Polycyclic aromatic hydrocarbons (PAHs) enriching antibiotic resistance genes (ARGs) in the Soils

Baowei Chen^{§,§}, Rong He^{§,§}, Ke Yuan^{§,§}, Enzhong Chen[†], Lan Lin[†], Xin Chen^{§,§}, Sha Sha[‡],
Jianan Zhong[‡], Li Lin[‡], Lihua Yang^{§,§}, Ying Yang^{§,§}, Xiaowei Wang^{§,§}, Shichun Zou^{§,§},
Tiangang Luan^{‡,*}

[§]South China Sea Resource Exploitation and Protection Collaborative Innovation Center, School of Marine Sciences, Sun Yat-sen University, Guangzhou 510275, China

[†]Zhujiang Hospital of Southern Medical University, Guangzhou 510282, China

[‡]MOE Key Laboratory of Aquatic Product Safety, School of Life Sciences, Sun Yat-Sen University, Guangzhou 510275, China

[§]Guangdong Provincial Key Laboratory of Marine Resources and Coastal Engineering, School of Marine Sciences, Sun Yat-Sen University, Guangzhou 510275, China

*Corresponding author

Dr. Tiangang Luan

Telephone: +86-20-84112958; Fax: +86-20-84037549

Email address: cesltg@mail.sysu.edu.cn (T. Luan)

ABSTRACT

The prevalence of antibiotic resistance genes (ARGs) in modern environment raises an emerging global health concern. In this study, soil samples were collected from three sites in petrochemical plant that represented different pollution levels of polycyclic aromatic hydrocarbons (PAHs). Metagenomic profiling of these soils demonstrated that ARGs in the PAHs-contaminated soils were approximately 15 times more abundant than those in the less-contaminated ones, with Proteobacterial being the preponderant phylum. Resistance profile of ARGs in the PAHs-polluted soils was characterized by the dominance of efflux pump-encoding ARGs associated with aromatic antibiotics (e.g., fluoroquinolones and acriflavine) that accounted for more than 70% of the total ARGs, which was significantly different from representative sources of ARG pollution due to wide use of antibiotics. Most of ARGs enriched in the PAHs-contaminated soils were not carried by plasmids, indicating the low possibilities of them being transferred between bacteria. Significant correlation was observed between the total abundance of ARGs and that of Proteobacteria in the soils. Proteobacteria selected by PAHs led to simultaneously enriching of ARGs carried by them in the soils. Our results suggested that PAHs could serve as one of selective stresses for greatly enriching of ARGs in the human-impacted environment.

Keywords: Antibiotic resistance genes, Polycyclic aromatic hydrocarbons, Soils, Enrichment, Metagenomics

1. Introduction

Antibiotic resistance genes (ARGs) have been recognized as an emerging environmental ‘pollutant’ (Pruden et al., 2006). Antibiotic resistance encoded by ARGs is popular and abundant in various matrices collected from the highly human-impacted environment, e.g., soils, water, and sediments (Chen et al., 2013a; Chen et al., 2013b; Knapp et al., 2010). Potential transfer of ARGs from environmental bacteria to human pathogens may seriously impair the efficacy of antibiotic therapy, thereby posing a deleterious health risk to the public (Pehrsson et al., 2016). In some clinical cases, some antibiotics are out of use due to the prevalence of their resistance and treatment options for certain pathogens have been becoming very limited (Wenzel and Edmond, 2000), whereas bacterial infections still annually claim hundred thousands of lives worldwide (Nathan, 2004).

The soils are considered as the largest reservoir of antibiotic resistant bacteria (ARB) and ARGs (Nesme and Simonet, 2015). A large variety of novel ARGs were frequently discovered in the soils and represent all major resistance mechanisms, which largely expanded our knowledge on antibiotic resistome in the environment (Forsberg et al., 2014). Analysis of archived soils demonstrated that ARG abundance was substantially increased since antibiotic era began at 1940 (Knapp et al., 2010). However, the question is that what type of anthropogenic contaminant can result in the dissemination and accumulation of ARGs in the soils. First of all, a huge amount of commonly-used antibiotics in human medicine, husbandry and aquaculture establish the major human impacts on the spread and enrichment of antibiotic resistance in various environment niches (Chee-Sanford et al., 2009; Larson, 2007). Meanwhile, other contaminants with significant antimicrobial activities could not be ignored in shaping bacterial communities and accordingly selecting antibiotic resistance. Growing evidence has demonstrated the selection of ARB and ARGs by heavy metals (Baker-Austin et al., 2006), as well as by some organic contaminants, e.g., polychlorinated biphenyls (Lo Giudice et al., 2013), pesticides (Anjum and

Krakat, 2016), and disinfectants (Karatzas et al., 2007). It is well believed that more organic contaminants will be found to have significant selection for the ARB and ARGs.

Polycyclic aromatic hydrocarbons (PAHs) are ubiquitous in various environments (Bosch et al., 2015; Choi et al., 2013), and their concentrations are unexpectedly high in some pollution accidents, e.g., oil spills and coking wastewater (Brandt et al., 2002; Wang et al., 2014; Zhang et al., 2012). As a result, community structure of indigenous bacteria in the soils was considerably changed due to the contamination of PAHs (de Menezes et al., 2012; Yang et al., 2014; Zhang et al., 2010). PAHs-tolerant bacterial isolates from the highly-impacted environment often exhibited strong resistance to metals and antibiotics as well (Ben Said et al., 2008a; Mathe et al., 2012). However, it is unknown how many types of ARGs could be enriched in the soils by the PAHs. The ongoing rapid development of sequencing-based descriptive metagenomic approaches is providing the means to analyze the occurrence and abundance of enormous well-recognized ARGs in the soils that could not be solved previously by culture-based strategies, conventional PCR technique specific to several ARGs, and combination of them (Li et al., 2015b; Monier et al., 2011). Metagenomic profiling of ARGs in the PAHs-contaminated soils could help to better understand the emergence and diversity of ARGs under PAH stresses. Moreover, ARGs could be amplified in the environment under elevated selective stress via two distinct pathways, viz. the overgrowth of ARGs-carrying indigenous bacteria, or horizontal gene transfer of ARGs between bacteria using mobile gene elements (MGEs) as the carriers (e.g., plasmids, transposons and integrons) (Engelstadter et al., 2016; Knapp et al., 2008; Wang et al., 2015). Therefore, analysis of bacterial community and identification of genetic location of ARGs could be beneficial to gain more insights about how ARGs are accumulated in the PAHs-contaminated soils.

The overall objective of the present study was to delineate comprehensive profile of ARGs (that is to say, the diversity and relative abundance of ARGs) enriched in the

PAHs-contaminated soils using the metagenomic method. Further, analysis of the bacterial community structure and genetic locations of ARGs in the PAHs-contaminated soils were carried out to elucidate the underlying mechanisms for ARGs being enriched in the PAHs-contaminated soils via the changes of bacterial community or alternatively via horizontal transfer of genes between bacterial hosts.

