

RESEARCH ARTICLE

Open Access



# Metagenomic analysis of microbial community and function involved in cd-contaminated soil

Gang Feng<sup>1†</sup>, Tian Xie<sup>1†</sup>, Xin Wang<sup>1</sup>, Jiuyuan Bai<sup>1</sup>, Lin Tang<sup>2</sup>, Hai Zhao<sup>3</sup>, Wei Wei<sup>1</sup>, Maolin Wang<sup>1</sup> and Yun Zhao<sup>1\*</sup>

## Abstract

**Background:** Soil contaminated with the heavy metal Cadmium (Cd) is a widespread problem in many parts of the world. Based on metagenomic analysis, we investigated the functional potential and structural diversity of the microbial community in Cd-contaminated and non-contaminated soil samples and we explored the associated metabolic pathway network in cluster of orthologous groups (COG) and Kyoto Encyclopedia of Genes and Genomes (KEGG).

**Results:** The results showed that microorganisms in these soils were quite abundant, and many of them possessed numerous physiological functions. However, Cd-contamination has the potential to reduce the microbial diversity and further alter the community structure in the soil. Notably, function analysis of the crucial microorganisms (e. g. *Proteobacteria*, *Sulfuricella* and *Thiobacillus*) indicated that these bacteria and their corresponding physiological functions were important for the community to cope with Cd pollution. The COG annotation demonstrated that the predominant category was the microbial metabolism cluster in both soil samples, while the relative abundance of metabolic genes was increased in the Cd-contaminated soil. The KEGG annotation results exhibited that the non-contaminated soil had more genes, pathways, modules, orthologies and enzymes involved in metabolic pathways of microbial communities than the Cd-contaminated soil. The relative abundance of some dominant KEGG pathways increased in the Cd contaminated soil, and they were mostly enriched to the metabolism, biosynthesis and degradation of amino acids, fatty acids and nucleotides, which was related to Cd tolerance of the microorganisms.

**Conclusions:** Cd-contamination can decrease the taxonomic species of microbes in soil and change the soil microbial composition. The functional pathways involved in the soil change with microbial structure variation, many of which are related to the heavy metal tolerance of soil microbes. The Cd-contaminated soil microbes is a potential resource for exploring cadmium resistant or tolerant bacteria.

**Keywords:** Metagenomic, Microbial community, Cadmium-contaminated, Functional potentials

## Background

Heavy metal contamination of soil derived from agricultural or industrial activities is one of the most severe environmental pollution issues around the world [1, 2]. During the past several decades, soil contaminated by heavy metals has become a serious phenomenon in China [3], which is

not easy to solve because heavy metals are non-degradable and get accumulated indefinitely in the environment [4]. In particular, the mining activities generate high concentrations of heavy metals (Ni, Cd, Pb, Mn, Cu, and Zn) that are exposed in open areas or agricultural soil. For example, high concentration of cadmium (1.8–7.6 mg/kg) has been found in paddy near the mining sites [5]. Moreover, it has been calculated that about 20% of China's arable lands are being contaminated by heavy metals, and this situation will deteriorated steadily in the near decades due to human activities, especially with respect to the agricultural soils [6]. Heavy metals gathering in soil would cause a series

\* Correspondence: zhaoyun@scu.edu.cn

†Equal contributors

<sup>1</sup>Key Laboratory of Bio-Resource and Eco-Environment of Ministry of Education, College of Life sciences, Sichuan University, No. 24 South Section 1, Yihuan Road, Chengdu 610065, China

Full list of author information is available at the end of the article



problems, including reducing the microbial biodiversity, soil quality and crops yield, and then threats animals and humans health via the food chain [7, 8].

It is clear that cadmium is not an essential nutrient element but a harmful factor to plant growth and development. Due to its good solubility in soil, cadmium can be absorbed by the crops planted in the soil and affect the crops' stomatal opening, transpiration, and photosynthesis [9]. In addition, the oxidative stress is involved in Cd toxicity have been proved in many researches, then increased the oxygen free radical production, or induced the enzymatic and nonenzymatic antioxidants [10–12], which could affect the normal development and the quality of the crops. Cd contamination has also been found to significantly decrease the microbial numbers and diversity in the soil [13, 14], leading to the reduction of soil microbial biomass, the respiration rate of microbes and enzyme activity [15–19], and decreased utilization of carbon [20].

Studies involved in soil microbial community composition could not only help us to explore the potential risks associated with contaminated soils but also provide insight into possible soil remediation strategies using indigenous resistant bacteria. It is well known that the some rhizosphere microorganisms can be used for phytoremediation [21, 22]. Furthermore, as the most active element of the soil ecosystem, microbes can rapidly respond to anthropogenic pressures, making it possible to be an indicator of soil quality and health [23, 24]. However, excess cadmium in soil has been shown to change the taxonomic diversity and function of the native microbial species [25–27]. For example, Hong and Si showed that varying concentrations of heavy metals surrounding an iron mining areas increased bacterial *alpha* diversity and resulted in a shift of the dominant genera [28]. Tonin et al. determined that *Arbuscular Mycorrhizal* (AM) fungi and plants associated with heavy metal tolerance can effectively enhance the heavy metal accumulation of plant roots [29].

To remedy the soil problem, the mechanism of metal resistance should be dissected as clear as possible. However, current knowledge about this is mostly based on cultured microorganisms, which results in less metal resistance determinants to be discovered. Because traditional microbiological techniques only isolate and culture a small proportion (0.1 to 1%) of the microorganisms in soil samples [30–33], therefore obtaining incomplete information about soil bacteria. However, some uncultivable bacteria may confer metal resistance, which was a vital environmental importance because of their potential application in bio-remediation area. Recently, efforts have been tried to identify genes from environmental samples via culture-independent methods, but they were amplified or detected because of their similarity to the previously

known genes [34–36], which is invalidate for exploiting novel mechanisms of metal resistance. As the advancement of cultivation-independent metagenomic approaches, it has been used to analysis the soil microbial community [37, 38], and improved our knowledge of the soil microbes and its potential significance. Moreover, the local soil microbial communities are mainly comprised of some dominant species and many other rare taxa [39]. Some of them with a low-abundance may be the novel microbial lineages [40, 41], and could play a crucial role in biogeochemical interactions of the soil-plant system [42–44]. Thus, the wealth of information obtained from the deep metagenomic sequencing makes it possible to capture the genomic information of low-abundance populations and to reveal the multiple activities in soil. It has been successfully introduced into investigating many diverse microbial niches in marine water, grassland soil, human gut, and artificial systems. Additionally, recent attempts have been made to analyze the abundance of bacteria and the associated genes involved in copper resistance from agricultural soils via metagenomic sequencing [45]. It is a useful tool to help us understand the unique characteristics of bacteria and genes associated with heavy metal contaminated soil and to aid in developing novel biological methods that could remedy soil ecosystem damage result from heavy metals.

The environmental troubles in mining areas are mainly related to physical disturbances surrounding spilled mine tailings, emitted dust and metal-containing waste, which results in toxic heavy metals enriched in soils [46]. Thus, microbial communities in these soils are optimal for the discovery of new determinants of metal resistance. Therefore, samples in this study were collected near phosphate rock chemical plants, Sichuan province, China.

In this study, we used metagenomic to analysis the microbial community structure and function in Cd-contaminated soil and reveal the variations of the microbial community, which provided an ecological methodology for investigating potential heavy metal tolerant or resistant bacteria that are able to cope with Cd-polluted soil.

## Methods

### Sample collection

The sample plot was near phosphate rock chemical plants in Shuangsheng town, Shifang County, Sichuan province, China. Two sites, designated site 1 ( $S_1$ ) and site 2 ( $S_2$ ), were sampled in triplicate from idle soils and without plants, on 7th May, 2015. Our samples were obtained from private land and our sampling was permitted by the land owner.  $S_1$  consisted of unpolluted soil and  $S_2$  consisted of Cd-contaminated soil collected next to the sewage outlet of three phosphate rock chemical plants. Samples were collected by removing the topsoil

(0 to 5 cm), gathering approximately 500 g of soil at a depth of 5 to 10 cm and then immediately transferring the samples into sterile plastic bags that were kept on ice. Upon returning to the laboratory, half of the soils were immediately frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$  for subsequent molecular analysis. The remaining half was air-dried at room temperature for one week and subsequently sieved through a 100-mesh sifter to remove stones and visible plant fragments. The soils were then stored at  $4^{\circ}\text{C}$  for further chemical analysis.

### Soil physicochemical analysis

In order to assess the soil quality and the microbial community structure, we closely followed the methods reported by Bloem et al. [47] that suggests indicators for soil quality assessment and microbial analysis.

