N) | Longitude (E) | Distance from the smelter (km) | $TI_{we}$[a] | $TI_{tot}$[a] | $pH_{water}$ | Number of sequenced 16S genes | Number of OTUs | Shannon Index (H) | Chao-1 richness | No. of detected functional genes |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **O1** | 50° 16′ | 19° 31′ | 1.9 | 142 | 43.5 | 5.81 | 19 087 | 2568 | 6.78 | 3219 | 35 809 |
| **O2** | 50° 17′ | 19° 29′ | 3.9 | 84.0 | 22.6 | 4.59 | 19 441 | 2309 | 6.57 | 3058 | 35 987 |
| **O3** | 50° 18′ | 19° 29′ | 5.3 | 52.3 | 10.6 | 4.97 | 27 671 | 2964 | 6.77 | 3694 | 39 468 |
| **O4** | 50° 19′ | 19° 30′ | 7.9 | 40.7 | 8.18 | 4.40 | 17 581 | 1970 | 6.45 | 2797 | 43 107 |
| **O5** | 50° 19′ | 19° 32′ | 19.6 | 29.8 | 4.84 | 5.10 | 23 111 | 2911 | 6.84 | 3650 | 40 250 |
| **O6** | 50° 25′ | 19° 37′ | 31.8 | 11.3 | 3.00 | 4.02 | 19 656 | 1571 | 5.99 | 2034 | 35 809 |
| **M1** | 50° 29′ | 18° 57′ | 2.1 | 232 | 41.3 | 4.59 | 21 990 | 2414 | 6.54 | 3175 | 42 896 |
| **M2** | 50° 31′ | 18° 56′ | 2.6 | 599 | 39.8 | 4.16 | 24 302 | 2508 | 6.66 | 3235 | 43 033 |
| **M3** | 50° 31′ | 18° 57′ | 3.3 | 219 | 20.2 | 4.31 | 26 573 | 2360 | 6.49 | 2827 | 42 345 |
| **M4** | 50° 32′ | 18° 57′ | 5.1 | 32.2 | 3.46 | 4.17 | 27 957 | 1824 | 6.03 | 2213 | 45 274 |
| **M5** | 50° 34′ | 19° 58′ | 8.7 | 16.5 | 3.62 | 3.91 | 27 994 | 1832 | 6.03 | 2227 | 37 781 |
| **M6** | 50° 32′ | 19° 39′ | 52.6 | 8.89 | 1.03 | 4.74 | 28 500 | 2471 | 6.50 | 3033 | 39 327 |

[a]toxicity indices calculated on the basis of total ($TI_{tot}$) and water-extractable ($TI_{we}$) concentrations of Cd, Zn, Cu and Pb, according to the equation $TI = \Sigma(C_i/EC50_i)$, where $C_i$ is the total concentration of metal i in soil (mg $kg^{-1}$ dry weight soil) and $EC50_i$ is the concentration of that metal causing a 50% reduction in dehydrogenase activity.

All physicochemical analyses were performed in triplicate on subsamples taken from the same pooled sample and presented results are the mean values with standard deviations.

## Soil physicochemical analyses

Complete soil analyses and detailed soil biological characterization of the sites are available in Azarbad *et al.* (2013). Water holding capacity was measured as described by Boivin (2005). The dry weight of the organic soil was measured after drying the samples for 12 h at 105°C, and organic matter content was determined as loss on ignition at 550°C (12 h). Soil pH was measured in water and 1 M KCl at 1:10 ratio (soil: liquid, w:v) using a digital pH meter (Nester Instr.). The total concentrations of metals (K, Ca, Mn, Na, Fe, Zn, Cd, Pb and Cu) were measured after wet digestion of 0.5 g subsamples in 10 ml concentrated $HNO_3$ (Merck, Darmstadt, Germany) with a gradual temperature increase from 50 to 150°C. Water-extractable concentrations of metals (K, Mn, Fe, Zn, Cd and Pb) were measured after 1 h shaking of soil samples in water at 1:10 (w: v) ratio and filtration of the extracts through a Whatman ffilter. Total and water-extractable metal concentrations were measured by using a flame or graphite furnace atomic absorption spectrophotometer (Perkin-Elmer, model AAnalyst 800, USA). At least three blanks accompanied every run of analysis. Additionally, each analytical run was checked for precision with a certified standard material (Chinese soil: NCS ZC 73001). The measured concentrations for the reference material were within 20% of the certified reference values, indicating accurate determination. Concentrations of the metals in the polluted soils were highly correlated to each other, and thus determination of the effects of any individual metal contaminant on soil microbial community structure is confounded. The complex metal pollution measured in field soils was summarized for each site by a toxicity index (TI) as defined by Stefanowicz, Niklińska and Laskowski (2008), for subsequent use in statistical analysis. This was done for water-extractable metal concentrations (leading to $TI_{we}$) and total metal concentrations ($TI_{tot}$). Briefly, TI was calculated as $\Sigma$(Ci/EC50i), where Ci is the concentration of metal i in soil and EC50i is the concentration of that metal causing 50% reduction in dehydrogenase activity as reported by Welp (1999). Characteristics of each of the soils collected, including total and soluble heavy metal levels, have been reported previously (Azarbad *et al.*, 2013). For the purpose of discussion, levels of heavy metals used to calculate the toxicity indices have been included in Table S1 (Supporting Information).

## DNA extraction

DNA was extracted from 0.25 g soil using the PowerSoil DNA kit (MoBio Laboratories, Solana Beach, CA) following instructions of the manufacturer. DNA quality was assessed by measuring UV absorbance at 230, 260 and 280 nm (respectively A230, A260 and A280) using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies Inc., Wilmington, DE). UV absorbance ratios A260/A280 and A260/A230 in the range of 1.7–2.0 indicate 'clean DNA'. DNA concentrations were measured with a PicoGreen method.

## Sample tagging, PCR amplification and sequencing

Primers 515F (5′-GTGCCAGCMGCCGCGG-3′) and 806R (5′-GGACTACHVGGGTWTC TAAT-3′) which target the hypervariable V4 region of the bacterial 16S rRNA gene were employed (Caporaso *et al.*, 2012). Both forward and reverse primers were

tagged with adapter, pad and linker sequences. A sample-specific barcode sequence (12 mer) was added to the reverse primer to allow for pooling multiple samples in one run of Illumina MiSeq sequencing. All primers were synthesized by Invitrogen (Carlsbad, CA, USA).

