



Microbiome and antibiotic resistome in household dust from Beijing, China

Long-Jun Ding^a, Xin-Yuan Zhou^{b,c}, Yong-Guan Zhu^{a,b,c,*}

^a State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China

^b University of Chinese Academy of Sciences, Beijing 100049, China

^c Key Laboratory of Urban Environment and Health, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen 361021, China

ARTICLE INFO

Handling Editor: Frederic Coulon

Keywords:

Microbial diversity
Geographical distribution
Bacteria
Fungi
Antibiotic resistant genes
Household dust

ABSTRACT

We spend ever-increasing time indoors along with urbanization; however, the geographical distribution patterns of microbiome and antibiotic resistome, and their driving forces in household environment remains poorly characterized. Here, we surveyed the bacterial and fungal communities, and the resistome in settled dust gathered from 82 homes located across Beijing, China, employing Illumina sequencing and high-throughput quantitative PCR techniques. There was no clear geographical distribution pattern in dust-related bacterial communities although a slight but significant ($P < 0.05$) distance-decay relationship occurred in its community similarity; by contrast, a relatively distinct geographical clustering and a stronger distance-decay relationship were observed in fungal communities at the local scale. The cross-domain (bacteria versus fungi) relationships in the microbiome of the dust samples were mostly observed as robust co-occurrence correlations. The bacterial communities were dominated by *Proteobacteria* and *Actinobacteria* phyla, with human skin, soil and plants being potential major sources. The fungal communities largely comprised potential allergens (a median 61% of the fungal sequences), with *Alternaria* genus within *Ascomycota* phylum being the most predominant taxa. The profile of dust-related bacterial communities was mainly affected by housing factors related to occupants and houseplants, while that of fungal communities was determined by georeferenced environmental factors, particularly vascular plant diversity. Additionally, a great diversity (1.96 on average for Shannon index) and normalized abundance (2.22 copies per bacterial cell on average) of antibiotic resistance genes were detected across the dust samples, with the dominance of genes resistant to vancomycin and Macrolide-Lincosamide-Streptogramin B. The resistome profile exhibited no distinct geographical pattern, and was primarily driven by certain bacterial phyla and occupancy-related factors. Overall, we underline the significance of anthropogenic impacts and house location in structuring bacterial and fungal communities inside homes, respectively, and suggest that household dust is an overlooked reservoir for antibiotic resistance.

1. Introduction

With urbanization, people currently spend approximately 90% of their daily lives indoors, and this proportion can be even higher for the elderly and children. Household environment, a typical representative of indoor environments, is perceived as a complex ecosystem that comprises a myriad of organisms including trillions of microbes (Amend et al., 2010). Some of the microbial populations inside homes can adversely influence human health, either as pathogens (e.g., bacterial taxa *Streptococcus pneumoniae*) or functioning as inducements of allergies and asthma symptoms (e.g., fungal taxa *Alternaria alternata*) (Barberán et al., 2015a; Kakuta et al., 2020). In contrast, other populations such as immunoregulation-inducing microbes which accompany mammalian

evolution (termed as 'old friends') could have positive health effects (Rook et al., 2013). A series of studies have suggested that limited early exposure to microbial diversity inside homes could partially account for the increased risks of allergies and other chronic inflammatory and even psychiatric disorders in many developed countries (Ege et al., 2011; Rook et al., 2013; Mubanga et al., 2017; Böbel et al., 2018). Therefore, exploring the ecology of microbial life inside homes is of vital importance for human health and wellbeing.

The microbiome living in a home is considered to be an intricate system depending on several aspects, including microbiota of the occupants, selective pressures exerted by lifestyle of the occupants and the building itself, and dispersal from the biome outdoors (Gilbert and Stephens, 2018). It is apparent that the household microbial community

* Corresponding author at: State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China.

E-mail address: ygzhu@rcees.ac.cn (Y.-G. Zhu).

<https://doi.org/10.1016/j.envint.2020.105702>

Received 20 February 2020; Received in revised form 26 March 2020; Accepted 28 March 2020

Available online 02 April 2020

0160-4120/ © 2020 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

structure is linked to the microbiota of human and non-human (e.g., dogs and cats) occupants mostly through skin-to-surface touch and/or direct emission of microbial particles, and the number, gender and behaviors of human occupants could affect the composition of microbial assemblages inside homes somewhat (Barberán et al., 2015a; Gupta, 2017; Gilbert and Stephens, 2018). Besides, the attributes of the house itself such as the design, ventilation type, and moisture have been reported to shape the microbiota found inside homes to some extent (Dannemiller et al., 2016). It has also been proposed that the geographical location of a house might play an important part in shaping the microbiome living in a given home when home is regarded as a microbial ecosystem, which are similar to the biogeography of plants or animals (Barberán et al., 2015b). On the one hand, many environmental factors such as urbanization level, soil pH, and climatic variables have been found to, more or less, affect the types of bacterial and fungal taxa detected in outdoor air, lead to an expected geographical structure of these airborne microbial communities found outdoors, and this might eventually influence the surrounding microbes live indoors via dispersal (Bowers et al., 2011; Barberán et al., 2015a, 2015b). On the other hand, several environmental factors like climatic variables could affect the indoor environmental conditions by air exchange, and then exert impacts on the microbiota found indoors directly (Martin et al., 2015). However, the knowledge about the biogeography of microbiome inside homes and their driving forces remains limited, especially in developing countries where the environmental (e.g., the level of urbanization) and housing variables (e.g., house design) are different from those in the developed countries that previous related studies mainly focused on.

It is well known that bacteria can carry antibiotic resistance genes (ARGs). The development of antibiotic resistance in pathogenic bacteria resulted from abuse and misuse of antibiotics has been regarded as an emerging threat to modern public health by World Health Organization (2014). This development primarily involves the recruitment of ARGs into pathogenic bacteria through mobile genetic elements (MGEs) like integrons and transposons (van Hoek et al., 2011). Both ARGs and MGEs are widely distributed in diverse environments including the built environment (Yooseph et al., 2013; Pal et al., 2016). For instance, several ARGs [e.g., *tet(K)* gene conferring resistance to tetracycline] have been reported to be significantly and positively correlated with specific antimicrobial chemicals (e.g., triclosan) in dust collected from a mixed-use athletic and educational facility (Hartmann et al., 2016). What is more, Gandara et al. (2006) have isolated antibiotic resistant pathogens *Staphylococcus aureus* from air inside and outside homes, and found these airborne *S. aureus* isolates were more concentrated inside than outside the homes. Although the antibiotic resistance in host-related bacterial taxa and the occurrence of priority antibiotic-resistant pathogenic bacteria in the built environment have attracted growing attention, little is yet known about the profile of ARGs (i.e., the antibiotic resistome) indoors, particularly in the household environment. In addition, how does the resistome found inside homes vary at a local scale and what factors are major contributors to the household resistome profile are also knowledge gaps.