Physicochemical property of the soil samples was detected at the beginning of the experiment in order to have a baseline for soil characterization. The following metrics were evaluated: soil organic matter (SOM), pH, total phosphorous (TP), total nitrogen (TN), total potassium (TK) and Cd. SOM and TN concentrations were determined using the method described by Sun et al. [48]. We measured the soil pH using an INESA pH meter (Shanghai REX Instrument Factory, Shanghai, China) in a 1:5 suspension of ultrapure water. The TP content was measured using the ammonium molybdate spectrometry [49]. TK content was evaluated via the method in *Soil agricultural chemistry analysis* [50].

All Cd concentrations were determined using inductively coupled plasma mass spectrometry (ICP-MS), digestion of the soil samples with 5 mL HCL at  $150^{\circ}\text{C}$  for 60 min, 5 mL  $\text{HNO}_3$  at  $150^{\circ}\text{C}$  for 60 min, 3 mL HF at  $150^{\circ}\text{C}$  for 10 min, and 3 mL  $\text{HClO}_4$  at  $190^{\circ}\text{C}$  for 100 min in a graphite digestion apparatus. All mixed soil samples were tested three times; and results are reported as the mean of the data.

### DNA extraction, library construction, and metagenomic sequencing

DNA for metagenomic analysis was extracted from the soil samples using the Power Soil DNA kit (MoBio Laboratories, Solana Beach, CA), according to the manufacturer's protocols. The concentration and quality ( $A_{260}/A_{280}$ ) of extracted DNA were determined using a NanoDrop ND-2000 spectrophotometer (NanoDrop, Wilmington, DE, USA), and evaluated with a 2% agarose gel. To minimize DNA extraction bias, three replicate DNA isolation of each sample were pooled. Afterwards, DNA was sheared into fragments of approximately 300 bp using an M220 Focused-ultrasonicator (Covaris Inc., Woburn, MA, USA) to build a paired-end library. DNA templates were then pretreated using the TruSeqTM Kit according to the manufacturer's instructions

(<http://www.illumina.com/>). Following libraries were pooled and loaded onto an Illumina cBot [51]. Pair-end sequencing ( $2 \times 150$  bp) was completed at Majorbio Bio-Pharm Technology Co., Ltd. (Shanghai, China) according to the standard protocol (<http://www.illumina.com/>).

### Selection and assembly of sequencing reads

The 3'-adaptors and 5'-adaptors was stripped. The raw reads were trimmed with a minimum quality score 20 and a minimum length 50 bp to maintain the reliability of the data [52]. Clean reads were then assembled using SOAPdenovo software [52] was used to assembled the clean reads at a Kmer range of 39 to 47. Statistical tests were conducted based on the scaffolds with a length over 500 bp [52]. Scaffolds with a length over 500 bp were then extracted and broken into contigs without gaps [52]. Contigs were used for the following predictions and annotations. The statistical information for the contigs was exhibited in Table 2. To confirm the assembled metagenomic data, all the sequencing reads were mapped to contigs using the Burrows-Wheeler Aligner (<http://bio-bwa.sourceforge.net/>) BWA tool [53]. Pread\_-contig were calculated as the number of reads successfully aligned to the assembled contigs [52], divided by the total number of clean reads.

### Gene prediction, taxonomy and functional assignment

The open reading frames (ORFs) of contigs and annotations in each sample were predicted with MetaGene Annotator [53]. The predicted ORFs with lengths over 100 bp were retained and translated into amino acid sequences via NCBI ORF finder [53], which were subsequently annotated by BLASTP (BLAST Version 2.2.28+) against NCBI-nr database including SwissProt. The optimized E-value threshold for BLAST Alignment was  $1e-5$ .

In this study, the taxonomy of the metagenomes was assigned by the small-subunit (SSU) rRNA gene tags and predicted proteins [54]. We obtained the SSU rRNA gene tags by align the raw reads with the Silva SSU rRNA database (<https://www.arb-silva.de/>) with an optimized E-value threshold of  $1e-5$ , alignment length of 100 bp, and identification cutoff 80%. The taxonomic annotation of the samples were obtained by blast against NCBI-nt database. For each protein-coding read, alignment with protein sequences in the NCBI-nr database was performed using the BLAST tool, and results with the best hit to the NCBI microbial taxonomy database were obtained .

The ORFs were aligned using the eggNOG database (evolutionary genealogy of genes: Non-supervised Orthologous Groups, Version 4.0) via BLASTP (BLAST Version 2.2.28+) at the optimized E-value threshold of  $1e-5$  [54], and the corresponding cluster of orthologous groups of protein (COG) were obtained. A pair-wise

statistical comparison of COG classification was carried out using STAMT [54], and the relevance between samples was evaluated using the Mann-Whitney rank sum test in SigmaPlot (Version 12.5, Systat Software, USA). The KEGG pathway annotation was conducted using BLAST search (BLAST Version 2.2.28+) against the Kyoto Encyclopedia of Genes and Genomes database (Kyoto Encyclopedia of Genes and Genomes, KEGG GENES,) at the optimized E-value threshold of  $1e-5$ . The functional annotation of genes was conducted via KOBAS 2.0 (KEGG Orthology Based Annotation System) [53]. Hierarchical clustering analysis of the KEGG annotation results among different samples was carried out using Qiime (version 1.17) and R Programming Language [54].

## Results

### Soil physicochemical background

The initial soil characteristics and Cd concentrations were determined for each site (Table 1). The results showed that the pH value, organic matter, total nitrogen, total phosphorus and total potassium content in the two samples coincide with the local large area soil. The concentrations of Cd in the non-contaminated soil ( $S_1$ ) and the contaminated soil ( $S_2$ ) were 0.072 and 1.639 mg/kg, respectively. There were no significant differences in all soil characteristics between the two samples except for the Cd concentration (Table 1).

### Overview of assembly

We obtained 15.8 GB of clean reads from the metagenomic sequencing. After cleavage into 47 bp-Kmers, we subsequently gained 122,977 and 40,775 contigs (> 500 bp) from the scaffolds of  $S_1$  and  $S_2$ , respectively. Statistical information of contigs (Table 2) indicated that these assemblies were sufficient for our study [55]. We obtained 204,739 and 68,701 ORFs via MetaGene Annotator (<http://metagene.cb.k.u-tokyo.ac.jp/>) from the contigs of  $S_1$  and  $S_2$ , respectively. All ORFs with a length over 100 bp were subsequently translated into amino acids.  $S_1$  had more ORFs when compared to  $S_2$ , which consisted with the distribution of contigs.

### COG annotation and analysis

The taxonomic assignments of the two samples were obtained by blast against NCBI-nt database. Notably, 39.8% of the predicted ORFs identified in the two samples can be assigned to putative functions. The classification of

**Table 2** Summary of the metagenomic sequencing

Category	$S_1$	$S_2$
Raw reads	100,895,940	94,080,764
Clean reads	96,712,790	90,359,739
Assembled contigs	122,977	40,775
Largest contig length (bp)	47,750	144,494
Contig_N50 length <sup>a</sup> (bp)	1014	949
Contig_N90 length <sup>b</sup> (bp)	553	548
Predicted ORFs	204,739	68,701

<sup>a</sup>Contig\_N50 length, the length of the smallest contig in the set of largest contigs that have a combined length that represents at least 50% of the assembly [74]

<sup>b</sup>Contig\_N90 length, the length of the smallest contig in the set of largest contigs that have a combined length that represents at least 90% of the assembly [74]

the functional potential genes was subsequently conducted via COG analysis. Amino acid transport and metabolism (E), energy production and conversion (C), general function prediction only (R) and function unknown (S) were the dominant functions among the 25 categories (Fig. 1). We summarized the results into four categories: information storage and processing (cluster I); cellular processes and signaling (cluster II); metabolism (cluster III); and poorly characterized function (cluster IV). Notably, cluster III was predominant in both samples and related to growth of the microbial community. However, the relative abundance of these metabolic genes in  $S_1$  was observably lower than in  $S_2$ . Cluster IV was the next most abundant in both samples, followed by cluster II and I. In  $S_2$ , the relative abundance of cluster III was higher than that of  $S_1$ . The percentage of reads assigned to cell motility, cell cycle control, cell division, and chromosome partitioning were the lowest ( $\leq 1.0\%$ ) among the two samples. These sequencing data indicated that Cd-contamination inspired the metabolism of the microorganisms, as demonstrated by increase in relative abundance of metabolic genes detected in the Cd-polluted soil.