PCR amplification was performed in triplicate using a Gene Amp PCR-System® 9700 (Applied Biosystems, Foster City, CA, USA) in a total volume of 25 $\mu$l containing 2.5 $\mu$l 10 × PCR buffer and 0.5 unit of AccuPrime™ Taq DNA Polymerase High Fidelity (Invitrogen, Carlsbad, CA, USA), 0.4 $\mu$M of each primer and 10 ng template DNA.

Thermal cycling conditions were as follows: an initial denaturation at 94°C for 1 min, and 30 cycles at 94°C for 20 s, 53°C for 25 s and 68°C for 45 s, with a final extension at 68°C for 10 min. The triplicate PCR reactions were combined and quantified with PicoGreen. From each sample, 200 ng of PCR product was combined with other purified samples as a library, and re-quantified with PicoGreen. The mixture library was diluted, denatured to obtain 8 pM sample DNA library and mixed with an equal volume of 8 pM PhiX (Illumina, San Diego, CA, USA), according to the MiSeq™ Reagent Kit Preparation Guide (Illumina, San Diego, CA, USA). Finally, 600 $\mu$l of mixture library was loaded with read 1, read 2 and index sequencing primers (Caporaso *et al.*, 2012) on a 500-cycle (2 × 250 paired ends) kit, and run on a MiSeq at the Institute for Environmental Genomics of the University of Oklahoma. For sequencing data processing, firstly PhiX sequences were removed using BLAST (Altschul *et al.*, 1990) and sequences without assigned barcodes were discarded. To minimize the number of sequences with sequencing errors and ensure the quality of the sequences, both forward and reverse reads were trimmed based on sequence quality scores using Btrim (Kong 2011), where the sequences were trimmed if the average quality score of five continuous bases was less than 20. After trimming, the forward and reverse reads with 50 to 250 bp overlapping were combined to obtain longer sequences (250 ± 20 bp) using FLASH; reads that could not be joined together were removed. Also, unqualified sequences were removed if they were too short or contained undetermined base 'N'. Following this, potential chimeric sequences were detected and removed by UCHIME (Edgar *et al.*, 2011) using the Greengenes database for 16S rRNA genes (DeSantis *et al.*, 2006) as reference. Sequences were then clustered into operational taxonomic units (OTUs) at 97% sequence similarity using UCLUST (Edgar *et al.*, 2011). Finally, the RDP Classifier was used to assign 16S rRNA sequences to the bacterial taxa (Wang *et al.*, 2007).

Sequencing depth (number of sequences obtained for each sample) differed between samples (Table 1), and samples with the highest number of sequences were expected to show a comparatively higher diversity (Lundin *et al.*, 2012). Therefore, resampling was performed for each sample, with 17 581 sequences which corresponded to the smallest sampling effort in our datasets (sample O4; Table 1), as described by Galand *et al.* (2010). Sequence data have been deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under the accession number SRP028927.

## Functional gene array

GeoChip 4.2 was employed to assess metabolic potential of the bacterial communities from the study sites. It contains over 80 000 probes covering 140 000 coding sequences from 410 functional gene families related to microbial carbon (C), nitrogen (N), sulfur (S) and phosphorus (P) cycling, energy metabolism, antibiotic resistance, metal resistance/reduction, organic remediation, stress responses, bacteriophage and virulence (Tu *et al.*, 2014). Extracted DNA (1 $\mu$g) was labeled with Cy-3 using random primers and the Klenow fragment of DNA polymerase I. The labeled products were dried (45°C, 45 min; ThermoSavant) and rehydrated (with 2.68 $\mu$l sample tracking control to confirm sample identity). Hybridization solution (7.32 $\mu$l), containing 40% formamide, 25% SSC, 1% SDS, 2% Cy5-labeled common oligonucleotide reference standard (CORS) target, and 2.38% Cy3-labeled alignment oligo (NimbleGen) and 2.8% Cy5-labeled CORS target for data normalization and comparison (Liang *et al.*, 2010), was then added to the samples. Labeled DNA was hybridized to GeoChip 4.2 on a HS4800 Hybridization Station (MAUI, BioMicro Systems, Salt Lake City, UT, USA) at 43°C for 16 h. All hybridizations were carried out in triplicate on different modules. After hybridization, the arrays were scanned with a laser power of 100% and a photomultiplier tube gain of 100% (MS 200 Microarray Scanner, NimbleGen).

Scanned images were processed using ImaGene version 6.1 (BioDiscovery, El Segundo, CA) by averaging the intensities of every pixel inside the target region (segmentation method). The mean signal intensity was determined for each spot, and the local background signals were subtracted automatically from the hybridization signal of each spot. The spot signals, spot quality and background fluorescence intensities of scanned images were quantified with ImaGene. After this step all bad spots, which were flagged by the image processing software using predetermined criteria (defined as spots whose signals could not be accurately quantified due to their irregular shapes and/or contaminations), were removed and acceptable spots intensity data were extracted from ImaGene output files. Raw data obtained using ImaGene were uploaded to the laboratory's microarray data manager and analyzed using a GeoChip 4.2 data analysis pipeline (http://ieg.ou.edu/microarray). Spots with a signal-to-noise ratio [SNR = (signal intensity-background intensity)/standard deviation of the background] greater than 2 were used for further analysis (including potential empty spots and good spots).

Normalization was performed as described by He *et al.* (2007) and Liang *et al.* (2010). The hybridization signal was normalized by the mean signal intensity across all genes on the array. The across-array mean was calculated based on all intensities on the arrays after the removal of poor spots and outliers. Then, a ratio was calculated for each positive spot by dividing the signal intensity of the spot by the mean signal intensity to obtain the normalized ratio. The average signal intensities across all of the genes were expected to be approximately equal because the same amount of DNA was used for labeling and hybridization. A matrix was then generated from the normalized pixel intensities of all protein-encoding genes. Microarray data has been deposited in NCBI GEO and are accessible through GEO series accession number GSE59620.

## Denaturing Gradient Gel Electrophoresis (DGGE) profiling

DGGE analyses of specific groups (Fungi, Bacteria, *Acidobacteria*, *Actinobacteria*, *Alphaproteobacteria* and *Gammaproteobacteria*) were performed as described by de Boer *et al.* (2012) [see Table S2 (Supporting Information) for primers and amplification details].

## Statistical analysis

The number of OTUs in the subset was used for estimating total species richness as predicted by the Chao-1 estimator

(Colwell and Coddington 1994). The Shannon index (H) was calculated according to the formula: $H = \Sigma$ pi ln pi, where pi is the proportion of OTU i relative to the total number of OTUs.