In this study, given that the settled dust method has been widely applied in indoor air surveys (Rintala et al., 2012), we employed dust samples gathered from 82 homes randomly located in Beijing, China to characterize profiles of bacterial and fungal communities, together with antibiotic resistome inside homes, explore their geographical patterns at the local scale, and disentangle factors determining the geographical structure of the microbes and ARGs living in the homes, using Illumina MiSeq sequencing in combination with high-throughput quantitative PCR techniques.

2. Materials and methods

2.1. Sample collection and characterization

Indoor dust samples were collected by volunteers from 82 households in Beijing city, northern part of China in spring 2017 (from March to April). A map of sampling house locations is shown in Fig. S1. Each participant was made aware of nature of the study, and supplied with a sampling protocol and a sampling kit containing sterile cotton-tipped swabs (Shandong Longhai Medical Devices Co., Ltd., China). The interior windowsill in a living room was chosen as a sampling location as it exists in almost all homes. Participants were asked to close windows of the living room for about 3 days before dust sampling, hence the corresponding interior windowsill is considered as a collector of deposited dust inside home with possible contact with home occupants. The detailed information about sampling procedure is given in Supplementary Methods. The participants wore sterile latex gloves and surgical face masks throughout the sampling period to preclude pollution while talking or handling swabs. After sampling, the swabs were rapidly transported to the lab within 12 h, and stored at -80°C for DNA extraction. Meanwhile, un-used swabs in the sampling kit are regarded as control samples, and they were subjected to the same molecular analyses (e.g., DNA extraction and amplification) as used swabs. Additionally, a questionnaire on general conditions of the house and occupancy (both human and pet) was completed by each participant and also returned to the lab.

Geographical coordinates (latitude and longitude) were derived from house location information (i.e., home address). Coordinates were then used to obtain the information of georeferenced environmental variables for each household. In addition, the information of the home design and occupancy was obtained through the returned questionnaire. The detailed information of environmental and housing variables are shown in Supplementary Methods and Table S1.

2.2. Extraction of DNA for dust samples

Genomic DNA for dust samples was extracted from swabs using the FastDNA® Spin Kit for soil (Bio 101 Inc., USA) based on manufacturer's protocols with the modifications that the cotton tips of frozen swabs were directly broken off into bead tubes and the lysis procedure was implemented twice. Additionally, un-used sterile swabs were also subjected to DNA extraction as a control. The concentration and purity of the extracted DNA were measured by a UV-Vis Spectrophotometer (NanoDrop Technologies Inc., Wilmington, DE, USA).

2.3. High-throughput quantitative PCR (HT-qPCR) of antibiotic resistance genes

To survey abundances of ARGs in dust samples, HT-qPCR reactions were carried out with the extracted DNA on the Takara SmartChip Real-time PCR system which encompasses 5184 reaction wells with 100-nL volume for each well (Su et al., 2015). 296 verified primer pairs were employed during HT-qPCR reactions as reported by our recent study (Xiang et al., 2020), including 283 pairs targeting ARGs conferring resistance to nearly all major types of antibiotics, 12 pairs targeting MGEs (8 for transposase genes and 4 for integron-integrase genes), and a pair for bacterial 16S rRNA gene (Table S2). The qPCR thermal conditions consisted of 10 min at 95°C , and 40 cycles of denaturation at 95°C for 30 s and annealing at 60°C for 30 s for amplification, followed by a melting curve analyses automatically generated by Takara system (Chen et al., 2016). All samples were performed in technical triplicates. HT-qPCR results were preprocessed by SmartChip qPCR Software excluding the wells with multiple melting peaks or amplification efficiency beyond the range 1.7–2.3. A threshold cycle (C_T) of 31 was used as the detection limit, and only samples with detected amplification products in all three technical replicates were considered as positive

and used in subsequent data analyses. Relative gene copy number was calculated according to Eq.1 (Looft et al., 2012).

$$\text{Relative gene copy number} = 10^{(31 - C_T)(10/3)} \quad (1)$$

To minimize biases resulted from differences in background bacterial abundances and yields of the extracted DNA, the abundance of each ARG/MGE was transformed to 'relative abundance' (copies per copy of 16S rRNA gene) through dividing copy number of each ARG/MGE by that of 16S rRNA gene (Zhu et al., 2017). Subsequently, 'normalized abundance' of each ARG/MGE (copies per bacterial cell) was calculated through multiplying its relative abundance by 4.1, as the mean copy number of 16S rRNA genes per bacterial cell is estimated to be 4.1 at present (Su et al., 2017).

2.4. Illumina sequencing and raw data processing

To explore the microbial diversity of dust samples, the extracted DNA of dust samples was amplified using a primer pair 515F/806R that targets the V4 hypervariable region of the 16S rRNA gene for bacteria, and a primer pair ITS1F/ITS2 that targets the internal transcribed spacer (ITS) region 1 of the rRNA operon for fungi, respectively (Barberán et al., 2015b). The details about PCR amplification are given in [Supplementary Methods](#). Moreover, negative controls including no-template (three replications) and template from un-used swabs (three replications) were also subjected to amplification. Consequently, neither the bacterial 16S rRNA sequences nor fungal ITS sequences of these negative controls could be amplified, indicating that both PCR reagents and sample DNA were not polluted. The PCR products successfully amplified were then purified and pooled in equimolar concentration for paired-end sequencing on an Illumina MiSeq PE300 platform by Majorbio Bio-Pharm Technology Co., Ltd. (Shanghai, China).