2. Material and methods

2.1 Study area and sampling

Soil samples were collected from a petrochemical plant located at Guangzhou, China in July 2015 (N23.123, E113.4771). These samples were taken from three regions that represented different contamination degree of PAHs (the garden at the front gate (Site P1), tank storage area (Site P2), and wastewater pond (Site P3)). At each of sampling regions, the top 5-cm layer of soils were taken from 5 different positions using a sampling shovel, and were thoroughly mixed to avoid heterogeneous difference caused by a single time of sampling. All soil samples were immediately stored in sterilized polyethylene plastic bags and delivered back to the lab. A part of fresh soils collected at Site P3 was used to screen and isolate culturable bacterial strains with high tolerance to PAHs because this site was the most polluted by PAHs among the three sampling sites. The rest samples were frozen at -20 °C until DNA extraction and analysis of PAHs.

In the present study, an active sludge sample was collected from a sewage treatment plant in Shatin, Hong Kong in July 2007 (Yang et al., 2013), a swine feces sample was taken from a 7-8 month-old pig in a swine feedlot in Zhuhai, China in January 2013 (Li et al., 2015b), and a sediment sample was collected from an aquaculture pond in Shantou, China in December 2015. The above three samples that represented typical pollution sources of ARGs due to wide use of anthropogenic antibiotics were used as the comparisons with the soils collected from petrochemical plant. Moreover, sequencing datasets of six soil samples contaminated with oil spills were downloaded from the MG-RAST (Table S1) (Supplementary material).

2.2 Analysis of PAHs

Sixty PAHs in the soils were analyzed using the method reported in our previous study (Yuan et al., 2015). Briefly, PAHs were extracted from the soils using accelerated solvent extraction (ASE 350, Dionex) according to US EPA 3545a procedure. Approximately 10 g of dry and ground soils were mixed with certain diatomaceous earth, and then added into extraction cells. The m-Terphenoyl (200 mL, 1.0 mg L⁻¹), as surrogate standard, was spiked into each of samples before PAH extraction. Extraction cells were heated to 100 °C, pressurized with dichloromethane and acetone mixture (1:1, v/v), and maintained for 10 min. Subsequently, the extracts were concentrated using a rotary evaporator. The final volume was adjusted to 1.0 mL with acetone, and then internal standards (fluoranthene-d10, pyrene-d10, and chrysene-d12) were spiked at the concentration of 0.1 mg g⁻¹. Finally, the concentrations of PAHs were determined using gas chromatography coupled with mass spectrometry (GC-MS) in electron impact mode.

The selected ions, limit of detections (LOD), limit of quantification (LOQ), recoveries and relative standard deviation (RSD) of PAHs were presented in Table S2. The recoveries of PAHs were in the range of 64.2% to 132.3%. The RSD of all PAHs was lower than 12.9%, with an exception for 2,3+1,4-Dimethyl naphthalene (24.8%). The limit of quantification (LOQ) with a signal-to-noise ratio (S/N) of 10 was ranged from 0.03 to 5.83 ng g⁻¹.

2.3 Isolation of culturable bacteria from the PAHs-contaminated soils

Isolation of PAHs-tolerant bacteria from the soils was carried out according to the following procedure. The soils collected from Site P3 were chosen for isolation of bacteria because this site was the most polluted by PAHs among all three sites. Approximately 5 g of soils collected from Site P3 nearby wastewater ponds were added into 50 mL of sterile sodium pyrophosphate solution (2.8 g L⁻¹), and resulted solution was shaken overnight on a horizontal rotator with 150 rpm at 28 °C. An 5-mL aliquot of bacterial culture was mixed with 45 mL minimal salt medium (MSM)

((NH₄)₂SO₄, 1.0 g; MgSO₄·7H₂O, 0.2 g; FeSO₄·7H₂O, 0.01 g; CaCl₂, 0.1 g; K₂HPO₄·3H₂O 8.22 g and KH₂PO₄ 1.91 g *per* L, pH 7.2) containing 100 mg L⁻¹ of 1-methyl phenanthrene (1-MP). Bacterial culture (0.5 mL) was inoculated to 50 mL of fresh MSM with 100 mg L⁻¹ of 1-MP every 2 weeks. After the above inoculation was repeated 7 times, culturable bacterial consortium from the soils of Site P3 was obtained.

2.4 Testing antibiotic resistance of bacterial isolates

Bacterial consortium isolated from the soils of Site P3 was daubed on the Luria–Bertani (LB) agar plates (1.0% tryptone, 0.5% yeast extract, 1.0% NaCl, 2.0% agar). Ten bacterial colonies were picked up from the plates for antibiotic resistance tests. Antibiotic resistance was tested on Mueller-Hinton agar using the disc diffusion method according to the guidelines of the Clinical and Laboratory Standards Institute (CLSI, 2007). Antibiotic disc diffusion assays were performed with the following antibiotics commonly used in human medicine, including sulfonamides (trimethoprim/sulfamethoxazole), fluoroquinolones (ciprofloxacin, moxifloxacin and levofloxacin), *beta*-lactams (ampicillin, cefepime, ceftazidime, ceftriaxone, cefoxitin, cefazolin, imipenem and penicillin), glycopeptides (teicoplanin), nitrofurantoin, and macrolide (erythromycin). The strains of *Escherichia coli* ATCC 25922 and *Staphylococcus aureus* ATCC25923 were used as the quality controls.

2.5 Analysis of 16S rRNA genes of PAHs-tolerant bacterial isolates

Genomic DNA of bacterial isolates was extracted using DNA extraction kit (Omega Bio-tek). The 16S rRNA genes were amplified using a pair of universal primers 27F (5'-AGAGTTTGATCCTGGCTCAG-3') and 1492R (5'-TACGGTTACCTTACGACTT-3') under the subsequent conditions: an initial denaturation of 4 min for 94 °C, followed by 32 cycles of denaturation for 60 s at 94 °C, annealing for 30 s at 56 °C and elongation for 90 s at 72 °C, and a final extension for 10 min at 72 °C. PCR products purified with gel extraction kit (TaKaRa) were ligated on the pEAST-T1 vector (Transgen), and were transformed into competent

cells. The cells containing 16S rRNA gene were selected. The 16S rRNA genes were amplified using the primers M13F (5'-TGTAACGACGGCCAGT-3') and M13R (5'-CAGGAAACAGCTATGACC-3'), and then were sequenced using the dideoxynucleotide chain-termination method by a DNA Analyzer (3730XL, Applied Biosystems). Sequencing results of 16S rRNA genes of 6 PAHs-tolerant bacterial isolates were presented in the supplementary material.