Canonical correspondence analysis (CCA) was performed to link microbial communities to environmental variables. The relative abundances of OTUs, signal intensities of individual functional gene variants in Geochip or functional gene categories (summed intensities of all genes belonging to a certain functional gene category) (Zhou *et al.*, 2008) were used for CCA analysis. Selection for CCA modeling was conducted by an iterative procedure of eliminating redundant environmental variables based on significance level of the effect of each variable which was tested by a Monte Carlo permutation test (Ter Braak and Verdonschot 1995).

To estimate community similarity among samples, hierarchical cluster analysis was applied on the basis of the abundance of all OTUs observed using Bray–Curtis similarity in the program PAST (PAleontological STatistics) (Hammer, Harper and Ryan 2001) and the Jaccard coefficient which takes only presence/absence of OTUs into account (Jaccard 1912). As described in the result section, large species richness was observed and many OTUs that did not appear in all samples. Their absence might be due to either chance or be significant. In order to correct for absence by chance, we determined the total number $N$ of a particular OTU i required to ensure that each sample contains at least one count of that OTU i. A permutation approach was used, coded in Maple 9 (Maplesoft, Waterloo, Canada). $N$ counts were randomly divided over the 12 samples, to obtain a series of counts. This procedure was repeated $10^6$ times after which the number of series ($n$) was determined that contained at least one sample with zero counts. Probability $P$ was calculated as $n \times 10^{-6}$. The permutation approach was repeated with increasing values for $N$ until a $P < 0.00001$ was achieved (derived from the usual value $P = 0.05$, after correcting for multiple testing by dividing with the number of different OTUs observed ($5 \times 10^3$)). $N$ was found to be 165. Therefore, all OTUs with a total abundance of less than 165 were removed from the data set, and the trimmed data set was used to calculate Bray–Curtis and Jaccard similarities.

The Mantel test was used to test for significance ($P \leq 0.05$) of relationships between different microbial phyla (relative abundance of OTUs) and TI. A partial Mantel test, controlling for the confounding effect of pH, was applied to establish the correlation between microbial community composition (either based on presence/absence of OTUs or structure taking relative abundance of all OTUs into consideration) and TI. Correlations between relative abundances of bacterial phyla and TI were determined by using Spearman's rank correlation (with correction for multiple testing).

DGGE profiles were analyzed in the GelCompar II software package (Applied Maths, Kortrijk, Belgium) using a band assignment-independent method (Pearson product-moment correlation coefficient and unweighted pair-group clustering method using arithmetic averages). The correlation matrix generated was used for non-metric multidimensional scaling to visualize how microbial community structure changed in both gradients. Statistical analyses were performed using the vegan package in R 2.9, PAST (Hammer *et al.*, 2001) and Statgraphics Centurion XV software (StatPoint, Herndon, VA, USA).

# RESULTS

## Soil characteristics

Soil properties of the study sites were described in detail by Azarbad *et al.* (2013). In brief, the toxicity indices including the four main metals ($TI_{tot}$) at the most polluted sites in the Miasteczko Śląskie area, M1 and M2, were 41.3 and 39.8, respectively, and in the Olkusz area 43.5 (O1) and 22.6 (O2) (Table 1). For comparison, at the unpolluted sites M6 and O6, $TI_{tot}$ reached 1.0 and 3.0, respectively. Also $TI_{we}$ values were substantially elevated near the smelters: the difference between the most and least polluted sites in the Olkusz area was over 12-fold, and in the Miasteczko Śląskie area over 67-fold (Table 1). Soil $pH_{water}$ ranged from 3.91 (M5) to 5.81 (O1). Values of $pH_{water}$ were more or less similar at all study sites (acidic or highly acidic), except for O1 (pH 5.81) and O5 (pH 5.10), where pH was higher than at the other sites (Table 1).

## Microbial diversity and community structure in relation to metal contamination

After quality filtering the raw reads, we obtained 17 581 to 28 500 16S rRNA gene sequences per sample, with an average of 23 655. The numbers of different bacterial OTUs at the 97% similarity level ranged from 1571 to 2964 per sample with an average of 2308 OTUs (Table 1). Neither Shannon H' diversity nor Chao-1 richness indices were correlated with contamination level (TI) ($P > 0.05$).

Representative sequences of each OTU were classified into the domains bacteria (99.8% of the total data set; 20 phyla) and archaea (0.2%; 2 phyla). The major microbial phyla, i.e. those represented by ≥1% of the identified sequences, are presented in Table 2. Proteobacteria were the most dominant phylum observed, accounting for 38.2% of the sequences. Acidobacteria were the second most abundant phylum, at 10.6%.

Several multivariate analyses revealed a significant impact of $TI_{tot}$ on the microbial community structure. CCA indicated a significant relationship between the bacterial community structure and $TI_{tot}$ ($P = 0.037$). Besides pollution level, there were significant and stronger correlations of overall community composition with soil pH ($P = 0.001$) and concentrations of several nutrients (Na: $P = 0.018$, Fe: $P = 0.008$, Mn: $P = 0.004$) (Fig. 1). A partial Mantel test, controlling for the confounding effect of pH, showed that microbial community composition (based on presence/absence of OTUs only) and structure (taking relative abundance of all OTUs into consideration) significantly correlated with $TI_{we}$ and $TI_{tot}$ ($P = 0.008$ and $P = 0.004$, respectively).

To reveal how the 11 most abundant bacteria phyla were affected by soil properties, relative abundances of OTUs belonging to a particular phylum together with all soil properties were used in CCA (Table 2). TI significantly explained the community structures of *Actinobacteria*, *Bacteroidetes*, *Chloroflexi*, *Firmicutes*, *Planctomycetes* and *Proteobacteria* (Table 2). pH again appeared to be the most significant explanatory factor (Table 2). Partial Mantel tests, controlling for the confounding effect of pH, indicated nearly the same phyla as significantly affected by $TI_{tot}$ (*Acidobacteria*, *Actinobacteria*, *Bacteroidetes*, *Chloroflexi*, *Firmicutes*, *Planctomycetes* and *Proteobacteria*). $TI_{we}$ significantly affected the community structures of *Actinobacteria*, *Bacteroidetes*, *Chloroflexi*, *Firmicutes* and *Proteobacteria* (Table 2).