Raw sequencing data were analyzed by the Quantitative Insights into Microbial Ecology (QIIME) toolkit 2.0 (Caporaso et al., 2010). In brief, pairs of raw reads were quality-filtered by Trimmomatic and assembled using the PANDASEQ (Bartram et al., 2011). After trimming off the barcodes and primers, chimeric sequences were detected and removed via the Chimera Slayer. The resulted sequences were clustered into operational taxonomic units (OTUs) at 97% similarity cutoff, and singleton OTUs were then filtered (Ding et al., 2017). The taxonomic information of representative sequences from the OTUs was determined by the Ribosome Database Project (RDP) classifier trained on Silva 128 16S rRNA database (<http://www.arb-silva.de>) for bacteria (Quast et al., 2013), and the UNITE 6.0 ITS database (<http://unite.ut.ee/index.php>) for fungi (Abarenkov et al., 2010), both at a confidence level of 70%. The representative sequences identified as mitochondrion or chloroplasts were removed from OTU table before further analyses.

To minimize potential biases caused by different sequencing depth, samples with fewer than 10,000 sequences were discarded. Consequently, the number of samples for downstream analyses was 82 for bacteria and 75 for fungi. The bacterial and fungal sequencing depth were considered to be sufficient to delineate the profiles of dust-associated bacterial and fungal communities as their rarefaction curves of Shannon index had both reached a plateau (Fig. S2).

The raw sequencing data of the bacterial and fungal communities have been uploaded to the National Center for Biotechnology Information (NCBI) Sequence Read Archive (<http://trace.ncbi.nlm.nih.gov/Traces/sra/>) with accession number PRJNA573814.

2.5. Analyses of source tracking and bacterial source indicators

In order to discern the relative contribution of different putative sources to the bacterial communities living in the investigated homes, two complementary methods were used, including Bayesian source tracking and indicator taxa analysis (Flores et al., 2013). Six putative source environments were chosen for both methods, including soil,

plants, human skin, human oral cavity, human/livestock stool and insects, as they are regarded as potential important sources of indoor microbes (Barberán et al., 2015b). The Bayesian source tracking approach was carried out based on the SourceTracker algorithm in QIIME (Knights et al., 2011). The indicator taxa analysis was carried out using packages of vegan 2.3–5, indicspecies 1.7.6, and labdsv 1.8–0 in R 3.1.3. The detailed information of these two methods are shown in [Supplementary Methods](#). The list of bacterial indicator taxa of different putative source environments identified herein is presented in [Table S3](#).

2.6. Identification of potential bacterial pathogens and fungal allergens

Potential bacterial pathogens in the dust samples were determined based on a database of bacterial pathogens reported by our earlier study (Chen et al., 2016), in which the 16S rRNA gene sequences of bacterial pathogens are all available from the NCBI (<http://www.ncbi.nlm.nih.gov/>). The 16S rRNA gene sequences of each sample were blasted against those of bacterial pathogens in the database, with an e -value $< 1 \times 10^{-10}$. The results of blast hit were then filtered to identify the potential bacterial pathogens with a sequence similarity threshold $> 99\%$. Additionally, potential fungal allergens were determined according to previously published literature and the directory of pathogenic microbes infecting human beings issued by the Ministry of Health of the People's Republic of China (Barberán et al., 2015a; Du et al., 2018; Hernandez and Martinez, 2018).

2.7. Statistical and network analyses

For all statistical analyses, significance level was set as $P < 0.05$ or with listed P -values. SPSS V20.0 (IBM, USA) software was used to implement standard statistical tests such as nonparametric Mann-Whitney test for significance and correlation analysis.

The collinearity of all the examined candidate explanatory factors was reduced by discerning strongly correlated factors through principle component analysis (PCA) and computing variance inflation factors (VIFs) (Barberán et al., 2015b). After collinearity analysis, 20 independent explanatory variables (including environmental and housing variables) were left for subsequent analyses (listed in [Table S1](#), except information of main vascular plant species outdoors and houseplant species). The relationships between dust-associated microbial community structure and the remaining explanatory factors were investigated using Mantel test. The relative contribution of each explanatory factor to the variations of bacterial or fungal community structure in dust samples was then calculated by DistLM forward3 ($P < 0.05$, permutation = 999) (Shi et al., 2018). In addition, the dissimilarities of microbial (i.e., bacterial and fungal) communities and ARGs among all dust samples were analyzed by a non-metric multidimensional scaling (NMDS) ordination based on a Bray-Curtis distance matrix. PCA, Mantel test and NMDS mentioned above, together with diversity index, canonical correspondence analysis (CCA) and variation partitioning analysis (VPA) were performed with packages of vegan 2.3–5 and ecodist 2.0.1 in R. Furthermore, distance-decay analyses of microbial communities and ARG profiles were carried out with packages of labdsv 1.8–0, vegan 2.3–5, and geosphere 1.5–10 in R. ArcGIS V10.0 software was employed to map the NMDS scores and the proportional abundances of microbial taxa at the spatial scale examined (Su et al., 2017). Network analysis was conducted using R package pheamap 1.0.12 and Gephi 0.9.2 as previously reported (Chen et al., 2016; Su et al., 2017). Details of statistical and network analyses are presented in [Supplementary Methods](#).