### Taxonomic composition of the microbial communities

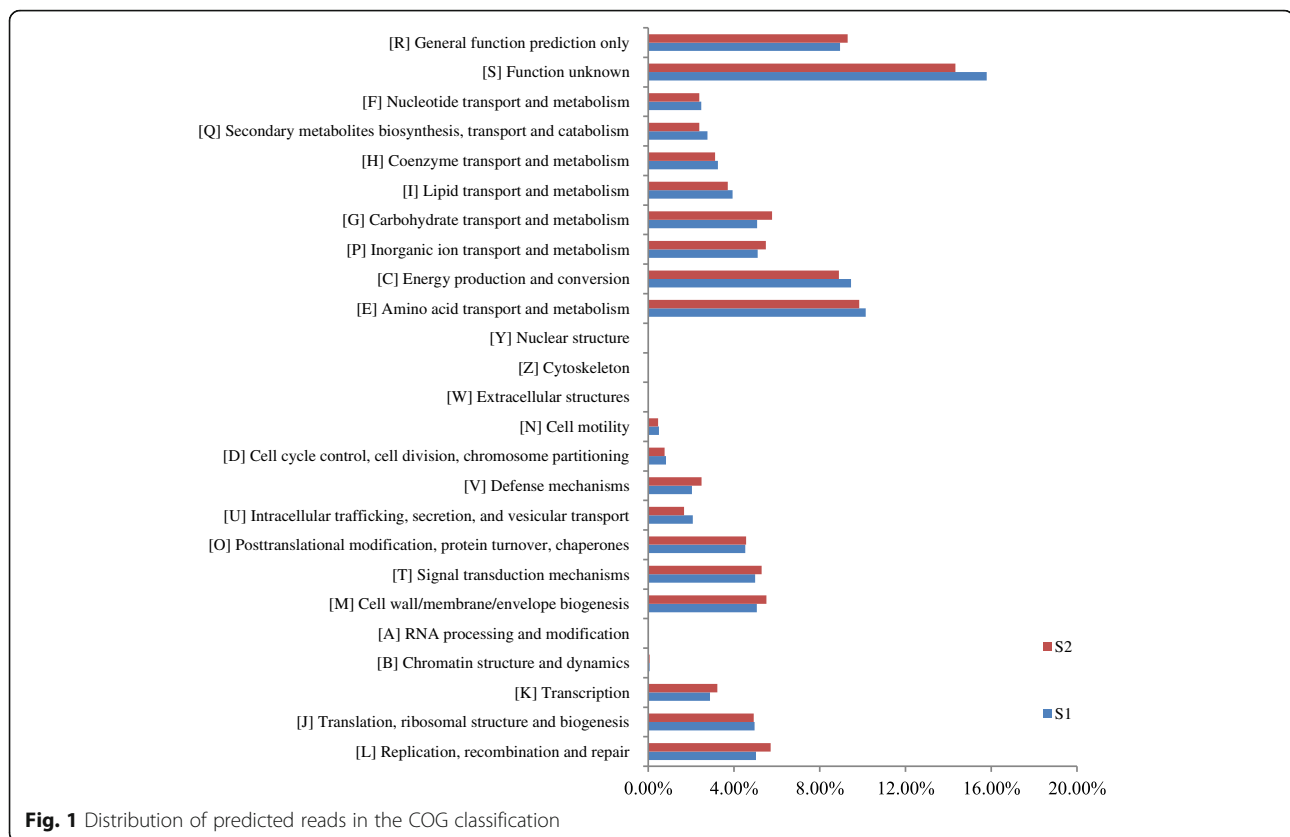
The predicted ORFs with a length over 100 bp was selected and translated into amino acid sequence via the NCBI ORF finder, then annotated by BLASTP search (BLAST Version 2.2.28+) against NCBI-nr database [56] at the optimized E-value threshold of  $1e-5$ . According to the corresponding taxonomic information in NCBI-nr database, we could obtain the taxonomy annotation and

**Table 1** Physical and chemical properties of soil

Sample	pH	Organic Matter (g/kg)	Total nitrogen (g/kg)	Total phosphorus (g/kg)	Total potassium (g/kg)	Cd concentration (mg/kg)
$S_1$	6.66 ± 0.039a	21.07 ± 1.949a	2.13 ± 0.254a	3.62 ± 0.172a	17.83 ± 0.331a	0.072 ± 0.003a
$S_2$	6.72 ± 0.024a	18.94 ± 2.589a	1.82 ± 0.176a	3.94 ± 0.030a	17.83 ± 0.324a	1.639 ± 0.225b

The different letter a and b behind the data means significant ( $p < 0.05$ ) differences between the groups



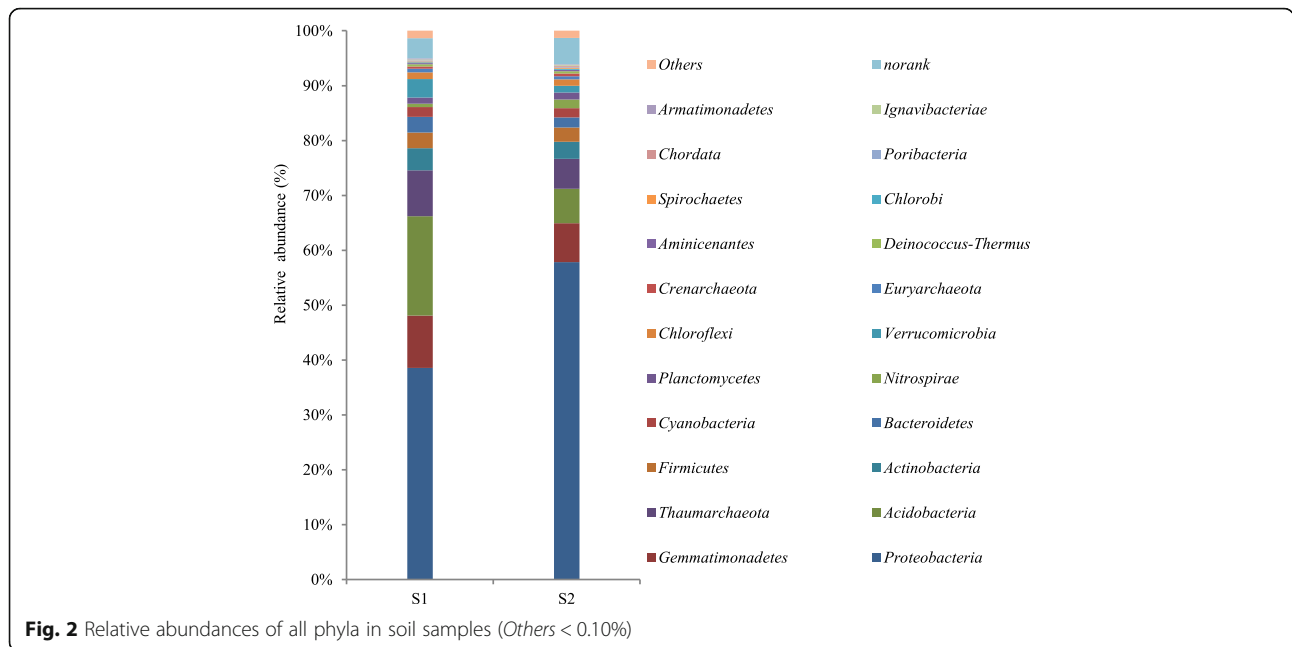


abundance of the species derived from the two samples. We assigned the classification confidence threshold in a lineage without classified information in the database as “*unclassified*”. We assigned “*norank*” to all sequences that could not be exactly classified into any known group at a taxonomic level. We classified the microorganisms with a relative abundance lower than 0.10% as “*Others*” (Fig. 2).

At phylum level, we summarized 77 taxa (apart from the *norank*) were from the two soil samples. According to the annotation, we defined *Proteobacteria*, *Gemmatimonadetes*, *Thaumarchaeota* and *Acidobacteria* as the dominant phyla in both  $S_1$  and  $S_2$ , accounting for over 75% of the total population (Fig. 2). *Proteobacteria* was the most abundant phylum in both samples, comprising of 38.56% of all phyla in  $S_1$  and 57.85% in  $S_2$ . These bacteria are important in soil, providing some basic functions related to the biogeochemical cycle. The *Proteobacteria* in the two samples embraced *Alphaproteobacteria*, *Betaproteobacteria*, *Gammaproteobacteria*, *Deltaproteobacteria* and *Epsilonproteobacteria*. Among them, *Betaproteobacteria* was the most abundant in the two soil samples. *Acidobacteria* was the second abundant phylum in  $S_1$  (18.13%), but was less abundant in  $S_2$  (6.30%). The high abundance of *Proteobacteria*, *Gemmatimonadetes*, *Thaumarchaeota* and *Acidobacteria*

indicated that these phyla played key roles in the soil bacterial community.