**Table 2.** Relationships between environmental factors and community structures of the most abundant bacterial phyla at two metal-polluted transects in southern Poland: Miasteczko Śląskie and Olkusz. CCA analysis and Mantel test were used to relate the variation in the relative abundances of OTUs belonging to particular phyla to metal toxicity and other environmental factors. Spearman correlation with Bonferroni correction was used to check relationship between pollution level and relative abundances of bacterial phyla. $TI_{tot}$ and $TI_{we}$ represent metal pollution levels based on total or water-extractable concentrations in the soils (see Table 1).

| Taxonomic group | Relative Abundance% | CCA analysis *P*-values | | | | | | | Mantel test *P*-values | | Spearman correlation test *P*-values | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $TI_{we}$ | $TI_{tot}$ | $pH_{water}$ | Na | Ca | Fe | Mn | $TI_{we}$ | $TI_{tot}$ | $TI_{we}$ | $TI_{tot}$ |
| *Proteobacteria* | 38.2 | - | 0.047 | 0.001 | 0.012 | 0.043 | 0.018 | 0.005 | 0.01 | 0.007 | 0.045 | 0.008 |
| *Acidobacteria* | 10.6 | - | 0.036 | 0.001 | 0.013 | - | 0.009 | 0.001 | - | 0.012 | - | - |
| *Planctomycetes* | 8.9 | - | 0.036 | 0.001 | 0.016 | - | 0.014 | 0.002 | - | 0.03 | - | - |
| *Actinobacteria* | 8.5 | - | 0.036 | 0.001 | 0.016 | 0.013 | 0.013 | 0.002 | 0.002 | 0.036 | - | - |
| *Verrucomicrobia* | 5.9 | - | - | 0.001 | 0.026 | - | 0.024 | 0.01 | - | - | 0.007 | 0.014 |
| *Bacteroidetes* | 5.7 | - | - | 0.001 | 0.02 | 0.042 | 0.023 | 0.003 | 0.005 | 0.004 | - | - |
| *Chlamydiae* | 2.0 | - | - | - | - | 0.044 | - | 0.001 | - | - | - | - |
| *Firmicutes* | 1.8 | - | 0.031 | 0.001 | 0.031 | - | 0.006 | - | 0.044 | 0.004 | - | - |
| *Armatimonadetes* | 1.3 | - | - | 0.001 | - | 0.049 | 0.026 | 0.013 | - | - | - | - |
| *Chloroflexi* | 1.1 | 0.027 | 0.016 | 0.001 | 0.032 | - | 0.017 | 0.001 | 0.005 | 0.004 | 0.013 | 0.002 |
| *Gemmatimonadetes* | 1.1 | - | - | 0.001 | - | - | - | 0.004 | - | - | 0.014 | 0.002 |

Abbreviations: —, no significant correlation




**Figure 1.** Ordination triplot based on CCA (the model significant at $P = 0.005$) of the relationship between the structure of soil microbial communities (relative abundances of OTUs) and environmental variables along two metal-polluted gradients in southern Poland: Miasteczko Śląskie (M1–M6) and Olkusz (O1–O6); see Table 1. Environmental variables were chosen based on significance levels calculated from individual CCA results in comparison with other variables ($TI_{tot}$: $P = 0.037$, pH: $P = 0.001$, Na: $P = 0.018$, Ca: $P = 0.045$, Fe: $P = 0.008$, Mn: $P = 0.004$). The percentage of total variance explained by each axis is shown next to the axis title.

Spearman correlation tests revealed that relative abundances of *Chloroflexi*, *Gemmatimonadetes* and *Proteobacteria* increased with increasing pollution level, while the relative abundance of *Verrucomicrobia* was negatively correlated with TI (Table 2). Frequently, pollution-related changes in the relative abundances of classes within other phyla were observed, while the relative abundances of the phyla itself were not affected by metal concentrations. Among members of the dominant phylum *Acidobacteria*, the *Acidobacteria* subgroup 1 correlated negatively with pollution ($TI_{we}$: $P = 0.05$, $TI_{tot}$: $P = 0.013$), while positive correlations with TI were found for the *Acidobacteria* subgroups 7 ($TI_{tot}$: $P = 0.02$), 10, 15, 16 and 17 ($TI_{we}$: $P = 0.04$, $TI_{tot}$: $P = 0.009$). Within the *Firmicutes*, relative abundances of *Clostridia* were negatively correlated to pollution level ($TI_{tot}$: $P = 0.013$).

When all OTUs and their relative abundances were taken into account, the average Bray–Curtis similarity between microbial communities was 52% (± a standard deviation of 12%). When the absence and presence of OTUs were taken into account, average Jaccard similarity between all samples was 41 ± 7%. These relatively low similarities may relate to the high species richness of the samples (on average 2308 OTUs), so the ~20 000 reads we obtained per sample may not be sufficient to quantitatively represent the community structure. Illumina sequencing only recovers a subset of the 16S rRNA genes in the DNA template used. Instead DGGE analysis will profile a PCR product containing billions of 16S rRNA gene amplicons derived from a DNA template, and would allow for better representation and comparison of the structures of complex communities. Indeed, when DGGE analysis of bacteria was done, a much higher average similarity was observed: 79 ± 11% (Fig. S1A, Supporting Information). Therefore, we corrected the OTU dataset by removing all OTUs that might be absent in one or more samples due to chance (see 'Material and methods' section for procedure). After this correction, the cluster analysis revealed a higher Bray–Curtis similarity of 61 ± 12% and, more importantly, a much higher average Jaccard