3. Results

3.1. Characterization of bacterial and fungal communities in household dust

The overview of Illumina sequencing results of 82 samples for

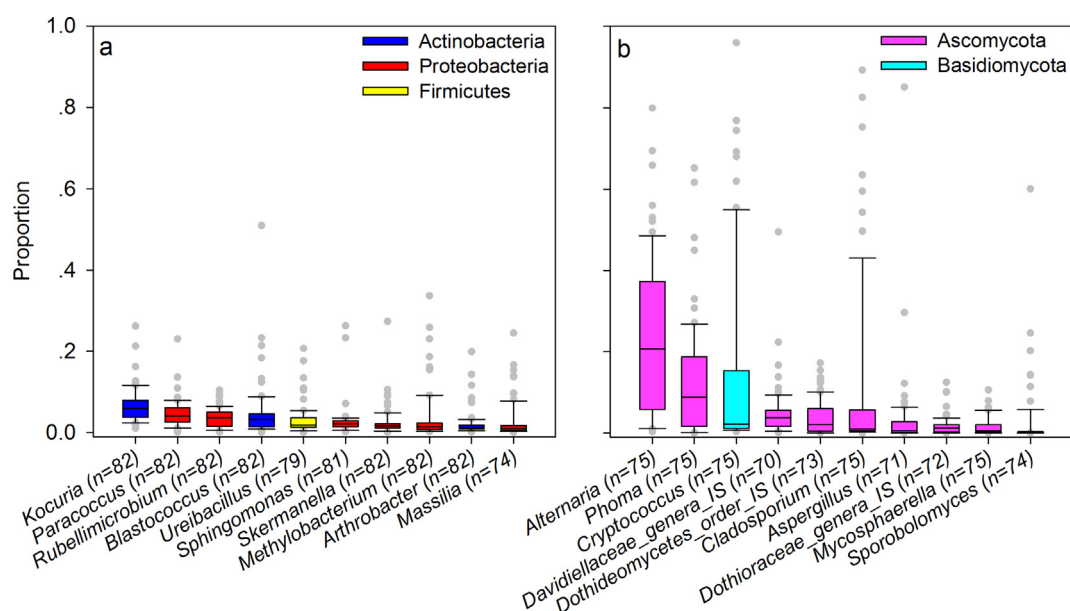


Fig. 1. Proportions of the top 10 abundant bacterial (a) and fungal genera (b) across dust samples. Genera were colored by their respective phylum. Number in parenthesis represents the frequency of occurrence of corresponding microbial taxa across the samples. *Davidiellaceae_genera_IS*, *Davidiellaceae_genera_incertain_sedis*, denotes an unnamed genus affiliating with *Davidiellaceae* family, the same below.

bacteria and 75 samples for fungi was shown in [Supplementary Results and Table S4](#) in the SI. Obviously, the bacterial communities across all samples had significantly higher overall diversity (represented by Shannon diversity; on average 4.87) and species richness (Chao1 estimator; on average 1333) than the fungal communities (Mann-Whitney test $P < 0.01$ for both indices) ([Fig. S3](#)).

For bacterial community composition, *Proteobacteria*, especially *Alphaproteobacteria* and *Actinobacteria* were predominant phyla across the dust samples, occupying 48.3–96.5% of the total bacterial abundance for each sample ([Fig. S4](#)). Accordingly, ten of the most abundant genera were almost belonged to these two phyla, with *Kocuria*, *Paracoccus* and *Rubellimicrobium* being the most abundant genera in terms of average relative abundance ([Fig. 1a](#)). These three predominant bacterial genera were also shared by all samples ([Fig. S5a](#)). The results of source tracking combined with indicator taxa analyses showed that the bacterial communities across the samples were mainly originated from bacterial populations probably observed in human skin, soil and plants

([Fig. 2](#) and [Table S3](#)), suggesting that human skin, soil and plants are potential primary sources of bacteria inside homes.

With respect to fungal community composition, most samples were dominated by *Pleiosporales* and *Capnoidiales* orders which were affiliated with *Dothideomycetes* class of *Ascomycota* phylum, accounting for an average of 78.7% of the total fungal abundance across the samples ([Fig. S6](#)). Furthermore, at the genus level, *Alternaria*, *Phoma* (both belonging to *Ascomycota* phylum) and *Cryptococcus* (belonging to *Basidiomycota* phylum) were the top three abundant populations in the light of average proportion despite their wide proportion ranges ([Fig. 1b](#)). As was the case with bacteria, these three dominant fungal genera were consistently detected across all samples ([Fig. S5b](#)).

3.2. Factors shaping bacterial and fungal communities in household dust

The composition of bacterial or fungal community was distinct among different dust samples, albeit the presence of a shared prevalence of certain bacterial or fungal taxa across the samples as described above ([Fig. 3a](#) and [b](#)). At the local scale, there was no clear geographical pattern in the structure of dust-related bacterial communities, while relatively obvious geographical gradient was detected in the fungal communities which was evidenced by that the dust samples collected from homes in the east of Beijing generally harbored different fungal communities from those found in the west of Beijing ([Fig. 3a](#) and [b](#)). Furthermore, significant ($P < 0.05$) distance-decay effects occurred for both groups of microbes, as the similarity of bacterial or fungal communities between any two homes decreased with increasing geographic distance (up to about 60 km), and this effect on the fungal community ($r = -0.063$, $P = 0.033$) was stronger than on the bacterial community ($r = -0.055$, $P = 0.041$) ([Fig. 3c](#) and [d](#)).

Based on Mantel test, the dust-related bacterial community structure exhibited a significant ($P < 0.05$) correlation mainly with housing variables, including numbers of habitants, dogs and cats, ratio of female to male number, and number of houseplants in order of correlation strength (i.e., r value) ([Table S5](#)). Moreover, all the surveyed housing variables explained greater variations of bacterial community structure (26.72%) than the environmental variables (13.83%), and the number of habitants contributed the most (6.98%) to the structural variations of bacterial communities across the samples ([Table 1](#)).

Unlike bacteria, fungal community structure was significantly

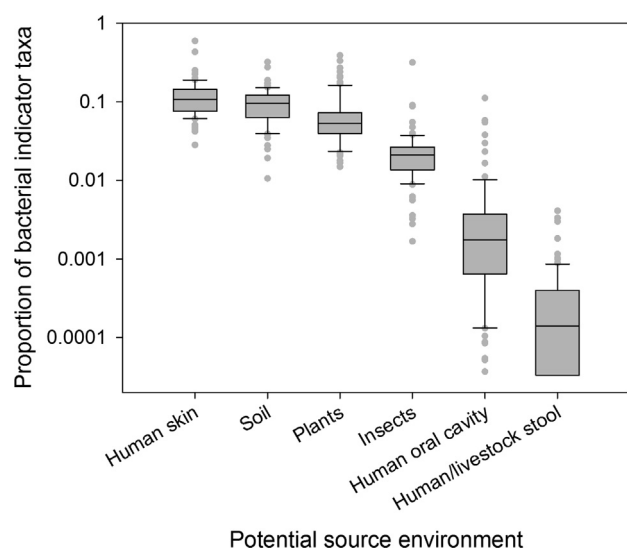


Fig. 2. Proportions of bacterial indicator taxa of different potential source environments across dust samples.