At the genus level, we obtained a total of 1736 genera (apart from the *norank*) from the two samples. Among of the whole genera, there were 1331 genera exist in both samples. There were 174 genera (apart from the *norank*) with a relative abundance greater than 0.10% of the total microorganisms in at least one of both samples. Figure 3 demonstrated the primary genera that each accounted for over 1.00% of the population in one of the two samples. The primary genera of  $S_1$  ( $n = 11$ ) and  $S_2$  (14) make up close 30% of the whole soil community. *Candidatus Koribacter* (6.44%), *Candidatus Solibacter* (3.11%) and *Bradyrhizobium* (3.07%) were the three most abundant genera in  $S_1$ . The taxonomic analysis showed that the bacteria belonged to *Nitrososphaera*, *Nitrosospira*, *Sulfuricella*, *Gemmatimonas*, *Candidatus Solibacter*, *Burkholderia*, *Bradyrhizobium*, *Candidatus Accumulibacter*, *Nitrospira*, *Candidatus Entotheonella*, *Cupriavidus*, *Azoarcus*, *Thiobacillus* and *Sideroxydans*, respectively. These bacteria accounted for 1.04% to 3.66% of the total community in  $S_2$  (Fig. 3), indicating that a majority of the core microbial community in soil responded to the Cd contamination. *Nitrososphaera* (3.66%), *Nitrosospira* (3.63%) and *Sulfuricella* (3.06%) were the three most abundant genera in  $S_2$ . The total



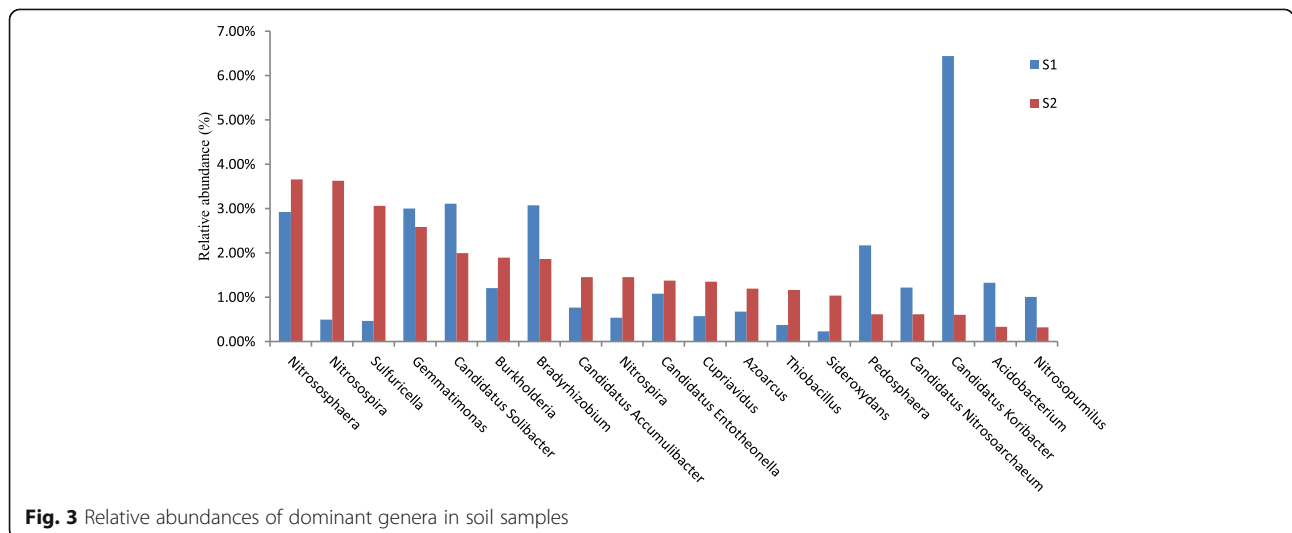
relative abundance of the three dominant genera in their own corresponding group accounted for over 10% of the entire 1736 detected genera. According to the distribution of the soil microbes at the phylum and genus level, we found that the abundance of microbial species in S1 was higher than that in S2. Furthermore, the species diversity of S1 was greater than that of S2. The results showed that Cd-contamination decreased the microbial species diversity of the soil, but made the composition of microorganisms in the soil more concentrated.

**KEGG function annotation and analysis**

Apart from phylogenic insights, metagenomic analysis also provided an opportunity to assess the functional

potentials associated with the soil microbial community. We blasted the predicted genes against KEGG GENES via BLAST search (BLAST Version 2.2.28+, <http://blast.ncbi.nlm.nih.gov/Blast.cgi>) at the optimized threshold of 1e-5. Then, we conducted the functional annotation of genes via KOBAS 2.0 according to the blast results. Results indicated that the unpolluted soil (S<sub>1</sub>) has more genes, pathways, modules, orthologies and enzymes involved in the metabolic pathways of microbial communities than the Cd-contaminated soil (S<sub>2</sub>) (Table 3).

We have defined dominant pathway as those with a relative abundance greater than 1.00% of the total observed pathways. Then we obtained 37 dominant pathways in the two soil samples (Fig. 4). Among the 37



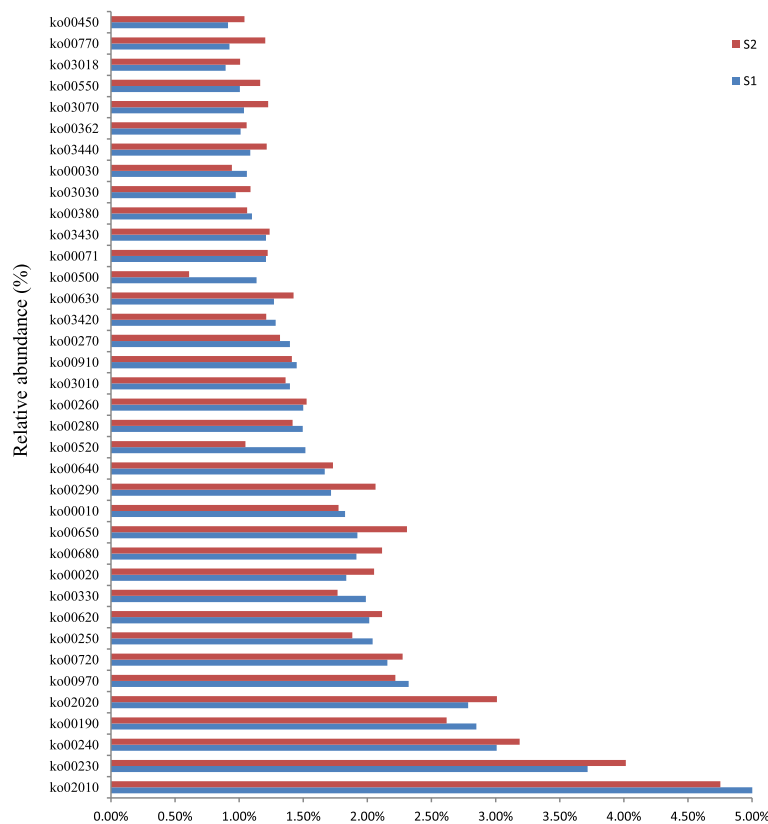
**Table 3** Summary of the KEGG annotation results

Annotation categories	KEGG Genes	KEGG Pathway	KEGG Module	KEGG Orthology	Enzyme
S <sub>1</sub>	63,026	230	376	3680	1469
S <sub>2</sub>	21,622	213	331	2978	1248

dominant pathways, the relative abundance of twenty-one KEGG pathways in the Cd-contaminated soil was higher than in the non-contaminated soil. There were 230 pathways determined in the two samples and the functions were depicted in the additional file. The most abundant pathway observed in the two samples was ko02010, which could be assigned to ABC transporters (Fig. 4 and Table 4). The relative abundance of ko02010 in S<sub>1</sub> was higher than that in S<sub>2</sub>. But the relative abundances of some modules belonging to ko02010 were increased in S<sub>2</sub> compared to S<sub>1</sub>, such as M00239 (Fig. 5). The ABC transporters could transport amino acids, lipids, lipid sugar, peptides and inorganic ions such as metal ions to the outside of the cells [56], and some of them also had a particular import function [57]. All the 37 dominant pathways were involved in some detoxification compound metabolisms (ko00230, ko00250), such as organic acid (ko00280), soluble amino acids (ko00330), fatty

acid (ko00071), proteins and selenocompounds (ko00450), which could combine with Cd and decrease its biological toxicity [58, 59]. The collective abundance of the 37 dominant pathways accounted for almost 60% of the entire 230 pathways in S<sub>1</sub> and 55% of the 213 pathways in the Cd-polluted soil (S<sub>2</sub>). This indicated that Cd contamination can inhibit part of the KEGG pathways in the soil microorganisms. Meanwhile we also found that the relative abundances of other pathways (about 45% of the total pathways) in S<sub>2</sub> increased when compared to S<sub>1</sub>, indicating that these pathways have the potential to help microorganisms cope with cadmium stress.