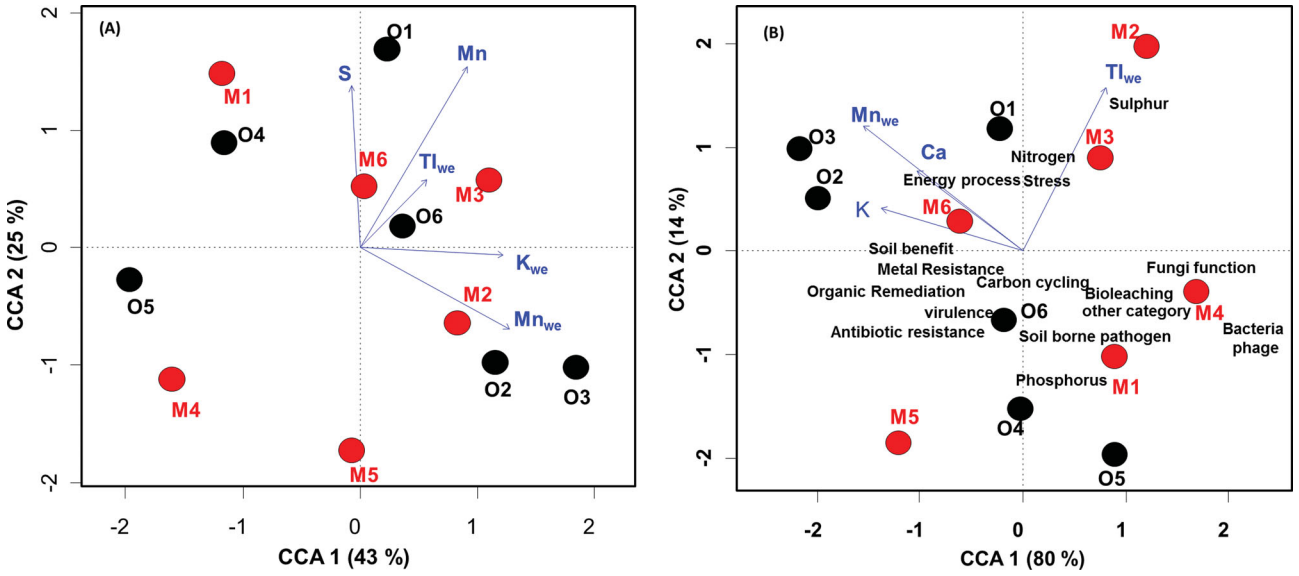


**Figure 2.** (**A**) Ordination triplot based on CCA (the model is significant at $P = 0.03$) of the relationship between GeoChip 4.2 hybridization signal intensities (circles) and environmental variables (arrows) along two metal-polluted transects in southern Poland: Miasteczko Śląskie (M1–M6) and Olkusz (O1–O6). Environmental variables were chosen based on significance levels calculated from individual CCA results in comparison with other variables (S: $P = 0.28$, $TI_{we}$: $P = 0.18$, K: $P = 0.18$, $Mn_{we}$: $P = 0.08$). Axis 1 explains 43% of variance, while Axis 2 describes additional 25% of the variance among samples. (**B**) CCA (the model significant at $P = 0.015$) for all functional gene categories (total number of positive genes in each category) (circles) and environmental variables (arrows) in the 12 study sites (red circles, Miasteczko Śląskie; black circles, Olkusz). Four environmental variables were chosen ($TI_{we}$, $P = 0.312$; $Mn_{we}$, $P = 0.02$; K, $P = 0.014$; Ca, $P = 0.131$). The percentage of variation explained by each CCA axis is shown.

similarity of 92 ± 5% among the 12 bacterial communities. This high Jaccard similarity suggests that the impact of pollution on community composition is only minor; we did not observe any OTU that solely occurred in either the heavily polluted (O1–O3, M1–M3) or less polluted soils (O4–O6, M4–M6) (Table S3, Supporting Information). The high Jaccard similarity is also in line with DGGE analysis of bacteria (Fig. S1A, Supporting Information). The high similarity observed for DGGE profiles of bacteria was also observed at higher phylogenetic resolution for *Gammaproteobacteria*, *Alphaproteobacteria* and *Actinobacteria* (Fig. S1B, C and D, Supporting Information), although considerable variation was observed in *Acidobacteria* communities (Fig. S1E, Supporting Information). The fungal community profiles revealed a high diversity within and between samples, but no significant relationships to pollution or location of the gradient were detected (Fig. S2, Supporting Information).

## Functional gene diversity and metabolic potentials in metal-impacted microbial communities

The number of detected functional genes ranged from 35 809 to 43 107 per sample with an average of 40 090 (Table 1). No relationships were observed between the total number of functional genes and the level of metal pollution as indicated by TI, suggesting that pollution level did not affect the overall functional gene diversity.

The average Bray–Curtis similarity between microbial communities was 80 ± 4%. There was no significant effect of metal pollution level on functional gene-based community structure when the individual signal intensities of all gene variants (individual genes) were used in CCA, as shown in Fig. 2A. However, when CCA was performed on distinct functional gene categories (based on summed intensities of all genes belonging to a particular functional gene category) a significant model was revealed ($P = 0.015$), explaining 94% of the total variance (Fig. 2B). Abundance of genes related to sulfur metabolism was positively affected by pollution level. Other functional gene categories, such as antibiotic resistance, virulence and soil-borne pathogen, were negatively affected by pollution level.

Subsequently, each functional gene category was examined separately, using signal intensities of individual genes belonging to that particular category. TI was only significant for metal-resistance genes ($P = 0.013$) (Fig. 31). GeoChip 4.2 contains 44 functional gene subgroups involved in metal resistance (e.g. genes that are involved in zinc, copper and lead resistance), accounting for 9% (9272 out of 95 847) of all probes. A total of 3931 to 4877 probes were positive for the 12 sampling locations investigated. These probes related to the resistance against a variety of metals, including Cd (corresponding to on average 23% of all probes relating to metal resistance), Cu (19%), Te (12%), Cr (12%), As (10%), Hg (9%), Zn (6%), Ag (6%), Al (1%), Pb (1%) and Ni (1%). Total signal intensity for these probes ranged from 3750 to 4785; however, no significant relationships between total metal-resistance gene intensities and pollution level were observed. Pollution level correlated positively with several genes, in particular *arsA*, *pbrT* and *cusF* (Fig. 31). Spearman correlation with Bonferroni correction corroborated this finding, indicating a positive relationship between pollution level and the genes *arsA* ($P = 0.0004$), *pbrT* ($P = 0.001$) and *cusF* ($P = 0.001$); however, no significant relation was observed with any other metal-resistance genes.

A Mantel test, comparing similarity matrices derived from sequencing data and GeoChip data, did not reveal a significant correlation between microbial community structure and functional potential gene diversity.

## DISCUSSION

Although microorganisms are major drivers of many soil processes, the relationship between microbial community