2.6 DNA extraction and high-throughput sequencing

DNA extraction was conducted using the PowerSoil® DNA Isolation Kit (MO BIO, CA) according to the manufacturer's manuals. DNA was multiple times extracted from each of samples to minimize possible heterogeneity in the samples and avoid possible bias from the process of extracting a single DNA sample. The purity and concentration of the DNA were analyzed using a Thermo Scientific NanoDrop1000 Spectrophotometer. Combined DNA from the preparation of multiple times was delivered to the Annoroad Gene Technology's headquarter (Beijing, China) for sequencing. Approximately 5 µg of DNA was sheared into ~200 nt fragments, and the overhangs of DNA fragments were end-repaired using T4 DNA polymerase, Klenow Fragment and T4 Polynucleotide Kinase. Adapters were ligated to the ends of the DNA fragments, and then the desired fragments were purified using gel-electrophoresis. Target DNA fragments were amplified by PCR, with index tags being introduced into the adapter to construct DNA libraries. The qualified DNA libraries were used for subsequent sequencing by Illumina HiSeq 2500. The data size of each of samples was approximately 3.0 Gb. Raw sequencing data was deposited on the GenBank and MG-RAST, and accession number was presented in Table S1.

2.7 Bioinformatic analysis

The raw reads (125 nt in length) with a quality score of 30 accounted for greater than 90% in all sequencing datasets of samples, and the clean rates were higher than 90%. The raw reads were trimmed to discard low-quality reads that contained ambiguous nucleotides or a quality score of lower than 30. The 16S rRNA gene-like

sequences were identified by blasting our datasets against the SILVA Small Subunit (SSU) database (version 10.4) with an *E*-value threshold of 10^{-20} through a local BLASTN program (Altschul et al., 1997; Mackelprang et al., 2011), and thereafter BLAST results were used to characterize bacterial community structure in the soils by MEGAN (version 4.70.4) with the Lowest Common Ancestor (LCA) algorithm, where the absolute cutoff was a BLAST bitscore of 50 and the relative cutoff was 10% of the top 50 hits (Huson et al., 2007). The BLAST programs (Linux Release 2.2.29) were obtained from the NCBI, and run on the local server to analyze our sequencing datasets against the databases available. An antibiotic resistance genes database (ARDB) was used to search ARG-like sequences by the local BLASTX program (Liu and Pop, 2009), with the best BLASTX hit being a similarity of higher than 90% and an alignment length of at least 25 amino acids (aa) (Yang et al., 2013). The local BLASTN program was employed to align our sequencing data against the database of plasmids that was developed according to the plasmid sequences deposited into the NCBI RefSeq database (Yang et al., 2013). The plasmid-like sequences were determined by the alignments with an identity of above 95% over a sequence length of greater than 90 nt (Zhang et al., 2011).

After quality filtering, sequencing datasets were also assembled using the IDBA (Version 1.1.0 released for Linux) with a Kmer value of 63 to 123 and a step of 20 (Peng et al., 2010). Assembled contigs with a length longer than 500 nt were selected to predict open read frames (ORFs) using Prodigal (Version 2.6.3) (Hyatt et al., 2010). An ORF was identified as an ARG-like sequence if the best BLASTX alignment was not lower than 80% identity with the query sequence coverage of at least 70%. This analytical procedure of bioinformatics was evaluated to be precise as high as 99.1% (Hu et al., 2013).

2.8 Statistical analysis

Nonparametric paired sample sign tests (NPSSTs), principal component analysis (PCA) and linear regression were carried out using the SPSS® for Windows Release

16.0 (SPSS Inc. U.S.). The NPSSTs were performed to explore the differences in the relative abundance of bacterial phyla, ARGs and plasmids among the samples. Differences at the $p < 0.05$ level (99% confidence interval) between the samples were considered statistically significant. With respect to PCA and linear regression, the data was checked to meet the normal distribution. PCA method was used to characterize and cluster the samples and ARG subtypes, where the relative abundance of resistance types and ARG subtypes were used as data inputs. Linear correlations between Proteobacterial phylum and ARGs were also completed using the absolute abundance of Proteobacterial phylum and ARGs in the samples.

3. Results and discussion

3.1 Concentrations of PAHs in the soils

A total 49 of 60 investigated PAHs were detected in the soils collected from petrochemical plant using GC-MS technique (Table S1). Fig. S1 showed that total concentration of investigated PAHs in the soils at Site P3 was highest among the three sampling sites (up to $529 \mu\text{g g}^{-1}$ (dry weight)), which was at least 50 times higher than those at Site P1 and P2. The levels of PAHs are generally low in the environment (Niepceron et al., 2013), whereas the total concentration of them are extremely high as several hundred $\mu\text{g g}^{-1}$ in some contamination accidents, e.g., oil spill (Brandt et al., 2002; Tian et al., 2014; Wang et al., 2014). Analysis of PAHs demonstrated that our samples collected from petrochemical plant represented different pollution levels of PAHs from mild degree to extremely high degree as several hundred $\mu\text{g g}^{-1}$, which could help to reveal the relationships between PAH contamination and bacterial community or ARGs in the soils.

3.2 Bacterial community structures in the soils

The 16S rRNA genes were identified by aligning our sequencing datasets against the SSU database, and further bacterial community structures in the soils were analyzed using MEGAN. Fig. 1 shows relative percentages of bacterial phyla in the soils and in the bacterial culture isolated from the soils of Site P3 with the highest levels of PAHs. Bacterial community structure in the soils of Site P3 was significantly

differentiated from those at Site P1 and P2 ($p < 0.05$). In particular, relative abundance of Proteobacteria in the soils at Site P3 was substantially higher than those at Site P1 and P2. Approximately 97% of culturable strains isolated from the soils of Site P3 were also assigned to Proteobacteria. Moreover, analysis of 16S rRNA genes of 6 PAHs-tolerant bacterial isolates demonstrated that they all belonged to the Proteobacteria.

PAHs released from anthropogenic sources (e.g., oil spill, industrial runoff, *etc.*) exert deleterious effects on indigenous bacterial communities in the soils (de Menezes et al., 2012; Yang et al., 2014; Zhang et al., 2010). Bacterial population in the reed rhizosphere soils was considerably shrunk due to the irrigation with oil-polluted water, with Proteobacteria being predominant (Tian et al., 2014). Thus GammaProteobacteria was proposed as a suitable bio-indicator for evaluating PAH impacts on bacterial communities in the agricultural soils (Niepceyron et al., 2013). From perspective of metaproteomics, proteobacterial proteins in the soils were largely induced by petroleum pollution in comparison to the controls (Bastida et al., 2016). All facts have evidenced that Proteobacterial strains in the soils are enriched by PAHs. It was probably attributed to that Proteobacteria may own some strategies to resist toxic PAHs, e.g., most of PAH bacterial degraders were phylogenetically assigned to Proteobacteria (Tauler et al., 2016; Yi et al., 2016). Meanwhile, relative abundance of Proteobacteria in the soils with a poor PAH degradation ability was significantly increased with PAH supplementation (Ren et al., 2015). In this case, the selection of PAHs on Proteobacteria implies that these bacteria probably have other PAH tolerance mechanisms different from the bio-degradation.