In order to investigate the precise functions of the entire 230 pathways, we divided the entire pathways into 399 different functional modules. Almost 90% of the total pathways relative abundances in S<sub>2</sub> were lower than those in S<sub>1</sub>, which indicated that Cd-contamination could affect the majority of the modules. There were 25 dominant modules (Fig. 5), as defined by the relative abundance greater than 1.00% of the whole modules in one or both of the samples. M00360 and M00359 were the two most abundant modules observed in the samples, and they were components of the ko00970 pathway, which was assigned to aminoacyl-tRNA biosynthesis (Table 5). The relative abundance of the 25



**Fig. 4** Relative abundance of dominant KEGG pathways (> 1.00%) in the two soil samples

**Table 4** Definition of the KEGG pathways (> 1.00%) in the two soil samples

Pathway	Definition
ko02010	ABC transporters
ko00230	Purine metabolism
ko00240	Pyrimidine metabolism
ko00190	Oxidative phosphorylation
ko02020	Two-component system
ko00970	Aminoacyl-tRNA biosynthesis
ko00720	Carbon fixation pathways in prokaryotes
ko00250	Alanine, aspartate and glutamate metabolism
ko00620	Pyruvate metabolism
ko00330	Arginine and proline metabolism
ko00020	Citrate cycle (TCA cycle)
ko00680	Methane metabolism
ko00650	Butanoate metabolism
ko00010	Glycolysis/Gluconeogenesis
ko00290	Valine, leucine and isoleucine biosynthesis
ko00640	Propanoate metabolism
ko00520	Amino sugar and nucleotide sugar metabolism
ko00280	Valine, leucine and isoleucine degradation
ko00260	Glycine, serine and threonine metabolism
ko03010	Ribosome
ko00910	Nitrogen metabolism
ko00270	Cysteine and methionine metabolism
ko03420	Nucleotide excision repair
ko00630	Glyoxylate and dicarboxylate metabolism
ko00500	Starch and sucrose metabolism
ko00071	Fatty acid metabolism
ko03430	Mismatch repair
ko00380	Tryptophan metabolism
ko03030	DNA replication
ko00030	Pentose phosphate pathway
ko03440	Homologous recombination
ko00362	Benzoate degradation
ko03070	Bacterial secretion system
ko00550	Peptidoglycan biosynthesis
ko03018	RNA degradation
ko00770	Pantothenate and CoA biosynthesis
ko00450	Selenocompound metabolism

dominant modules made up over 40% of the total modules identified in the two samples. M00009, M00011, M00017, M00019, M00048, M00051, M00087, M00173, M00178, M00237, M00239, M00307, M00335, M00374, M00376 and M00570 were the dominant modules in the Cd-contaminated soil. These sixteen (64%) of the dominant

modules involved in the Cd-contaminated soil were of a higher relative abundance than those in the unpolluted soil, which indicated that these modules had a potential role in microbial resistance or tolerance to cadmium. Nine of the dominant modules involved in the unpolluted soil were of a higher relative abundance than those detected in Cd-contamination soil, suggesting that most of the dominant modules were inhibited under the Cd contamination. The results of the metabolism module analysis revealed that the relative abundance of some metabolisms decreased significantly in Cd-polluted soil.

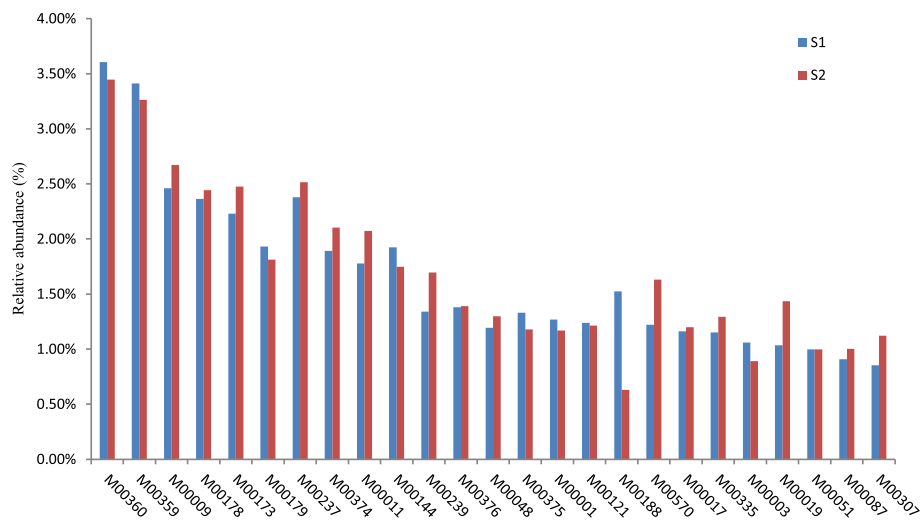
In order to completely understand the metabolism variation in the two soil samples, we summarized the enzymes involved in all the identified 230 pathways, then, we achieved 1581 enzymes from the metagenomic sequencing. According to the International Enzymatic Commission classification system, the enzymes associated with these soil samples were classified into six categories: oxidoreductase (I), transferase (II), hydrolase (III), lyase (IV), isomerase (V) and ligase or synthetase (VI). The relative abundance of these enzymes was calculated via BLAST against the KEGG database (Fig. 6). Transferase was the most abundant enzymatic category, followed by oxidoreductase. Isomerase was the least abundant enzyme observed in the samples. The relative abundance of hydrolase, isomerase and ligase in S<sub>2</sub> was lower than that of S<sub>1</sub>, but the relative abundance of oxidoreductase, transferase, and lyase in S<sub>2</sub> was higher than that in S<sub>1</sub>, which indicated that these enzymes had a function potential of tolerance to Cd-contamination and could help the soil microorganisms endure Cd stress.

## Discussion

In this study, high-throughput metagenomic sequencing provided a powerful strategy to investigate the microbial community structure and functional potential associated in Cd-contaminated soil. Previous studies reported that *Nitrosospira*-like organisms were the major nitrifiers of the community that may possess cadmium tolerance [60, 61]. By characterizing the abundances of genes belonging to these genera, metagenomic sequencing revealed the variations of microbial abundance and diversity under different environmental conditions. For example, the relative abundances of genera belonging to *Nitrosospira* and *Sulfuricella* in S<sub>1</sub> were 0.49% and 0.47%, respectively (Fig. 3), while their relative abundances in S<sub>2</sub> increased to 3.63% and 3.06%, respectively. These results revealed that Cd contamination could increase the relative abundances of some microorganisms with a potential tolerance to heavy metal, and subsequently decreased the abundance of other microorganisms sensitive to heavy metals.

All of the core microorganisms were Gram-negative *Proteobacteria*, and most of them were related to the





**Fig. 5** Relative abundance of dominant KEGG modules (> 1.0%) in the two soil samples

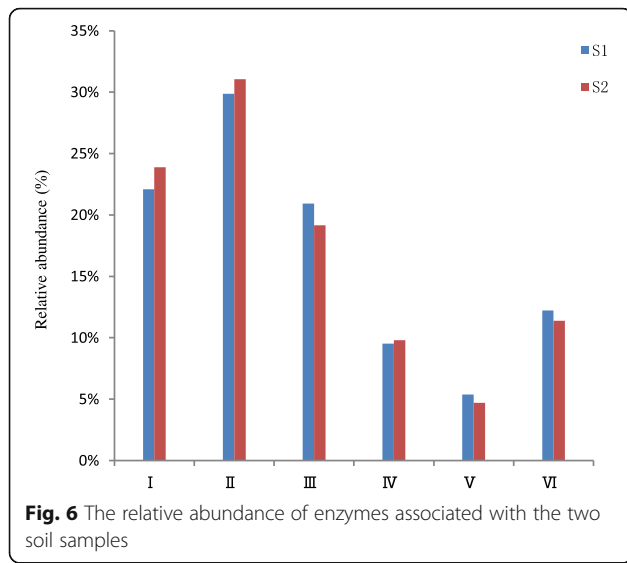
nitrogen cycle, which was an important process within the soil microbial community. It was reported that *Sulfuricella* and *Thiobacillus* mainly caused denitrification and dissimilatory sulfate reduction [62]. *Sideroxydans* belonged to a kind of iron oxide bacteria, *Candidatus Accumulibacter* was a phosphorus accumulating bacteria [63], and *Cupriavidus* was capable of a tolerance to copper [64]. All functions involved in these bacteria contributed to the tolerance of microorganisms to heavy metals [65–67].