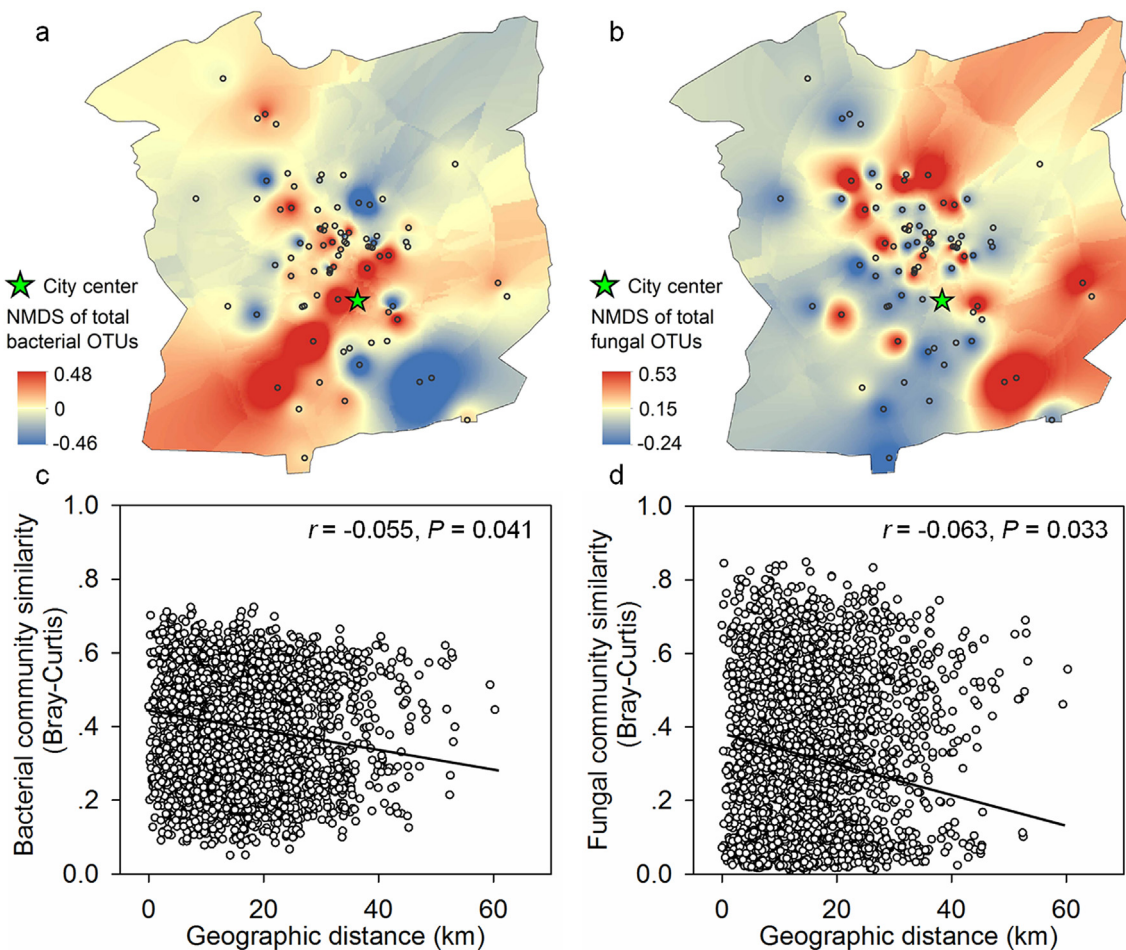


Fig. 3. Overall profiles of microbial communities in dust samples. **a** and **b**. ArcGIS map showing the variations of bacterial communities (**a**) and fungal communities (**b**) based on scores of non-metric multidimensional scaling (NMDS) ordination axis. Different colors indicate that bacterial or fungal communities are distinct in their community structure. Points denote sampling home locations in Beijing city. **c** and **d**. Relationship between geographical distance and similarity of bacterial (**c**) or fungal community (**d**) based on Bray-Curtis distance.

($P < 0.05$) correlated with environmental variables, including vascular plant diversity, population density, fine particulate matter ($PM_{2.5}$) concentration and woody vegetation cover percentage in order of correlation strength (Table S5). In addition, the environmental variables were more important contributors (explained 24.17%) to structural variations of fungal communities found in the investigated homes than the housing variables (14.23%), and the vascular plant diversity had the strongest explanatory power (4.95%) of the structural variations (Table 1).

3.3. Cross-domain relationship between bacterial and fungal communities in household dust

The network analysis was carried out to unveil potential cross-domain relationships in the household dust (Fig. 4). In brief, over 4800 significant and robust correlations occurred between OTUs assigned to 423 bacterial genera and those assigned to 261 fungal genera, with the majority (4840/4851, 99.8%) being involved in co-occurrence correlations (blue edge, Fig. 4 and Table S6). Among these significantly correlated bacterial and fungal genera, the OTUs of *Deinococcus* (within bacteria) and *Cryptococcus* genera (within fungi) were involved in the largest number of strong cross-domain correlations, respectively. Furthermore, OTUs from the most prominent fungal genus *Alternaria* engaged in different patterns of possible relationships with their bacterial neighbors, as they not only presented strong co-occurrence relationships with *Acinetobacter*, *Deinococcus* and *Flavobacterium* genera but also

strong co-exclusion correlations with *Bifidobacterium* and *Lactobacillus* genera (Figs. 1b and 4; Table S6). By contrast, OTUs of *Kocuria* genus, the most prominent bacterial population, showed a sole strong co-occurrence correlation with *Cryptococcus* genera within fungi (Figs. 1a and 4; Table S6).

3.4. Potential bacterial pathogens and fungal allergens in household dust

A total of 105 potential bacterial pathogenic species were identified from all dust samples and belonged to 29 genera (Table S7), with members of *Staphylococcus* (e.g., *Staphylococcus lugdunensis*), *Bacillus* (e.g., *Bacillus megaterium*), and *Pseudomonas* genera (e.g., *Pseudomonas putida*) being the top three predominant potential pathogenic bacteria, accounting for 2.06%, 0.62% and 0.46% on average, respectively (Fig. S7a). The total relative abundances of the potential pathogenic bacteria detected varied from 0.04% to 35.3% (median = 2.3%) across the samples, and were below 2.5% in more than half of the samples (Fig. S7c). In addition, 13 known potential fungal allergens at the genus level were detected in all samples, with *Alternaria* (on average 23.1%), *Phoma* (13.3%) and *Cladosporium* (9.73%) being the three most abundant allergenic fungi (Fig. S7b). Compared with the potential bacterial pathogens, the potential fungal allergens occupied a quite larger proportion of the total fungal abundance (0.78–95.6%, median = 61%), which was above 60% in approximately half of the samples (Fig. S7d).