3.3 Abundance and resistance types of ARGs in the PAHs-contaminated soils

ARGs were categorized according to resistance types that were referred as target antibiotics associated with their resistance. ARG abundance was obtained from dividing the number of ARGs-like reads by the total number of reads in the same sample. Fig. 2A showed that total abundance of ARGs in the soils of Site P3 was

approximately 15 times higher than those at Site P1 and P2, indicating that PAH contamination could result in the accumulation of ARGs in the soils. Total ARG abundance in the culturable bacteria was comparable with that in the whole soil metagenome of Site P3. In comparison to representative pollution sources of ARGs (e.g., active sludge (AS) and aquaculture sediments (AQS)), ARGs in the soils of Site P3 and corresponding culturable bacteria were even more abundant.

Fig. 2B shows relative percentages of resistance types of ARGs in the PAHs-contaminated soils and representative pollution sources of ARGs. ARGs in the PAHs-contaminated soils were related to more than 10 groups of antibiotics, including aminoglycosides, *beta*-lactams, chloramphenicols, fluoroquinolones, macrolides, polypeptides, tetracyclines, *etc.* More than half of ARGs were only associated with a single group of antibiotics such as chloramphenicols, polypeptides, fluoroquinolones. Herein, ARGs specific to chloramphenicols were most abundant among all resistance types. In addition, approximately 50% of ARGs in the PAHs-contaminated soils were able to simultaneously encode the resistance to different groups of antibiotics, and these ARGs took up more than 80% of the resistance to commonly-used antibiotics including aminoglycosides, *beta*-lactams and macrolides. Similar resistance profiles of ARGs were also observed in the oil spills-contaminated soils, as shown in Fig. S2. Discerned from those in the PAHs-contaminated soils, resistance profiles in representative pollution sources of ARGs (Swine feces (SF), AS, and AQS) were featured with higher abundance of ARGs specific to a single anthropogenic antibiotic, such as tetracyclines, sulfonamides, *beta*-lactams, *etc.*

The samples were further clustered using a PCA approach according to resistance profiles of ARGs. Fig. 3 showed that the first two principal components accounted for about 72.3% of the total variance in antibiotic resistance types among all samples. Representative pollution sources of ARGs were significantly differentiated from the soils collected from petrochemical plant and oil spills-contaminated area. Next, P3,

A5 and A6 were grouped together mainly due to high abundance of ARGs and similar antibiotic resistance profiles in these soils, which were distinctly separated from other soils containing the low abundance of ARGs.

Resistance profiles of ARGs are highly dependent on selective stresses that bacterial communities need to overcome. In representative pollution sources of ARGs, e.g., sewage treatment plants and husbandry operations (Chee-Sanford et al., 2009; Larson, 2007), most of ARGs and ARB derived from human and animal digestive tracks, where anthropogenic antibiotics are administered, enter the sewage and the receiving environments. Consequently, resistance profiles of ARGs in these pollution sources were characterized by the dominance of resistance types specific to a single group of anthropogenic antibiotics over others, which could effectively protect bacteria from these antibiotics. In the present study, PAHs were considered as the main factor for shaping bacterial communities and their antibiotic resistance profiles in the soils. Significant correlation between antibiotic resistance and hydrocarbon degradation potential was observed for some bacterial strains (Mathe et al., 2012). However, there was lack of comprehensive knowledge on the ARGs carried by PAHs-tolerant bacteria in the soils. Assumed that ARGs enriched in the PAHs-contaminated soils might confer the resistance to PAHs, the genes related to aromatic antibiotics (such as acriflavine, chloramphenicols, and fluoroquinolones) could be more preferentially developed and amplified than others. Our results indeed demonstrated that ARGs associated with aromatic antibiotics accounted for more than 70% of the total ARGs in both the PAHs-contaminated soils and culturable bacterial consortium, and their abundance was increased almost by 12 times in comparison to the less-contaminated soils. Genetic recombination also corroborated the underlying relationships between PAH resistance and aromatic antibiotic resistance, e.g., the disruption of PAH resistance gene in *Pseudomonas fluorescens* cLP6a did reduce its tolerance to chloramphenicol (Hearn et al., 2003).

3.4 ARG subtypes in the PAHs-contaminated soils

Fig. 4A showed that the diversity of ARG subtypes in the soils of Site P3 was much higher than those in other soils. The amount of ARG subtypes hosted by culturable bacteria was 3 times lower than that in soil metagenome at Site P3, attributing to that a significant fraction of bacterial strains in the soils could not be incubated. A total of 93 ARG subtypes were identified in the samples, and 29 subtypes thereof concurrently existed in all samples (Fig. 4B). Twenty-one of the 29 concurrent ARG subtypes could encode the resistance to aromatic antibiotics, and relative percentage of concurrent subtypes in each of samples was greater than 75.0%. Moreover, about 86% of concurrent ARG subtypes could encode efflux pumps.

Furthermore, PCA was also carried out to categorize ARG subtypes according to their relative abundance in and between the samples (Fig. S3), where 23 ARG subtypes with the ratio of gene abundance at Site P3 to that at Site P1 larger than 2.0 were chosen. The first two principal components accounted for 96.0% of the total variance in ARG subtypes among all samples. Fig. S3 demonstrated that 8 major ARG subtypes (*acrA*, *acrB*, *amrB*, *bacA*, *ceoB*, *macB*, *mexB* and *tolC*), each of which accounted for at least 4% of the total ARGs in the PAHs-contaminated soils, were significantly distributed along Component 1. It indicated that *amrB*, *bacA* and *ceoB* were more abundant in the genome of culturable bacteria whereas *macB*, *mexB* and *tolC* were more enriched in the whole metagenome of PAHs-contaminated soils.

Herein, six ARG subtypes (*acrA*, *acrB*, *amrB*, *ceoB*, *mexB* and *tolC*) could encode the resistance to aromatic antibiotics via efflux pumps (Fischer and Kandt, 2013; Liu and Pop, 2009; Nair et al., 2004), and comprised of 80% of aromatic antibiotic resistance genes in the soils at Site P3. In contrast to the less-contaminated soils with PAHs, above-mentioned subtypes related to aromatic antibiotics were enriched by about 14 times in the soils of Site P3. Completed sequences of the above ARG subtypes were also identified in assembling contigs from the soils and culturable bacteria. These ARG subtypes were ubiquitous in both minimally and highly human-impacted settings, e.g., Tibetan plateau (Chen et al., 2016), sewage treatment

plant (Yang et al., 2013), and manured soils (Fang et al., 2015).