We found that most of the COG annotations were involved in microbial metabolism (Fig. 1), and the relative abundance of cluster III (metabolism) was the highest

among the four clusters but the relative abundance of metabolic genes in  $S_1$  was lower than that in  $S_2$ . This indicated that Cd-contamination increased the activity of microorganisms in the soil. An increase in the relative abundance of metabolism genes suggested that the corresponding bacteria possessed a potential tolerance to Cd, and carried out normal metabolism in Cd-contaminated soil. The most abundant pathway among the entire 230 pathways affected by the Cd contamination was ko02010, which was assigned to ABC transporter. The current research on microbial ABC transporters showed that it involved in many biological functions, including bacterial drug resistance, pheromone secretion,

**Table 5** Definition of dominant KEGG modules (> 1.0%) in the two soil samples

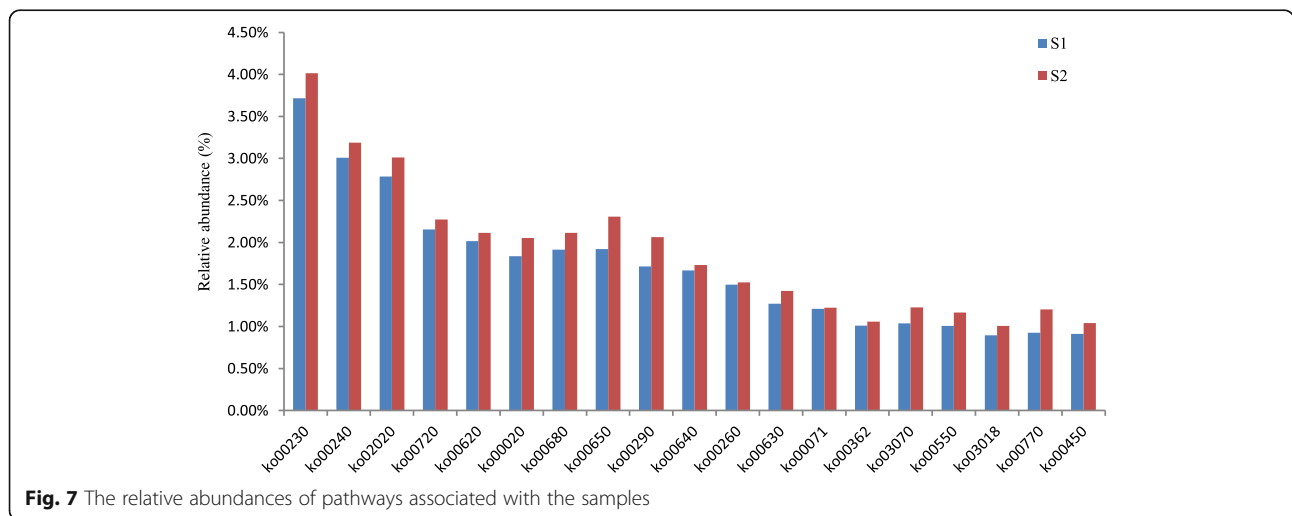
Module	Definition
M00009	Citrate cycle (TCA cycle, Krebs cycle) [PATH:map01200 map00020]
M00178	Ribosome, bacteria [PATH:map03010]
M00173	Reductive citrate cycle (Arnon-Buchanan cycle) [PATH:map01200 map00720]
M00237	Branched-chain amino acid transport system [PATH:map02010] [BR:ko02000]
M00374	Dicarboxylate-hydroxybutyrate cycle [PATH:map01200 map00720]
M00011	Citrate cycle, second carbon oxidation, 2-oxoglutarate = > oxaloacetate [PATH:map01200 map00020]
M00239	Peptides/nickel transport system [PATH:map02010]
M00376	3-Hydroxypropionate bi-cycle [PATH:map01200 map00720]
M00048	Inosine monophosphate biosynthesis, PRPP + glutamine = > IMP [PATH:map00230]
M00570	Isoleucine biosynthesis, threonine = > 2-oxobutanoate = > isoleucine [PATH:map01230 map00290]
M00017	Methionine biosynthesis, aspartate = > homoserine = > methionine [PATH:map01230 map00270]
M00335	Sec (secretion) system [PATH:map03070]
M00019	Valine/isoleucine biosynthesis, pyruvate = > valine / 2-oxobutanoate = > isoleucine [PATH:map01210 map01230 map00290]
M00051	Uridine monophosphate biosynthesis, glutamine (+ PRPP) = > UMP [PATH:map00240]
M00087	beta-Oxidation [PATH:map01212 map00071]
M00307	Pyruvate oxidation, pyruvate = > acetyl-CoA [PATH:map01200 map00010 map00020 map00620]



mitochondria biological function and detoxification of heavy metals, etc. [68–70]. It was reported that ABC transporter played an important role in heavy metal detoxification, and was regarded to be resistant to Cd contamination in *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae* [68–70]. It was generally acknowledged that the mechanism of microbes coping with the heavy metal could be classified into two modes: in vitro and in vivo [71, 72]. Microbial organism could synthesis and secrete some chelates to forming co-precipitation with heavy metals under heavy metal stress. Then, these co-precipitations were prevented entering into the cells by microbial membrane. On the other hand, the organic acid or some proteins in vivo could combine with the heavy metals entered into the cells to form complex precipitation and detoxify the heavy metal poisoning through some enzymatic reaction or fixing them in the low active part [73]. Most of the dominant pathways whose abundance

increased in the Cd contaminated soil were related to the metabolism, biosynthesis and degradation of amino acids, fatty acids and nucleotides, suggesting that these metabolic activities were related to Cd tolerance in the microorganisms or could help them survive from Cd-contamination at least (Fig. 7 and Additional file 1). Our results revealed that some soil microorganisms had a potential function in heavy metal tolerance and provided a new way to explore indigenous genera capable of bio-remediation.

The metagenomic analysis used in our experiment indicated that the method was adequate to assess the influence of cadmium pollution on microbial community and functional potential. The soil microorganisms were quite abundant, and many of them possessed numerous physiological functions. A stable microbial community is a critical factor to keep the bioavailability of soil. The total metagenomic sequences were assigned to 77 different phyla and 1736 genera. Among them, *Proteobacteria*, *Gemmatimonadetes*, *Thaumarchaeota* and *Acidobacteria* were the most common phyla in the samples, and comprised more than 75% of the total population in each sample. The results suggested that Cd-contamination had the potential to decrease the taxonomic species of microbes in soil, and further changed the soil microbial composition. It was also revealed that the functional pathways involved in the soil changed with microbial structure variation, many of which were related to the heavy metal tolerance of the soil microbes. These findings illustrated that Cd-contaminated soil was potential resource for exploring cadmium resistant or tolerant bacteria. In our study, all the core microorganisms were Gram-negative bacteria, belonging to *Proteobacteria*, with most of them being related to the nitrogen cycle and heavy metal tolerance. These bacteria and their physiological functions were important for the soil microbial community to cope with cadmium, as well as the stability and biological availability of the soil.



## Conclusion

Our research indicates that Cd-contamination has the potential to decrease the taxonomic species of microbes in soil, and further change the soil microbial composition. The functional pathways involved in the soil change with microbial structure variation, many of which are related to the heavy metal tolerance of the soil microbes. The Cd-contaminated soil is a potential resource for exploring cadmium resistant or tolerant bacteria.

## Additional file

**Additional file 1:** Definition of the KEGG pathways in the two soil samples. (DOCX 22 kb)

## Abbreviations

AAS: Atomic absorption spectroscopy; AM: *Arbuscular Mycorrhizal*; Cd: Cadmium; COG: Cluster of orthologous groups; ICP-MS: Inductively coupled plasma mass spectrometry; KEGG: Kyoto Encyclopedia of Genes and Genomes; ORFs: Open reading frames; PDB: Protein Data Bank; PIR: Protein Information Resource; PRF: Protein Research Foundation; SOM: Soil organic matter; TK: Total potassium; TN: Total nitrogen; TP: Total phosphorous

## Acknowledgements

We would like to thank Majorbio Bio-Pharm Technology Co., Ltd. for assistance with Illumina genome Analyzing.

## Funding

This work was supported by the Science and Technology Department of Sichuan Province, P. R. China (Grant No. 2016NYZ0031 and 2017SZ0181). The funding bodies had no role in the design of the study, in data collection, analysis or interpretation, or in writing the manuscript.

## Availability of data and materials

The datasets used and analyzed within the current study are available from the corresponding author on reasonable request.

## Authors' contributions

YZ coordinated the study, collected data, interpreted data, drafted parts of the manuscript and co-wrote the manuscript. GF, TX, XW allocated funds, curated the metagenomic data, conducted data analysis, interpreted the data, drafted parts of the manuscript and co-wrote the manuscript. XW, JB, LT collected data and critically analyzed the manuscript. MW, HZ, allocated funds and critically analyzed the manuscript. WW allocated funds and co-analysis the data. All authors read and approved the final manuscript.