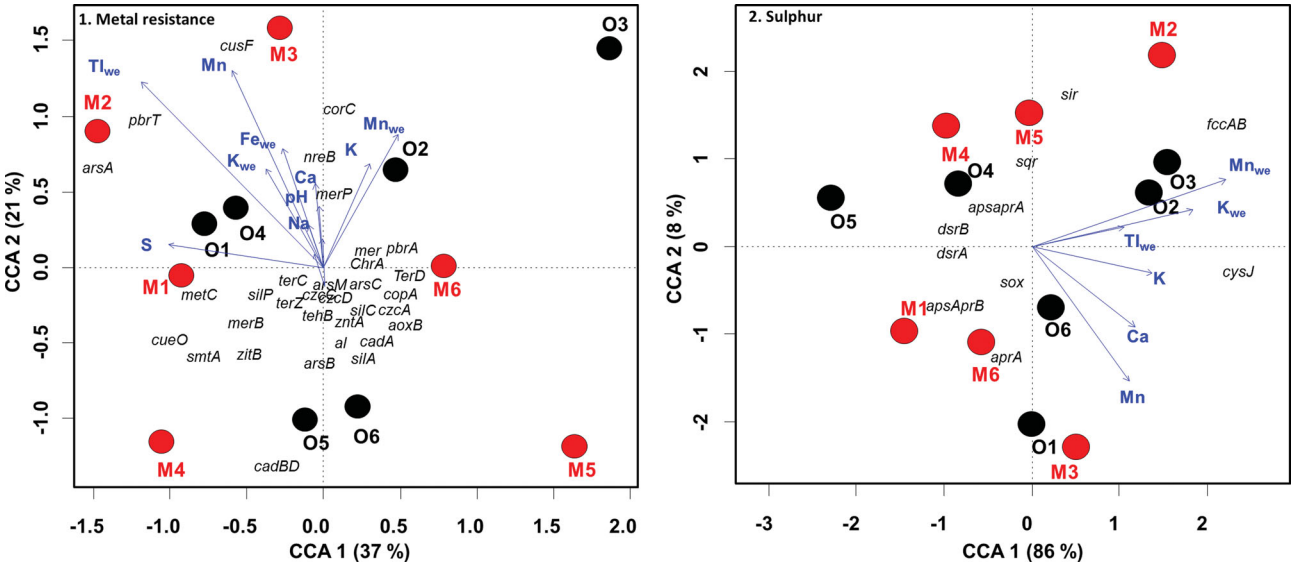


**Figure 3.** CCA of genes related to metal resistance (**1**; the model significant at $P = 0.013$, $TI_{we}$ $P = 0.013$) and sulfur metabolism (**2**; model $P = 0.17$, $Mn_{we}$ $P = 0.03$) obtained by using total normalized intensity of detected genes and soil properties along two metal-polluted transects in southern Poland: Miasteczko Śląskie (M1–M6, red circles) and Olkusz (O1–O6, black circles). The percentage of the total variation explained by each axis is shown.

composition and their functional potential in long-term polluted forest soils remains unclear. The response of soil microbial communities to metals in the natural field situation is in general confounded by soil properties. Our experimental design involving sampling along multiple gradients allows for better control over these confounding factors (Azarbad *et al.*, 2013). Both 16S rRNA gene sequencing and GeoChip data indicate that the soil microbial community had adapted to the elevated metal concentrations in the polluted soils as evidenced by changes in relative abundances of particular groups of microorganisms and increased frequency of certain metal-resistance genes (such as *arsA*, *pbrT* and *cusF*).

The remarkable high similarity in community composition along the two forest gradients suggests a great robustness of the soil bacterial communities. The findings are in agreement with the observations of de Boer *et al.* (2012), who studied the effects of nearly 27 years of Cu pollution on the microbial community structure (using DGGE) in agricultural fields. They showed that microbial community structure had stabilized after the shift caused by an artificial Cu amendment (<200 mg $kg^{-1}$). Adaptation of microbial communities to metal pollution has been suggested to be controlled by the diversity of metal-resistance genes (Lejon *et al.*, 2007). These genes can be located on plasmids (Monchy *et al.*, 2006) and might be exchanged between species leading to a community acquiring metal resistance but otherwise largely similar in composition to communities not exposed to metals (de Boer *et al.*, 2012). Using a high-resolution functional gene array, we observed that indeed differences were mainly in metal-resistance genes, and that other functional gene categories were not significantly affected, at least not to an extent that can be detected with the GeoChip.

While composition of the communities was rather similar, community structure was more variable along the two investigated gradients. It is worth noting that the sequencing suggested larger variation in community structure than indicated by DGGE analysis. Sequencing ~20 000 amplicons per sample may not be enough for a quantitative representation of the highly diverse microbial communities (with over 5000 different OTUs observed), impacting the comparison of community structures. Instead, DGGE profiled billions of amplicons per sample, and robust, whole curved-based analysis (Röling *et al.*, 2001) of these DGGE fingerprints suggested that the differences in communities were minor, even quantitatively (species and their abundances). This notion is also supported by qualitative comparison of the sequencing data after the removal of OTUs which might vary in their occurrence in the various samples simply due to chance. Methodological biases might also have dampened observing differences in community structure in relation to metal pollution. These are related to the low resolution of DGGE, which only detects the most abundant bacterial taxa, or the fact that we analyzed the structure of the total community, i.e. both the active and inactive members of the community are represented in the results.

## Effect of metal pollution on microbial diversity

Our results demonstrated that long-term pollution with metals (including Zn, Cd and Pb) does not reduce microbial diversity in forest soils. This conclusion is consistent with a study by Berg *et al.* (2012), who found that long-term copper exposure did not affect OTU richness at relatively crude taxonomic level, although it selected for copper-tolerant soil bacterial communities with changed composition. They used pyrosequencing of the 16S rRNA gene to elucidate the impacts of copper on bacterial community composition and diversity within a copper gradient (20–3537 mg $kg^{-1}$). Gough and Stahl (2011) investigated microbial community structure (by using terminal restriction fragment analysis of the 16S rRNA gene) in the metal-contaminated anoxic sediments of a eutrophic lake that were impacted for 80 years by nearby zinc-smelting activities. They showed that bacterial diversity and similarity remained high among microbial communities at the different sites along the gradient of metal contamination (Gough and Stahl 2011). In contrast, Sheik *et al.* (2012) observed a marked reduction in bacterial alpha diversity in soils contaminated for a long time with arsenic and chromium (three spatially isolated tanning facilities) compared to their paired control soils.

Addressing problems in studies relating biodiversity and ecosystem function, Bengtsson (1998) argued that for the management and development of sustainable ecosystems, it is probably more important to understand the linkages between key species or functional groups and ecosystem function, rather than focusing on species diversity. Therefore, we sought to describe the effects of metals at the community composition and phyla level in more details in the following section.

### Effect of metal pollution on microbial community structure

Although we did not detect a significant impact of TI on OTU richness, our results showed significant correlation between overall community structure and TI. Microbial communities shift in response to metal pollution by either selective growth or invasion of metal-resistant species (Turpeinen, Kairesalo and Häggblom 2004). In our study, potential population shifts were indicated by the correlation of relative abundances of some bacterial groups with pollution level (Table 2). Our results agree with earlier reports (Akob, Mills and Kostka 2007; Rastogi *et al.*, 2010) that *Proteobacteria*-related lineages constituted the most diverse group in the metal polluted sites.