Table 1

Structural variation of bacterial or fungal community in home dust explained by environmental and housing variables.

Community	Variable class	Variable ¹	SS ²	P	Percentage (%)	Cumulative (%)
Bacterial community	Environmental variables	Population density	2911	0.049*	2.30	2.30
		PM _{2.5} concentration	2162	0.268	1.88	4.18
		Vascular plant diversity	2046	0.325	1.61	5.79
		Woody vegetation cover	2004	0.365	1.58	7.37
		PM ₁₀ concentration	1900	0.418	1.52	8.89
		MAP	1885	0.432	1.41	10.30
		MAT	1750	0.593	1.13	11.43
		Hog density	1558	0.725	0.92	12.35
		Distance to city center	1434	0.839	0.75	13.10
		Elevation	1399	0.845	0.73	13.83
	Housing variables	Number of habitants	10,336	0.001*	6.98	20.81
		Number of dogs	4303	0.010*	3.42	24.23
		Number of cats	3809	0.023*	3.13	27.36
		Number of houseplants	3411	0.036*	2.89	30.25
		Female:Male ratio	3083	0.042*	2.70	32.95
		Floor height	2411	0.128	2.02	34.97
		Construction age	2323	0.208	1.97	36.94
		Living room space	1871	0.454	1.39	38.33
		House space	1808	0.525	1.33	39.66
		Number of carpets	1524	0.738	0.89	40.55
Fungal community	Environmental variables	Vascular plant diversity	8453	0.007*	4.95	4.95
		Population density	6375	0.012*	3.62	8.57
		PM _{2.5} concentration	6282	0.017*	3.56	12.13
		Woody vegetation cover	6261	0.021*	3.35	15.48
		PM ₁₀ concentration	3969	0.054	2.36	17.84
		Hog density	3541	0.072	2.12	19.96
		Elevation	2593	0.616	1.21	21.17
		MAP	2385	0.644	1.07	22.24
		MAT	2350	0.549	1.05	23.29
		Distance to city center	2080	0.779	0.88	24.17
	Housing variables	Number of habitants	4061	0.052	2.18	26.35
		Number of dogs	3881	0.059	2.03	28.38
		Number of houseplants	3586	0.052	1.84	30.22
		Number of cats	3375	0.232	1.71	31.93
		Female:Male ratio	3354	0.183	1.69	33.62
		Floor height	2649	0.504	1.24	34.86
		Living room space	2508	0.134	1.15	36.01
		Number of carpets	2365	0.598	1.06	37.07
		House space	1908	0.846	0.77	37.84
		Construction age	1584	0.955	0.56	38.40

¹ PM_{2.5} concentration: mean annual concentrations of fine particulate matter; Woody vegetation cover: percentage of woody vegetation cover; PM₁₀ concentration: mean annual concentrations of inhalable particles; MAP: mean annual precipitation; MAT: mean annual temperature.

² SS: sum of square; percentage: percentage of the variation explained by each variable; cumulative: cumulative percentage of the explained variation.

* $P < 0.05$.

3.5. Characterization of antibiotic resistance genes and mobile genetic elements in household dust

A total of 135 ARGs were observed in all 82 dust samples. The detected number and Shannon index of ARGs both dramatically varied from one sample to another, with a mean of 14 and 1.85, respectively (Fig. S8). Additionally, 11 MGEs (including 8 transposase genes and 3 integron-integrase genes) were detected only in 61 samples, with an average number of 2 for each sample, and their Shannon index ranged from 3.38×10^{-3} to 1.46, and averaged 0.75 (Fig. S8).

The total normalized abundance of ARGs ranged between 0.02 and 26.6 copies per bacterial cell in different samples, with an average of 2.10 copies per cell, and that of MGEs also spanned over three orders of magnitude (from 1.34×10^{-3} to 1.39 copies per cell) (Fig. 5a). The predominant resistance gene types across the samples were almost genes resistant to vancomycin and macrolide-lincosamide-streptogramin B (MLS_B), accounting for an average of 56.3% of the total normalized abundances of ARGs and MGEs across the samples (Fig. 5b and c). For instance, *vanXD* (vancomycin resistant), *vate-01* (MLS_B resistant) and *vanSB* genes (vancomycin resistant) were consistently observed at high normalized abundances, averaging 0.83, 0.47 and 0.15 copies per cell, respectively (Fig. S9a). Moreover, a shared prevalence of the ARGs and MGEs was found across the samples,

contributing to approximately 51% of the total normalized abundances of ARGs and MGEs on average for each sample (Fig. S9b).

As was the case with bacteria, no clear geographical clustering was observed in the structure of ARGs across the samples via NMDS analysis (Fig. S10a). Nevertheless, a mild but significant ($r = -0.052$, $P = 0.047$) distance-decay impact was still detected between the similarity of ARGs and geographic distance of any two investigated homes (Fig. S10b).

3.6. Factors determining structures of antibiotic resistance genes in household dust

The normalized abundance of total ARGs exhibited a positive and significant correlation (Pearson correlation coefficient = 0.608, $P < 0.01$) with that of MGEs (Table S8). CCA was further applied to reveal the relationships between the dust-associated ARG structures and four groups of factors including bacterial community (at the phylum level), environmental and housing variables, and MGEs. Nine significant ($P < 0.05$) variables, including the relative abundances of *Actinobacteria*, *Proteobacteria* and *Firmicutes* phyla, the numbers of habitants, cats and dogs, floor height, population density, and MGEs were shown in CCA plot (Fig. 6a). The first axis (CCA1), which explained 28.5% of structural variations of ARGs presented positive correlations