## Ethics approval and consent to participate

Not applicable.

## Consent for publication

Not applicable.

## Competing interests

All authors declare that they have no competing interests.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Author details

<sup>1</sup>Key Laboratory of Bio-Resource and Eco-Environment of Ministry of Education, College of Life sciences, Sichuan University, No. 24 South Section 1, Yihuan Road, Chengdu 610065, China. <sup>2</sup>Tibet Academy of Agriculture and Animal Husbandry Sciences, Lhasa 850002, China. <sup>3</sup>Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu 610041, China.

Received: 14 March 2017 Accepted: 7 February 2018

Published online: 13 February 2018

## References

1. Facchinelli A, Sacchi E, Mallen L. Multivariate statistical and GIS-based approach to identify heavy metal sources in soils. *Environ Pollut*. 2001;114(3):313–24.
2. Solgi E, Esmaili-Sari A, Riyahi-Bakhtiari A, Hadipour M. Soil contamination of metals in the three industrial estates, arak, Iran. *Bull Environ Contam Toxicol*. 2012;88(4):634–8.
3. Dong WQY, Cui Y, Liu X. Instances of soil and crop heavy metal contamination in China. *Soil & Sediment Contamination An International Journal*. 2001;10(5):497–510.
4. Salt DE, Blaylock M, Kumar NPBA, Dushenkov V, Ensley BD, Chet I, Raskin I. Phytoremediation: a novel strategy for the removal of toxic metals from the environment using plants. *Nat Biotechnol*. 1995;13(5):468–74.
5. Zhuang P, McBride MB, Xia H, Li N, Li Z. Health risk from heavy metals via consumption of food crops in the vicinity of Dabaoshan mine, South China. *Sci Total Environ*. 2009;407(5):1551–61.
6. Li Z, Ma Z, van der Kuijp TJ, Yuan Z, Huang L. A review of soil heavy metal pollution from mines in China: pollution and health risk assessment. *Sci Total Environ*. 2014;468-469:843–53.
7. Fu X, Dou C, Chen Y, Chen X, Shi J, Yu M, Jie X. Subcellular distribution and chemical forms of cadmium in *Phytolacca Americana* L. *J Hazard Mater*. 2011;186(1):103–7.
8. Maron PA, Mougel C, Ranjard L. Soil microbial diversity: methodological strategy, spatial overview and functional interest. *Comptes rendus biologies*. 2011;334(5–6):403–11.
9. Toppi LSD, Gabbriellini R. Response to cadmium in higher plants. *Environmental & Experimental Botany*. 1999;41(2):105–30.
10. Balestrasse KB, Gardley L, ML GSMT. Response of antioxidant defence system in soybean nodules and rootssubjected to cadmium stress. *Funct Plant Biol*. 2001;28(6):497–504.
11. Fornazier RF, Ferreira RR, Vitória AP, Molina SMG, Lea PJ, Azevedo RA. Effects of cadmium on antioxidant enzyme activities in sugar cane. *Biol Plant*. 2002;45(1):91–7.
12. Cho UH, Seo NH. Oxidative stress in *Arabidopsis Thaliana* exposed to cadmium is due to hydrogen peroxide accumulation. *Plant Sci*. 2005;168(1):113–20.
13. Gao L, Zhang J, Shen G. Effect of cd contamination on soil microbial community structure in flue-cured tobacco rhizosphere. *Science News*. 2014;164(1):7–9.
14. Almás AR, Bakken LR, Mulder J. Changes in tolerance of soil microbial communities in Zn and cd contaminated soils. *Soil Biol Biochem*. 2004;36(5):805–13.
15. Ohya H, Komai Y, Yamaguchi M. Zinc effects on a soil bacterial flora characterized by fatty acid composition of the isolates. *Biol Fertil Soils*. 1986;2(2):59–63.
16. Frostegard A, Tunlid A, Bååth E. Phospholipid fatty acid composition, biomass, and activity of microbial communities from two soil types experimentally exposed to different heavy metals. *Applied & Environmental Microbiology*. 1993;59(11):3605–17.
17. Frostegard A, Tunlid A, Baath E. Changes in microbial community structure during long-term incubation in two soils experimentally contaminated with metals. *Soil Biol Biochem*. 1996;28(1):55–63.
18. Yang Y, Liu C, Xu L, Wu P, Zhang G. Effects of heavy metal contamination on microbial biomass and community structure in soils. *Chin J Geochem*. 2004;23(4):319–28.
19. Wang Y, Shi J, Wang H, Lin Q, Chen X, Chen Y. The influence of soil heavy metals pollution on soil microbial biomass, enzyme activity, and community composition near a copper smelter. *Ecotoxicology & Environmental Safety*. 2007;67(1):75.
20. Kumpiene J, Guerri G, Landi L, Pietramellara G, Nannipieri P, Renella G. Microbial biomass, respiration and enzyme activities after in situ aided phytostabilization of a Pb- and cu-contaminated soil. *Ecotoxicology & Environmental Safety*. 2009;72(1):115–9.
21. Navarro-Noya YE, Jan-Roblero J, Gonzalez-Chavez Mdel C, Hernandez-Gama R, Hernandez-Rodriguez C. Bacterial communities associated with the rhizosphere of pioneer plants (*Bahia Xylopoda* and *Viguiera Linearis*) growing on heavy metals-contaminated soils. *Antonie Van Leeuwenhoek*. 2010;97(4):335–49.