Various strategies are adopted by bacteria to cope with PAHs. Bacteria are able to degrade PAHs into less toxic metabolites using a group of intracellular enzymes, and even into the nutrients for supporting bacterial growth when PAHs are provided as the sole carbon source (Kanehisa et al., 2008; Kweon et al., 2011). Nevertheless, different strategies might be employed by bacterial population with a poor PAH degradability (Ren et al., 2015). Efflux pumps that are capable of expelling toxicants out of cells are the popular mechanism for microbes to fight against various pollutants, such as antibiotics, metals, and PAHs (Adebusuyi and Foght, 2011; Rosen, 1996; Ryan et al., 2001). Our results exhibited that ARG subtypes encoding aromatic antibiotic transporters via efflux pumps were highly abundant in the PAHs-contaminated soils, and were mainly carried by the strains belonging to Proteobacterial phylum. It probably implies that efflux proteins expressed by these ARGs are likely essential to bacterial survivals from PAH pollution. The flexibility of efflux transporter binding to multiple substrates eases exporting different xenobiotics from the cells, e.g., an efflux system that consists of *acrA*, *acrB* and *tolC* is able to transport acriflavine, aminoglycosides, *beta*-lactams, tetracyclines and macrolides (Liu and Pop, 2009). With unusually high degree of substrate promiscuity, efflux pumps are considered as one of the most effective mode to deal with toxic pollutants in terms of maintenance cost and functional versatility (Lomovskaya et al., 2007; Wright, 2007). Structural similarities among the substrates reasonably determine if these compounds could be pumped out by the same efflux system (Nikaido, 1994). Unexceptionally, PAHs and some aromatic antibiotics (e.g., chloramphenicols) could both act as the substrates of an active efflux system (EmhABC) owned by bacteria (Hearn et al., 2003). Therefore, analysis of ARG subtypes suggested that PAHs might potentially act as the substrates of efflux transporters (e.g., *acrA*, *acrB*, *amrB*, and *tolC*) related to antibiotics.

3.5 Potentials of horizontal transfer of ARGs in the PAHs-contaminated soils

ARG-like reads were extracted from all sequencing datasets, and then blasted

against the database of plasmids for indicating the potential for ARGs being transferred among bacteria. Fig. 5 showed that both abundance and relative percentage of plasmid-related ARGs in the PAHs-contaminated soils were substantially lower than those in representative pollution sources of ARGs. In the soils collected from petrochemical plants, ARGs potentially carried by plasmids only took up lower than 0.5% of the total ARGs. Total abundance of plasmids-related ARGs was almost 180 times higher in the SF than in the soils of Site P3 while total abundance of ARGs was comparable between them.

ARGs can be disseminated between bacterial hosts via horizontal gene transfers in the human-impacted environments under elevated stresses (Gootz, 2010). Thus it was often observed in various environmental settings that ARGs were significantly correlated with MGEs in terms of both diversity and abundance because MGEs play as the vehicles in the lateral transfer of ARGs (Chen et al., 2014; Chen et al., 2013b; Luo et al., 2010). Our results demonstrated that a small fraction of ARGs in the PAHs-contaminated soils were presented in the plasmids. It implies that most of ARGs enriched in the PAHs-contaminated soils have the low possibility to be horizontally transferred between bacteria. On the contrary, potentially plasmids-carrying ARGs are enriched in representative pollution sources of ARGs. It could be attributed to the consequence of the horizontal transfer of ARGs from the overuse of anthropogenic antibiotics, and undoubtedly continues to ease horizontally acquiring of ARGs by bacteria *vice versa* (Chen et al., 2016).

3.6 Mechanisms proposed for accumulation of ARGs in the PAHs-contaminated soils

ARGs are able to horizontally spread between bacteria via transformation or conjugation under antibiotic stresses, thereby ARGs being enriched in the human-impacted environment (Cottell et al., 2014; Turner et al., 2014). Such gene transfers require plasmids as the shuttling carriers between bacteria (Wang et al., 2015). For instance, the concurrence of metal and antibiotic resistance genes in the plasmids facilitates the dissemination of these genes under elevated stresses caused by

either heavy metals or antibiotics (Li et al., 2015a; Zhai et al., 2016). However, our results evidenced that most of ARGs in the PAHs-contaminated soils were inherently located at the chromosomes of PAHs-tolerant bacteria rather than plasmids. Furthermore, non-covalent binding of PAHs with genetic bases may significantly inhibit the *in vitro* horizontal transfer of ARGs between bacteria via plasmid transformation (Kang et al., 2015). Therefore, all facts suggested that plasmids-mediated gene transfer was not the main mechanism for greatly enriching of ARGs in the PAHs-contaminated soils.

Alternatively, it could be conceived that ARGs might be accumulated in the soils through the selection of PAHs on the ARGs-carrying bacteria. In the PAHs-contaminated soils, PAH degrading and tolerant bacteria were predominant over others, and further characterization of these strains exhibited their strong abilities to resist antibiotics (Ben Said et al., 2008a; Mathe et al., 2012). Our assays also demonstrated that most of bacterial isolates from the PAHs-contaminated soils were not susceptible to different groups of antibiotics including fluoroquinolones, *beta*-lactams, nitrofurans, glycopeptides, *etc* (Table S3). It was reasonably considered that ARGs hosted by PAHs-tolerant bacteria, with efflux pump-encoding ARGs being predominant, were genetic basis for the resistance to these antibiotics. Analysis of 16S rRNA genes evidenced that most of the 20 selected PAHs-tolerant bacteria were taxonomically assigned to Proteobacterial phylum (Ben Said et al., 2008b). In the present study, Proteobacteria was always extremely abundant in the PAHs-contaminated soils, as well as culturable bacteria thereof. Thus it deserved to explore the relationship between Proteobacteria and ARGs. The total abundance of ARGs and those encoding efflux pumps was both significantly correlated with that of Proteobacteria in the soils (Fig. 6 and S4). Approximately 74% of ARG subtypes associated with efflux pumps in the ARDB database are found to be carried by Proteobacteria. Molecular evidence is needed that ARGs associated with efflux pumps also play the key roles in bacterial tolerance to PAHs. However, it was

obvious that Proteobacteria selected by PAHs in the soils simultaneously resulted in enriching of ARGs carried by them, most of which represented the mechanism of efflux pump.

4. Conclusions

Various ARGs were enriched in the PAHs-contaminated soils, which were characterized with the predominance of aromatic antibiotic resistance genes encoding efflux pumps. This study heralded that PAHs could act as selective stresses for ARGs, and could contribute significantly to ARG pollution in the human-impacted environment. Resistance profiles of ARGs selected by PAHs may be completely differentiated from those of representative ARG pollution sources that are arisen from the overuse of anthropogenic antibiotic. Our results provided a supplementary insight into the prevalence of ARGs in modern environment.

Acknowledgments

This work was supported by the Provincial Natural Science Foundation of Guangdong, China (2014A030313195), Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences (KF2014-01), and the National Science Foundation of China (21625703, 21677183, 51579253).