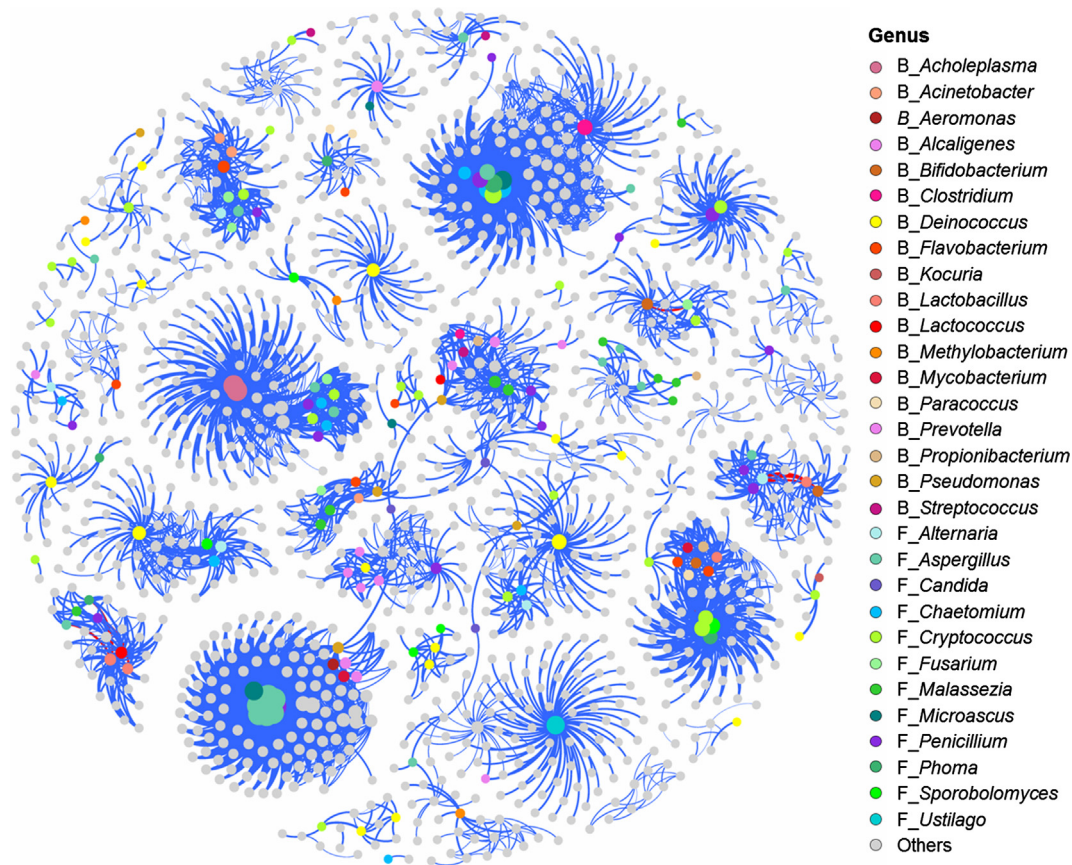


Fig. 4. Co-occurrence and co-exclusion patterns between bacterial and fungal OTUs visualized by network analysis. Only significant correlations ($P < 0.01$) with an absolute correlation coefficient magnitude (i.e., Spearman's ρ) ≥ 0.7 are shown to focus on merely strong correlations. The detailed information of these strong and significant correlations is given in Table S6. Nodes (circles) represent OTUs involved in either significant co-occurrence (correlation coefficient ≥ 0.7 ; blue edges, i.e., blue lines connecting the nodes) or co-exclusion (correlation coefficient ≤ -0.7 ; red edges) relationships. Each node is colored according to the genus of the OTU, and its size is proportioned to the number of connections. The width of the edges is weighted based on the correlation coefficient magnitude. The top 30 bacterial and fungal genera most involved in strong and significant relationships are shown, and the remaining genera are clustered as “Others”. *B_Acholeplasma*, denotes *Acholeplasma* genus was affiliated with *Bacteria* kingdom; *F_Alternaria*, denotes *Alternaria* genus was affiliated with *Fungi* kingdom, the same below. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

with the relative abundances of *Actinobacteria* and *Firmicutes*, and the numbers of cats and dogs, while negative correlations with the relative abundance of *Proteobacteria* and floor height. The second axis (CCA2), which explained 13.7% of the variations was positively correlated with the normalized abundance of MGEs, and negatively correlated with the number of habitants (Fig. 6a). The relative abundances of *Actinobacteria*, *Firmicutes* and *Proteobacteria* appeared to contribute more to structuring ARGs in the household dust as they were represented by longer arrows compared with the other significant variables.

CCA-based VPA results indicated that roughly 58% of the variations could be explained by these significant variables, with the bacterial community (including the relative abundances of *Actinobacteria*, *Firmicutes* and *Proteobacteria*) explaining much more (17.3%) than the housing factors (including the numbers of habitants, cats and dogs, and floor height; 10.5%), environmental factors (population density; 3.5%) and MGEs (1.7%; Fig. 6b). Moreover, the joint effect of bacterial community and housing factors was the greatest (14.4%) within all interactions between any two components. These results suggested the importance of bacterial community and housing factors in structuring indoor antibiotic resistance.

3.7. Co-occurrence patterns between ARGs and bacterial taxa in household dust

Mantel test uncovered a significant correlation ($r_M = 0.29$, $P = 0.032$) between ARGs and bacterial OTUs on the basis of Bray-

Curtis distance. Co-occurrence patterns between ARGs and bacterial genera with significant ($P < 0.05$) and robust correlations (Spearman's correlation coefficient ≥ 0.7) were further visualized via network analysis (Fig. S11). On the one hand, the nodes of ARGs conferring resistance to beta-lactamase (e.g., *ampc*-01, 04, 05 and 06), MLSB (e.g., *lnuA*-01) and multidrug (e.g., *mexE*) had markedly higher levels of bacterial genera-connectivity than those of other ARGs (Fig. S11 and Table S9). On the other hand, bacterial genera affiliated with *Proteobacteria*, *Firmicutes*, *Actinobacteria* and *Bacteroidetes* phyla were involved in much larger numbers of strong correlations with ARGs compared with other phyla (Table S9). Furthermore, different bacterial genera almost showed distinct profiles of ARGs even though they are from the same phylum. However, these co-occurrence patterns detected do not provide direct evidence that bacterial taxa actually carry the co-existed ARGs.