22. Lundberg DS, Lebeis SL, Paredes SH, Yourstone S, Gehring J, Malfatti S, Tremblay J, Engelbrekton A, Kunin V, del Rio TG, et al. Defining the core *Arabidopsis thaliana* root microbiome. *Nature*. 2012;488(7409):86–90.
23. Kaschuk G, Alberton O, Hungria M. Three decades of soil microbial biomass studies in Brazilian ecosystems: lessons learned about soil quality and indications for improving sustainability. *Soil Biol Biochem*. 2010;42(1):1–13.
24. Preston-Mafham J, Boddy L, Randerson PF. Analysis of microbial community functional diversity using sole-carbon-source utilisation profiles—a critique. *FEMS Microbiol Ecol*. 2002;42(1):1–14.
25. Li Y, Chen L, Wen H. Changes in the composition and diversity of bacterial communities 13 years after soil reclamation of abandoned mine land in eastern China. *Ecol Res*. 2014;30(2):357–66.
26. de Campos SB, Youn JW, Farina R, Jaenicke S, Junemann S, Szczepanowski R, Beneduzi A, Vargas LK, Goesmann A, Wendisch VF, et al. Changes in root bacterial communities associated to two different development stages of canola (*Brassica napus* L. var *oleifera*) evaluated through next-generation sequencing technology. *Microb Ecol*. 2013;65(3):593–601.
27. Hill G, Mitkowski N, Aldrich-Wolfe L, Emele L, Jurkonie D, Ficke A, Maldonado-Ramirez S, Lynch S, Nelson E. Methods for assessing the composition and diversity of soil microbial communities. *Appl Soil Ecol*. 2000;15(1):25–36.
28. Hong C, Si Y, Xing Y, Li Y. Illumina MiSeq sequencing investigation on the contrasting soil bacterial community structures in different iron mining areas. *Environ Sci Pollut Res Int*. 2015;22(14):10788–99.
29. Tonin C, Vandenkoornhuysse P, Joner E, Straczek J, Leyval C. Assessment of arbuscular mycorrhizal fungi diversity in the rhizosphere of *Viola calaminaria* and effect of these fungi on heavy metal uptake by clover. *Mycorrhiza*. 2001;10(4):161–8.
30. Amann RI, Ludwig W, Schleifer KH. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiological Rev*. 1995;59(1):143–69.
31. Giovannoni SJ, Britschgi TB, Moyer CL, Field KG. Genetic diversity in Sargasso Sea bacterioplankton. *Nature*. 1990;345(6270):60–3.
32. Hugenholtz P, Pitulle C, Hershberger KL, Pace NR. Novel division level bacterial diversity in a Yellowstone hot spring. *J Bacteriol*. 1998;180(2):366–76.
33. Torsvik V, Goksøyr J, Daae FL. High diversity in DNA of soil bacteria. *Applied & Environmental Microbiology*. 1990;56(3):782–7.
34. Bhadra B, Nanda AK, Chakraborty R. Inducible nickel resistance in a river isolate of India phylogenetically ascertained as a novel strain of *Acinetobacter junii*. *World J Microbiol Biotechnol*. 2005;22(3):225–32.
35. Rhee SK, Liu X, Wu L, Chong SC, Wan X, Zhou J. Detection of genes involved in biodegradation and biotransformation in microbial communities by using 50-mer oligonucleotide microarrays. *Appl Environ Microbiol*. 2004;70(7):4303–17.
36. Trajanovska S, Britz ML, Bhawe M. Detection of heavy metal ion resistance genes in gram-positive and gram-negative bacteria isolated from a lead-contaminated site. *Biodegradation*. 1997;8(2):113–24.
37. Jones DS, Albrecht HL, Dawson KS, Schaperdoth I, Freeman KH, Pi Y, Pearson A, Macalady JL. Community genomic analysis of an extremely acidophilic sulfur-oxidizing biofilm. *The ISME journal*. 2012;6(1):158–70.
38. Mason OU, Hazen TC, Borglin S, Chain PS, Dubinsky EA, Fortney JL, Han J, Holman HY, Hultman J, Lamendella R, et al. Metagenome, metatranscriptome and single-cell sequencing reveal microbial response to Deepwater horizon oil spill. *The ISME journal*. 2012;6(9):1715–27.
39. Sogin ML, Morrison HG, Huber JA, Mark Welch D, Huse SM, Neal PR, Arrieta JM, Herndl GJ. Microbial diversity in the deep sea and the underexplored “rare biosphere”. *Proc Natl Acad Sci U S A*. 2006;103(32):12115–20.
40. Castelle CJ, Hug LA, Wrighton KC, Thomas BC, Williams KH, Wu D, Tringe SG, Singer SW, Eisen JA, Banfield JF. Extraordinary phylogenetic diversity and metabolic versatility in aquifer sediment. *Nat Commun*. 2013;4:2120.
41. Kantor RS, Wrighton KC, Handley KM, Sharon I, Hug LA, Castelle CJ, Thomas BC, Banfield JF. Small genomes and sparse metabolisms of sediment-associated bacteria from four candidate phyla. *MBio*. 2013;4(5):e00708–13.
42. Musat N, Halm H, Winterholler B, Hoppe P, Peduzzi S, Hillion F, Horreard F, Amann R, Jorgensen BB, Kuypers MM. A single-cell view on the ecophysiology of anaerobic phototrophic bacteria. *Proc Natl Acad Sci U S A*. 2008;105(46):17861–6.
43. Wrighton KC, Thomas BC, Sharon I, Miller CS, Castelle CJ, Verberkmoes NC, Wilkins MJ, Hettich RL, Lipton MS, Williams KH. Fermentation, hydrogen, and sulfur metabolism in multiple uncultivated bacterial phyla. *Science*. 2012;337(6102):1661–5.
44. Hole JA, Ryberg T, Fuis GS, Bleibinhaus F, Sharma AK. Microbial population structures in the deep marine biosphere. *Science*. 2007;318(5847):97–100.
45. Altimira F, Yáñez C, Bravo G, González M, Rojas LA, Seeger M. Characterization of copper-resistant bacteria and bacterial communities from copper-polluted agricultural soils of central Chile. *BMC Microbiol*. 2012;12(1):193.
46. Dopson M, Baker-Austin C, Koppineedi PR, Bond PL. Growth in sulfidic mineral environments: metal resistance mechanisms in acidophilic micro-organisms. *Microbiology*. 2003;149(Pt 8):1959–70.
47. Bloem J, Hopkins DW, Benedetti A. Microbiological methods for assessing soil quality. *Microbiological Methods for Assessing Soil Quality 2006*, 36(1995):249–259. CABI Pub 2006, 36(1995): 249–259.
48. Sun JY, Li MZ, Zheng LH, Hu YG, Zhang XJ. Real-time analysis of soil moisture, soil organic matter, and soil total nitrogen with NIR spectra. *Spectroscopy & Spectral Analysis*. 2006;26(3):426–9.
49. Curtius AJ, Schlemmer G, Welz B. Determination of phosphorus by graphite furnace atomic absorption spectrometry. Part 2. Comparison of different modifiers. *J Anal At Spectrom*. 1987;2(2):115–24.
50. Bao SD. Soil and agricultural chemistry analysis. Beijing: China Agriculture Press; 2000.
51. Clark MJ, Chen R, Lam HY, Karczewski KJ, Chen R, Euskirchen G, Butte AJ, Snyder M. Performance comparison of exome DNA sequencing technologies. *Nat Biotechnol*. 2011;29(10):908–14.
52. Ma J, Wang Z, Li H, Park H-D, Zhichao W. Metagenomes reveal microbial structures, functional potentials, and biofouling-related genes in a membrane bioreactor. *Applied Microbiology & Biotechnology*. 2016;100(11):5109–21.
53. Li H, Durbin R. Fast and accurate short read alignment with burrows–wheeler transform. *Bioinformatics*. 2009;25(14):1754–60.
54. Parks DH, Beiko RG. Identifying biologically relevant differences between metagenomic communities. *Bioinformatics*. 2010;26(6):715–21.
55. Albertsen M, Hansen LB, Saunders AM, Nielsen PH, Nielsen KL. A metagenome of a full-scale microbial community carrying out enhanced biological phosphorus removal. *The ISME journal*. 2012;6(6):1094–106.
56. Ter BJ, Guskov A, Slotboom DJ. Structural diversity of ABC transporters. *J Gen Physiol*. 2014;143(4):419.
57. Higgins CF. ABC transporters: from microorganisms to man. *Annu Rev Cell Biol*. 1992;8(1):67–113.
58. Finlayson JC, Liao B, Droppo IG, Leppard GG, Liss SN. The relationship between the structure of activated sludge flocs and the sorption of hydrophobic pollutants. *Water Science & Technology*. 1998;37(4–5):353–7.
59. Kurek E, Francis AJ, Bollag JM. Immobilization of cadmium by microbial extracellular products. *Archives of Environmental Contamination & Toxicology*. 1991;21(1):106–11.
60. Demanou J, Sharma S, Weber A, Wilke BM, Njine T, Monkiedje A, Munch JC, Schlöter M. Shifts in microbial community functions and nitrifying communities as a result of combined application of copper and mephenoxam. *FEMS Microbiol Lett*. 2006;260(1):55–62.
61. Mertoglu B, Semerci N, Guler N, Calli B, Cecen F, Saatci AM. Monitoring of population shifts in an enriched nitrifying system under gradually increased cadmium loading. *J Hazard Mater*. 2008;160(2–3):495–501.
62. Carmalin Sophia A, Swaminathan K, Sandhya S. Microbially-influenced degradation of solidified/stabilized metal waste. *Bioresour Technol*. 2007;98(13):2562–7.
63. Seviour RJ, Mino T, Onuki M. The microbiology of biological phosphorus removal in activated sludge systems. *FEMS Microbiol Rev*. 2003;27(1):99–127.
64. Vandamme P, Coenye T. Taxonomy of the genus *Cupriavidus*: a tale of lost and found. *Int J Syst Evol Microbiol*. 2004;54(Pt 6):2285–9.
65. Chang IS, Kim BH. Effect of sulfate reduction activity on biological treatment of hexavalent chromium [Cr(VI)] contaminated electroplating wastewater under sulfate-rich condition. *Chemosphere*. 2007;68(2):218–26.
66. Cook WJ, Kar SR, Taylor KB, Hall LM. Crystal structure of the cyanobacterial metallothionein repressor SmtB: a model for metalloregulatory proteins 1. *J Mol Biol*. 1998;275(2):337–46.
67. Goulhen F, Gloter A, Guyot F, Bruschi M. Cr(VI) detoxification by *Desulfovibrio vulgaris* strain Hildenborough: microbe-metal interactions studies. *Appl Microbiol Biotechnol*. 2006;71(6):892–7.
68. Bauer BE, Wolfger H, Kuchler K. Inventory and function of yeast ABC proteins: about sex, stress, pleiotropic drug and heavy metal resistance. *Biochim Biophys Acta*. 2000;1461(2):217–36.

69. Lee JY, Yang JG, Zhitnitsky D, Lewinson O, Rees DC. Structural basis for heavy metal detoxification by an Atm1-type ABC exporter. *Science*. 2014;343(6175):1133–6.
70. Lee J, Yang J, Zhitnitsky D, Lewinson O, Rees D. Structural and functional characterization of a heavy metal detoxifying ABC transporter. *Sci-Tech Information Development & Economy*. 2014;24(3):308–13.
71. Nies DH: Bacterial transition metal homeostasis. *Springer Berlin Heidelberg* 2007, 3(2):117–142.
72. Prange A, Modrow H. *Rev Environ Sci Biotechnol*. 2015;
73. Silver S, Le TP. Bacterial Heavy Metal Resistance: new surprises. *Annu Rev Microbiol*. 1996;50(50):753.
74. Miller JR, Koren S, Sutton G. Assembly algorithms for next-generation sequencing data. *Genomics*. 2010;95(6):315–27.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