References

- Adebusuyi, A.A., Foght, J.M., 2011. An alternative physiological role for the EmhABC efflux pump in *Pseudomonas fluorescens* cLP6a. *BMC Microbiol.* 11, 252.
- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST. *Nucleic Acids Res.* 25, 3389-3402.
- Anjum, R., Krakat, N., 2016. Detection of Multiple Resistances, Biofilm Formation and Conjugative Transfer of *Bacillus cereus* from Contaminated Soils. *Curr. Microbiol.* 72, 321-328.
- Baker-Austin, C., Wright, M.S., Stepanauskas, R., McArthur, J.V., 2006. Co-selection of antibiotic and metal resistance. *Trends Microbiol.* 14, 176-182.
- Bastida, F., Jehmlich, N., Lima, K., Morris, B.E.L., Richnow, H.H., Hernandez, T., von Bergen, M., Garcia, C., 2016. The ecological and physiological responses of the microbial community from a semiarid soil to hydrocarbon contamination and its bioremediation using compost amendment. *J. Proteomics* 135, 162-169.
- Ben Said, O., Goni-Urriza, M.S., El Bour, M., Dellali, M., Aissa, P., Duran, R., 2008a. Characterization of aerobic polycyclic aromatic hydrocarbon-degrading bacteria from Bizerte lagoon sediments, Tunisia. *J. Appl. Microbiol.* 104, 987-997.
- Ben Said, O., Goni-Urriza, M.S., El Bour, M., Dellali, M., Aissa, P., Duran, R., 2008b. Characterization of aerobic polycyclic aromatic hydrocarbon-degrading bacteria from Bizerte lagoon sediments, Tunisia. *J. Appl. Microbiol.* 104, 987-997.
- Bosch, C., Andersson, A., Krusa, M., Bandh, C., Hovorkova, I., Klanova, J., Knowles, T.D., Pancost, R.D., Evershed, R.P., Gustafsson, O., 2015. Source Apportionment of Polycyclic Aromatic Hydrocarbons in Central European Soils with Compound-Specific Triple Isotopes ($\delta(13)\text{C}$, $\delta(14)\text{C}$, and $\delta(2)\text{H}$). *Environ. Sci. Technol.* 49, 7657-7665.
- Brandt, C.A., Becker, J.M., Porta, A., 2002. Distribution of polycyclic aromatic hydrocarbons in soils and terrestrial biota after a spill of crude oil in Trecate, Italy. *Environ. Toxicol. Chem.* 21, 1638-1643.
- Chee-Sanford, J.C., Mackie, R.I., Koike, S., Krapac, I.G., Lin, Y.F., Yannarell, A.C., Maxwell, S., Aminov, R.I., 2009. Fate and Transport of Antibiotic Residues and Antibiotic Resistance Genes following Land Application of Manure Waste. *J. Environ. Qual.* 38, 1086-1108.
- Chen, B., Liang, X., Huang, X., Zhang, T., Li, X., 2013a. Differentiating anthropogenic impacts on ARGs in the Pearl River Estuary by using suitable gene indicators. *Water Res.* 47, 2811-2820.
- Chen, B., Liang, X., Nie, X., Huang, X., Zou, S., Li, X., 2014. The role of class I integrons in the dissemination of sulfonamide resistance genes in the Pearl River and Pearl River Estuary, South China. *J. Hazard. Mater.* 282, 61-67.
- Chen, B., Yuan, K., Chen, X., Yang, Y., Zhang, T., Wang, Y., Luan, T., Zou, S., Li, X., 2016. Metagenomic Analysis Revealing Antibiotic Resistance Genes (ARGs) and Their Genetic Compartments in the Tibetan Environment. *Environ. Sci. Technol.* 50, 6670-6679.
- Chen, B.W., Yang, Y., Liang, X.M., Yu, K., Zhang, T., Li, X.D., 2013b. Metagenomic Profiles of Antibiotic Resistance Genes (ARGs) between Human Impacted Estuary and Deep Ocean Sediments. *Environ. Sci. Technol.* 47, 12753-12760.
- Choi, Y., Cho, Y.M., Gala, W.R., Luthy, R.G., 2013. Measurement and modeling of activated carbon performance for the sequestration of parent- and alkylated-polycyclic aromatic

525 hydrocarbons in petroleum-impacted sediments. *Environ. Sci. Technol.* 47, 1024-1032.

526 CLSI, 2007. Performance Standards for Antimicrobial Susceptibility Testing., Approved Standard
527 M100-S17. Clinical and Laboratory Standards Institute, Wayne, PA.

528 Cottell, J.L., Saw, H.T., Webber, M.A., Piddock, L.J., 2014. Functional genomics to identify the
529 factors contributing to successful persistence and global spread of an antibiotic resistance plasmid.
530 *BMC Microbiol.* 14, 168.

531 de Menezes, A., Clipson, N., Doyle, E., 2012. Comparative metatranscriptomics reveals
532 widespread community responses during phenanthrene degradation in soil. *Environ. Microbiol.* 14,
533 2577-2588.

534 Engelstadter, J., Harms, K., Johnsen, P.J., 2016. The evolutionary dynamics of integrons in
535 changing environments. *ISME J.* 10, 1296-1307.

536 Fang, H., Wang, H.F., Cai, L., Yu, Y.L., 2015. Prevalence of Antibiotic Resistance Genes and
537 Bacterial Pathogens in Long-Term Manured Greenhouse Soils As Revealed by Metagenomic
538 Survey. *Environ. Sci. Technol.* 49, 1095-1104.

539 Fischer, N., Kandt, C., 2013. Porter domain opening and closing motions in the multi-drug efflux
540 transporter AcrB. *Biochim. Biophys. Acta* 1828, 632-641.

541 Forsberg, K.J., Patel, S., Gibson, M.K., Lauber, C.L., Knight, R., Fierer, N., Dantas, G., 2014.
542 Bacterial phylogeny structures soil resistomes across habitats. *Nature* 509, 612-616.

543 Gootz, T.D., 2010. The global problem of antibiotic resistance. *Crit. Rev. Immunol.* 30, 79-93.

544 Hearn, E.M., Dennis, J.J., Gray, M.R., Foght, J.M., 2003. Identification and characterization of the
545 emhABC efflux system for polycyclic aromatic hydrocarbons in *Pseudomonas fluorescens* cLP6a.
546 *J. Bacteriol.* 185, 6233-6240.

547 Hu, Y., Yang, X., Qin, J., Lu, N., Cheng, G., Wu, N., Pan, Y., Li, J., Zhu, L., Wang, X., Meng, Z.,
548 Zhao, F., Liu, D., Ma, J., Qin, N., Xiang, C., Xiao, Y., Li, L., Yang, H., Wang, J., Yang, R., Gao,
549 G.F., Zhu, B., 2013. Metagenome-wide analysis of antibiotic resistance genes in a large cohort of
550 human gut microbiota. *Nat. Commun.* 4, 2151.

551 Huson, D.H., Auch, A.F., Qi, J., Schuster, S.C., 2007. MEGAN analysis of metagenomic data.
552 *Genome Res.* 17, 377-386.