The increased abundance of the phyla *Chloroflexi* and *Gemmatimonadetes* in metal-polluted sites, as found in our studies, is in line with the observations of Chodak *et al.* (2013). The presence of these groups suggests that they may play an important role in contaminated soils and that their presence may also provide a stabilizing element as they can be highly adapted to extreme environments (Spain, Krumholz and Elshahed 2009). A negative correlation between the relative abundance of *Verrucomicrobia* and pollution levels has been reported for copper-contaminated soils by Berg *et al.* (2012). However, as this phylum contains only few cultivable species, their physiological roles in natural environments remain unknown (Rastogi *et al.*, 2010).

### Effect of metal pollution on microbial functional potentials

Among functional gene categories, the effect of metal pollution was only significant for metal-resistance genes. As might be expected, metal contamination appeared to be the primary factor in the selection of metal-resistance genes since the functional genes of the less polluted soils were well separated from those of the metal-contaminated soils.

We related metal pollution to metal-resistance genes by using a TI that summarizes the concentrations of all encountered heavy metals (Stefanowicz *et al.*, 2008). This approach, instead of relating individual metals to individual resistance genes, was chosen for a number of reasons. Firstly, concentrations of most heavy metals were strongly correlated. Secondly, while concentrations of Cd, Pb and Zn were present at the highest concentrations in the investigated transects, and we thus would expect populations to be resistant to all these metals, very few probes for genes relating to resistance to these metals are currently present on the Geochip. Also other metals (e.g. Ni, As, Hg) can also be emitted by Zn-Pb smelters, but at much lower concentrations (Pasieczna and Lis 2008; Stefanowicz *et al.*, 2008). Therefore, we should also consider possible increase in resistance to these metals. Finally, while metal-resistance genes are generally named after the metal for which they were first described to provide resistance, there are hardly any metal-resistance genes that provide resistance against a single metal. Resistance systems often provide cross-resistance to several metals (Diaz-Ravina, Bååth and Frostegård 1994; Nies 2003).

A number of studies have used GeoChip to relate community functioning to metal pollution (Epelde *et al.*, 2010; Kang *et al.*, 2013; Singh *et al.*, 2014). The results of Epelde *et al.* (2010) and Singh *et al.* (2014) are not in agreement with our study. Likely reasons for the differences are types and concentrations of contaminating metals and environmental settings, e.g. the presence of metal hyperaccumulating plants (Epelde *et al.*, 2010). Our results are in agreement with Kang *et al.* (2013), who studied functional potential in sediments of Lake DePue (IL, USA), which has been contaminated for over 80 years by a Zinc smelter.

While stable functional composition was observed along the two metal gradients, we noted previously that total microbial biomass and functional potentials, i.e. substrate-induced respiration, were repressed at elevated metal concentrations. These declines may have increased the accumulation of organic matter in forest floor in polluted sites (Azarbad *et al.*, 2013).The work we report here supports the idea that shifts in the structure of the forest floor decomposer community have functional implications for C cycling and C storage in forest ecosystems but no effect on general functional potential gene categories.

### Linkages between taxonomy and functions

Knowing the composition of the microbial community alone does not necessarily lead to an understanding of its functional potential, as indicated by the absence of significant correlation between taxonomic and functional gene structure. We found that, despite the variance in phylogeny, functional potential appeared to be similar among microbial communities. Similar observations have been made for the human microbiome: despite the highly divergent compositions of gut microbiota communities across individuals, the functional gene profiles are alike (Turnbaugh *et al.*, 2009).

In an attempt to link the microbial community composition, structure and potential function along two metal pollution transects in forest soils, this study has shown that long-term exposure to varying levels of metal contamination influences abundance of several microbial populations while not resulting in large shifts in community composition nor in functional gene structure. Metal pollution was expected to exert a major stress on the community. However, since only abundances and diversity of predicted metal efflux proteins were significantly affected, acquiring metal resistance appears to be the defining force to adaptation. Our data also indicate that screens of phylogenetic marker genes, such as the 16S rRNA gene, and general functional genes may provide only limited insight into adaptation of microbial communities to metal contamination. Analysis of specific functional genes, e.g. involved in metal resistance, appears to be a more revealing strategy for monitoring changes in microbial communities.

## SUPPLEMENTARY DATA

Supplementary data is available at FEMSEC online.

## ACKNOWLEDGEMENTS

This study was performed within the 'Environmental stress, population viability and adaptation' project (No. MPD/2009-3/5) and supported by the DS759 of the Institute of Environmental Sciences, Jagiellonian University.

***Conflict of interest statement.*** None declared.

## REFERENCES


Akob DM, Mills HJ, Kostka JE. Metabolically active microbial communities in uranium-contaminated subsurface sediments. *FEMS Microbiol Ecol* 2007;**59**:95–107.

Altschul SF, Gish W, Miller W, *et al*. Basic local alignment search tool. *J Mol Evol* 1990;**215**, 403–10.

Azarbad H, Niklińska M, van Gestel CAM, *et al*. Microbial community structure and functioning along metal pollution gradients. *Environ Toxicol Chem* 2013;**32**:1992–2002.

Babich H, Stotzky G. Air pollution and microbial ecology. *Crit Rev Env Contr* 1974;**4**:353–421

Bengtsson J. Which species? What kind of diversity? Which ecosystem function? Some problems in studies of relations between biodiversity and ecosystem function. *Appl Soil Ecol* 1998;**10**:191–9.

Berg J, Brandt KK, Abu Al-Soud W, *et al*. Selection for Cu-tolerant bacterial communities with altered composition, but unaltered richness, via long-term Cu exposure. *Appl Environ Microb* 2012;**78**:7438–46.

Boivin MEY. Diversity of microbial communities in metal-polluted heterogeneous environments. *Ph.D. Thesis*. Vrije Universiteit Amsterdam, The Netherlands 2005.

Caporaso JG, Lauber CL, Walters WA, *et al*. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* 2012;**6**:1621–4.

Chodak M, Gołębiewski M, Morawska-Płoskonka J, *et al*. Diversity of microorganisms from forest soils differently polluted with heavy metals. *Appl Soil Ecol* 2013;**64**:7–14.

Colwell RK, Coddington JA. Estimating terrestrial biodiversity through extrapolation. *Philos T Roy Soc B* 1994;**345**: 101–18.

De Boer TE, Taş N, Braster M, *et al*. The influence of long-term copper contaminated agricultural soil at different pH levels on microbial communities and springtail transcriptional regulation. *Environ Sci Technol* 2012;**46**:60–8.

DeSantis TZ, Hugenholtz P, Larsen N, *et al*. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microb* 2006;**72**:5069–72.

Diaz-Ravina M, Bååth E, Frostegård A. Multiple heavy metal tolerance of soil bacterial communities and its measurement by a thymidine incorporation technique. *Appl Environ Microb* 1994;**60**:2238–47.