553 Hyatt, D., Chen, G.L., Locascio, P.F., Land, M.L., Larimer, F.W., Hauser, L.J., 2010. Prodigal:
554 prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11,
555 119.

556 Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima,
557 S., Okuda, S., Tokimatsu, T., Yamanishi, Y., 2008. KEGG for linking genomes to life and the
558 environment. *Nucleic Acids Res.* 36, D480-484.

559 Kang, F.X., Hu, X.J., Liu, J., Gao, Y.Z., 2015. Noncovalent Binding of Polycyclic Aromatic
560 Hydrocarbons with Genetic Bases Reducing the in Vitro Lateral Transfer of Antibiotic Resistant
561 Genes. *Environ. Sci. Technol.* 49, 10340-10348.

562 Karatzas, K.A., Webber, M.A., Jorgensen, F., Woodward, M.J., Piddock, L.J., Humphrey, T.J.,
563 2007. Prolonged treatment of *Salmonella enterica* serovar Typhimurium with commercial
564 disinfectants selects for multiple antibiotic resistance, increased efflux and reduced invasiveness. *J.*
565 *Antimicrob. Chemother.* 60, 947-955.

566 Knapp, C.W., Dolfing, J., Ehler, P.A.I., Graham, D.W., 2010. Evidence of Increasing Antibiotic

567 Resistance Gene Abundances in Archived Soils since 1940. *Environ. Sci. Technol.* 44, 580-587.
 568 Knapp, C.W., Engemann, C.A., Hanson, M.L., Keen, P.L., Hall, K.J., Graham, D.W., 2008.
 569 Indirect evidence of transposon-mediated selection of antibiotic resistance genes in aquatic
 570 systems at low-level oxytetracycline exposures. *Environ. Sci. Technol.* 42, 5348-5353.
 571 Kweon, O., Kim, S.J., Holland, R.D., Chen, H., Kim, D.W., Gao, Y., Yu, L.R., Baek, S., Baek,
 572 D.H., Ahn, H., Cerniglia, C.E., 2011. Polycyclic aromatic hydrocarbon metabolic network in
 573 *Mycobacterium vanbaalenii* PYR-1. *J. Bacteriol.* 193, 4326-4337.
 574 Larson, E., 2007. Community factors in the development of antibiotic resistance. *Annu. Rev. Publ.*
 575 *Health* 28, 435-447.
 576 Li, A.D., Li, L.G., Zhang, T., 2015a. Exploring antibiotic resistance genes and metal resistance
 577 genes in plasmid metagenomes from wastewater treatment plants. *Front Microbiol.* 6, 1025.
 578 Li, B., Yang, Y., Ma, L., Ju, F., Guo, F., Tiedje, J.M., Zhang, T., 2015b. Metagenomic and network
 579 analysis reveal wide distribution and co-occurrence of environmental antibiotic resistance genes.
 580 *ISME J.* 9, 2490-2502.
 581 Liu, B., Pop, M., 2009. ARDB-Antibiotic Resistance Genes Database. *Nucleic Acids Res.* 37,
 582 D443-D447.
 583 Lo Giudice, A., Casella, P., Bruni, V., Michaud, L., 2013. Response of bacterial isolates from
 584 Antarctic shallow sediments towards heavy metals, antibiotics and polychlorinated biphenyls.
 585 *Ecotoxicology* 22, 240-250.
 586 Lomovskaya, O., Zgurskaya, H.I., Totrov, M., Watkins, W.J., 2007. Waltzing transporters and 'the
 587 dance macabre' between humans and bacteria. *Nat. Rev. Drug Discov.* 6, 56-65.
 588 Luo, Y., Mao, D.Q., Rysz, M., Zhou, D.X., Zhang, H.J., Xu, L., Alvarez, P.J.J., 2010. Trends in
 589 Antibiotic Resistance Genes Occurrence in the Haihe River, China. *Environ. Sci. Technol.* 44,
 590 7220-7225.
 591 Mackelprang, R., Waldrop, M.P., DeAngelis, K.M., David, M.M., Chavarria, K.L., Blazewicz, S.J.,
 592 Rubin, E.M., Jansson, J.K., 2011. Metagenomic analysis of a permafrost microbial community
 593 reveals a rapid response to thaw. *Nature* 480, 368-371.
 594 Mathe, I., Benedek, T., Tancsics, A., Palatinszky, M., Lanyi, S., Marialigeti, K., 2012. Diversity,
 595 activity, antibiotic and heavy metal resistance of bacteria from petroleum hydrocarbon
 596 contaminated soils located in Harghita County (Romania). *Int. Biodeterior. Biodegrad.* 73, 41-49.
 597 Monier, J.M., Demaneche, S., Delmont, T.O., Mathieu, A., Vogel, T.M., Simonet, P., 2011.
 598 Metagenomic exploration of antibiotic resistance in soil. *Curr. Opin. Microbiol.* 14, 229-235.
 599 Nair, B.M., Cheung, K.J., Jr., Griffith, A., Burns, J.L., 2004. Salicylate induces an antibiotic efflux
 600 pump in *Burkholderia cepacia* complex genomovar III (*B. cenocepacia*). *J. Clin. Invest.* 113,
 601 464-473.
 602 Nathan, C., 2004. Antibiotics at the crossroads. *Nature* 431, 899-902.
 603 Nesme, J., Simonet, P., 2015. The soil resistome: a critical review on antibiotic resistance origins,
 604 ecology and dissemination potential in telluric bacteria. *Environ. Microbiol.* 17, 913-930.
 605 Niepceon, M., Martin-Laurent, F., Crampon, M., Portet-Koltalo, F., Akpa-Vinceslas, M., Legras,
 606 M., Bru, D., Bureau, F., Bodilis, J., 2013. GammaProteobacteria as a potential bioindicator of a
 607 multiple contamination by polycyclic aromatic hydrocarbons (PAHs) in agricultural soils. *Environ.*
 608 *Pollut.* 180, 199-205.

Nikaido, H., 1994. Prevention of drug access to bacterial targets: permeability barriers and active efflux. *Science* 264, 382-388.

Pehrsson, E.C., Tsukayama, P., Patel, S., Mejia-Bautista, M., Sosa-Soto, G., Navarrete, K.M., Calderon, M., Cabrera, L., Hoyos-Arango, W., Bertoli, M.T., Berg, D.E., Gilman, R.H., Dantas, G., 2016. Interconnected microbiomes and resistomes in low-income human habitats. *Nature* 533, 212-216.

Peng, Y., Leung, H.C.M., Yiu, S.M., Chin, F.Y.L., 2010. IDBA - A Practical Iterative de Bruijn Graph De Novo Assembler. *Research in Computational Molecular Biology, Proceedings* 6044, 426-440.

Pruden, A., Pei, R., Storteboom, H., Carlson, K.H., 2006. Antibiotic resistance genes as emerging contaminants: studies in northern Colorado. *Environ. Sci. Technol.* 40, 7445-7450.

Ren, G., Ren, W., Teng, Y., Li, Z., 2015. Evident bacterial community changes but only slight degradation when polluted with pyrene in a red soil. *Front Microbiol.* 6, 22.

Rosen, B.P., 1996. Bacterial resistance to heavy metals and metalloids. *J. Biol. Inorg. Chem.* 1, 273-277.

Ryan, B.M., Dougherty, T.J., Beaulieu, D., Chuang, J., Dougherty, B.A., Barrett, J.F., 2001. Efflux in bacteria: what do we really know about it? *Expert Opin. Inv. Drug* 10, 1409-1422.

Tauler, M., Vila, J., Nieto, J.M., Grifoll, M., 2016. Key high molecular weight PAH-degrading bacteria in a soil consortium enriched using a sand-in-liquid microcosm system. *Appl. Microbiol. Biotechnol.* 100, 3321-3336.

Tian, W.J., Zhao, Y.G., Sun, H.M., Bai, J., Wang, Y.M., Wu, C.L., 2014. The effect of irrigation with oil-polluted water on microbial communities in estuarine reed rhizosphere soils. *Ecol. Eng.* 70, 275-281.

Turner, P.E., Williams, E.S., Okeke, C., Cooper, V.S., Duffy, S., Wertz, J.E., 2014. Antibiotic resistance correlates with transmission in plasmid evolution. *Evolution* 68, 3368-3380.

Wang, Q., Mao, D., Luo, Y., 2015. Ionic Liquid Facilitates the Conjugative Transfer of Antibiotic Resistance Genes Mediated by Plasmid RP4. *Environ. Sci. Technol.* 49, 8731-8740.

Wang, W.P., Zhong, R.Q., Shan, D.P., Shao, Z.Z., 2014. Indigenous oil-degrading bacteria in crude oil-contaminated seawater of the Yellow sea, China. *Appl. Microbiol. Biotechnol.* 98, 7253-7269.

Wenzel, R.P., Edmond, M.B., 2000. Managing antibiotic resistance. *N. Engl. J. Med.* 343, 1961-1963.

Wright, G.D., 2007. The antibiotic resistome: the nexus of chemical and genetic diversity. *Nat. Rev. Microbiol.* 5, 175-186.

Yang, S., Wen, X., Zhao, L., Shi, Y., Jin, H., 2014. Crude oil treatment leads to shift of bacterial communities in soils from the deep active layer and upper permafrost along the China-Russia Crude Oil Pipeline route. *PLoS One* 9, e96552.

Yang, Y., Li, B., Ju, F., Zhang, T., 2013. Exploring Variation of Antibiotic Resistance Genes in Activated Sludge over a Four-Year Period through a Metagenomic Approach. *Environ. Sci. Technol.* 47, 10197-10205.

Yi, X.H., Jing, D.D., Wan, J., Ma, Y., Wang, Y., 2016. Temporal and spatial variations of contaminant removal, enzyme activities, and microbial community structure in a pilot horizontal subsurface flow constructed wetland purifying industrial runoff. *Environ. Sci. Pollut. Res. Int.* 23,

8565-8576.

Yuan, K., Wang, X., Lin, L., Zou, S., Li, Y., Yang, Q., Luan, T., 2015. Characterizing the parent and alkyl polycyclic aromatic hydrocarbons in the Pearl River Estuary, Daya Bay and northern South China Sea: influence of riverine input. *Environ. Pollut.* 199, 66-72.

Zhai, Y., He, Z., Kang, Y., Yu, H., Wang, J., Du, P., Zhang, Z., Hu, S., Gao, Z., 2016. Complete nucleotide sequence of pH11, an IncHI2 plasmid conferring multi-antibiotic resistance and multi-heavy metal resistance genes in a clinical *Klebsiella pneumoniae* isolate. *Plasmid* 86, 26-31.

Zhang, T., Zhang, X.X., Ye, L., 2011. Plasmid Metagenome Reveals High Levels of Antibiotic Resistance Genes and Mobile Genetic Elements in Activated Sludge. *Plos One* 6, e26041.

Zhang, W., Wang, H., Zhang, R., Yu, X.Z., Qian, P.Y., Wong, M.H., 2010. Bacterial communities in PAH contaminated soils at an electronic-waste processing center in China. *Ecotoxicology* 19, 96-104.

Zhang, W.H., Wei, C.H., Chai, X.S., He, J.Y., Cai, Y., Ren, M., Yan, B., Peng, P.A., Fu, J.M., 2012. The behaviors and fate of polycyclic aromatic hydrocarbons (PAHs) in a coking wastewater treatment plant. *Chemosphere* 88, 174-182.

Figures and Captions

Fig. 1. Bacterial community structure at the phylum level in the soils and culturable bacteria isolated from the soils of P3. P1, P2 and P3 represent soil samples collected from different sites at the petrochemical plant, respectively. CB represents culturable bacteria isolated from the soils (Site P3) containing the highest concentrations of PAHs among all soil samples.

Fig. 2. Abundance and resistance types of ARGs (A) and relative percentages of ARG resistance types (B) in the PAHs-contaminated soils and representative sources of pollution.

The abundance of ARGs was normalized to the size of the sequencing data, and “ppm” was calculated as one ARGs-like read in one million metagenomic sequencing reads. P1, P2 and P3 represent soil samples collected from different sites at the petrochemical plant, respectively. CB represents culturable bacteria isolated from the soils (Site P3). SF, AS and AQS represent swine feces, active sludge, and aquaculture sediments, respectively.

693

694 Fig. 3. PCA analysis of ARG profiles in the PAHs-contaminated soils and
695 representative sources of pollution.

696 Relative abundance of resistance types in the samples were used for the PCA analysis.
697 P1, P2 and P3 represent soil samples collected from different sites at the
698 petrochemical plant, respectively. CB represents culturable bacteria isolated from the
699 soils (Site P3). SF, AS and AQS represent swine feces, active sludge, and aquaculture
700 sediments, respectively. A1 to A6 represent the soils contaminated with oil spills
701 downloaded from MG-RAST.

702

703

704

705

706 Fig. 4. Diversity of ARG subtypes connected with each of resistance types (A), Venn
707 diagram of ARG subtypes (B) in the PAHs-contaminated soils and mixing bacterial
708 culture from the soils.

709 P1, P2 and P3 represent soil samples collected from different sites at the
710 petrochemical plant, respectively. CB represents culturable bacteria isolated from the
711 soils (Site P3). Subtypes indicate the types of antibiotic resistance genes described in
712 the ARDB database.

713

714

715
716
717
718
719
720
721

Fig. 5. Relative percentages of ARGs potential related to plasmids over the total ARGs in the PAHs-contaminated soils and representative sources of pollution. Log-scale was used for both abundance and relative percentage of ARGs potential related to plasmids.

722

723

724 Fig. 6. Correlation between the total abundance of ARGs and that of Proteobacteria in
725 the soils

726

Bacterial community

Proteobacteria