Edgar RC, Haas BJ, Clemente JC, *et al*. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 2011;**27**:2194–200.

Epelde L, Becerril JM, Barrutia O, *et al*. Interactions between plant and rhizosphere microbial communities in a metalliferous soil. *Environ Pollut* 2010;**158**:1576–83.

Galand PE, Potvin M, Casamayor EO, *et al*. Hydrography shapes bacterial biogeography of the deep Arctic Ocean. *ISME J* 2010;**4**:564–76.

Giesy JP. Cadmium inhibition of leaf decomposition in an aquatic microcosm. *Chemosphere* 1978;**6**:467–75.

Gough HL, Stahl DA. Microbial community structures in anoxic freshwater lake sediment along a metal contamination gradient. *ISME J* 2011;**5**:543–58.

Grodziński W, Greszta J, Laskowski R, *et al*. Effect of the chemical composition of industrial dusts on forest floor organic matter accumulation. *Water Air Soil Poll* 1990;**53**:169–78.

Hammer Ø, Harper DAT, Ryan PD. PAST: paleontological statistics software package for education and data analysis. *Palaeontol Electron* 2001;**4**:9.

He ZL, Gentry TJ, Schadt CW, *et al*. GeoChip: a comprehensive microarray for investigating biogeochemical, ecological and environmental processes. *ISME J* 2007;**1**:67–77.

Hemme CL, Deng Y, Gentry TJ, *et al*. Metagenomic insights into evolution of a heavy metal-contaminated groundwater microbial community. *ISME J* 2010;**4**:660–72

Hurlbert SH. Pseudoreplication and the design of ecological field experiments. *Ecol Monogr* 1984;**54**:187–211.

Jaccard P. The distribution of the flora in the alpine zone. *New Phytol* 1912;**11**:37–50.

Kang S, Van Nostrand JD, Gough HL, *et al*. Functional gene array-based analysis of microbial communities in heavy metals-contaminated lake sediments. *FEMS Microbiol Ecol* 2013;**86**:200–14.

Kong Y. Btrim: a fast, lightweight adapter and quality trimming program for next-generation sequencing technologies. *Genomics* 2011;**98**:152–3.

Laskowski R, Maryański M, Niklińska M. Effect of heavy metals and mineral nutrients on forest litter respiration rate. *Environ Pollut*. 1994;**84**:97–102.

Lejon DPH, Nowak V, Bouko S, *et al*. Fingerprinting and diversity of bacterial copA genes in response to soil types, soil organic status and copper contamination. *FEMS Microbiol Ecol* 2007;**61**:424–37.

Liang Y, He Z, Wu L, *et al*. Development of a common oligonucleotide reference standard for microarray data normalization and comparison across different microbial communities. *Appl Environ Microb* 2010;**76**:1088–94.

Lundin D, Severin I, Logue J, *et al*. Which sequencing depth is sufficient to describe patterns in bacterial $\alpha$- and $\beta$-diversity? *Environ Microbiol Rep* 2012;**4**:367–372.

Monchy S, Benotmane MA, Wattiez R, *et al*. Transcriptomic and proteomic analyses of the pMOL30-encoded copper resistance in *Cupriavidus metallidurans* strain CH34. *Microbiol* 2006;**152**:1765–76.

Nies DH. Efflux-mediated heavy metal resistance in prokaryotes. *FEMS Microbiol Rev* 2003;**27**:313–39.

Niklińska M, Chodak M, Laskowski R. Characterization of the forest humus microbial community in a heavy metal polluted area. *Soil Biol Biochem* 2005;**37**:2185–94.

Nowicki M. *Environment in Poland—Issues and solutions. Ministry of Environmental Protection, Natural Resources and Forestry*. Dordrecht, The Netherlands: Kluwer Academic, 1993.

Pasieczna A, Lis J. Environmental geochemical mapping of the Olkusz 1:25000 scale map sheet, Silesia–Cracow region, southern Poland. *Geochem-Explor Env A* 2008;**8**:323–31.

Rastogi G, Osman S, Vaishampayan PA, *et al*. Microbial diversity in uranium mining impacted soils as revealed by high-density 16S microarray and clone library. *Microbiol Ecol* 2010;**59**:94–108.

Roberts DR, Scheinost AC, Sparks DL. Zinc speciation in a smelter-contaminated soil profile using bulk and microspectroscopic techniques. *Environ Sci Technol* 2002;**36**:1742–50.

Röling WFM, van Breukelen BM, Braster M, *et al*. Relationships between microbial community structure and hydrochemistry in a landfill leachate-polluted aquifer. *Appl Environ Microb* 2001;**67**:4619–29.

Rühling Å, Tyler G. Heavy metal pollution and decomposition of spruce needle litter. *Oikos* 1973;**24**:402–16.

Singh BK, Quince C, Macdonald CA, *et al*. Loss of microbial diversity in soils is coincident with reductions in some specialized functions. *Environ Microbiol* 2014;**16**:2408–20.

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.

Spain AM, Krumholz LR, Elshahed MS. Abundance, composition, diversity and novelty of soil Proteobacteria. *ISME J* 2009;**3**:992–1000.

Stefanowicz AM, Niklińska M, Laskowski R. Metals affect soil bacterial and fungal functional diversity differently. *Environ Toxicol Chem* 2008;**27**:591–98.

Ter Braak CJF, Verdonschot PFM. Canonical correspondence analysis and related multivariate methods in aquatic ecology. *Aquat Ecol* 1995;**57**:255–89.

Tu Q, Yu H, He Z, *et al*. GeoChip 4: a functional gene arrays-based high throughput environmental technology for microbial community analysis. *Mol Ecol Resour* 2014;**14**:914–28.

Turnbaugh PJM, Hamady T, Yatsunenko BL, *et al*. A core gut microbiome in obese and lean twins. *Nature* 2009;**457**:480–4.

Turpeinen R, Kairesalo T, Häggblom MM. Microbial community structure and activity in arsenic-, chromium- and copper-contaminated soils. *FEMS Microbiol Ecol* 2004;**47**:39–50.

Wang Q, Garrity GM, Tiedje JM, *et al*. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microb* 2007;**73**:5261–7.

Welp G. Inhibitory effects of the total and water-soluble concentrations of nine different metals on the dehydrogenase activity of a loess soil. *Biol Fert Soils* 1999;**30**:132–9.

Zhou J, Kang S, Schadt CW, *et al*. Spatial scaling of functional gene diversity across various microbial taxa. *P Natl Acad Sci USA* 2008;**105**:7768–73.