4. Discussion

4.1. Microbial diversity in household dust

We found household dust collected from Beijing harbored a great richness and diversity of bacterial and fungal communities, which were comparable to those detected in PM_{2.5} of heavy haze from Beijing as well (Du et al., 2018). Not surprisingly, bacterial taxa related to skin of humans and nonhuman mammals were prominent component of the dust-related bacterial communities. For example, the actinobacterial

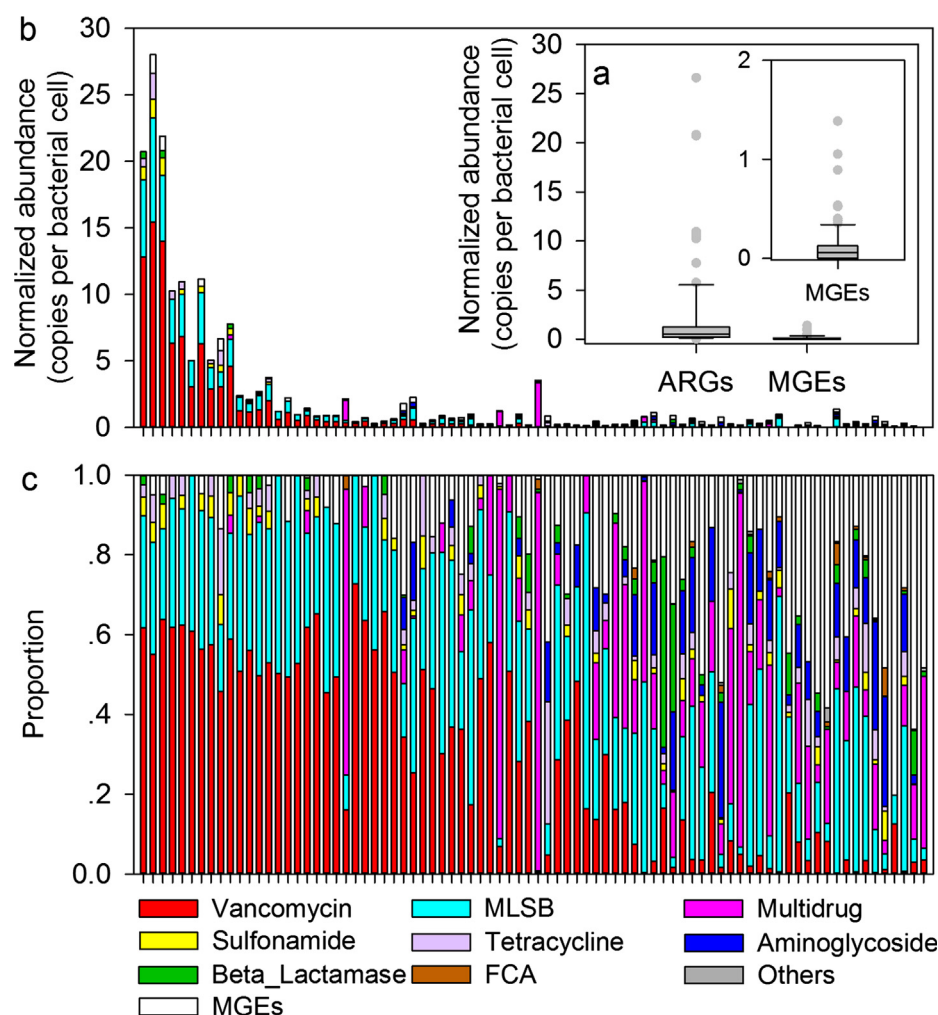


Fig. 5. ARG profiles in dust samples. (a) Total normalized abundance (copies per bacterial cell) of all ARGs and MGEs detected across dust samples. (b) Normalized abundance of ARG types and MGEs in each sample. Samples were ordered by their scores along the non-metric multidimensional scaling (NMDS) ordination axis of ARGs, the same below. (c) Proportion of ARG types and MGEs in each sample. Data are shown as the proportion of the normalized abundance of ARGs conferring resistance to a specific class of antibiotics to the total normalized abundance of all ARGs detected. MLSB: Macrolide-Lincosamide-Streptogramin B resistance genes; FCA: (flor)/(chlor)/(am)phenicol resistance genes; Others: ARGs conferring resistance to other antibiotic classes like bacitracin, fosfomycin, nitroimidazole, pyrazinamide and streptothricin.

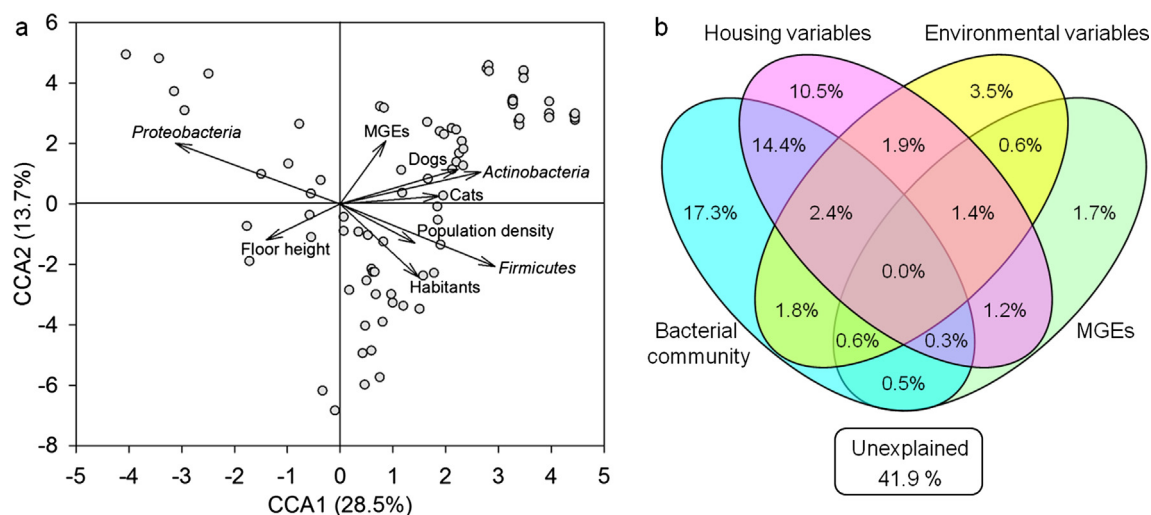


Fig. 6. Factors influencing ARG profiles in dust samples. (a) Canonical correlation analysis (CCA) revealing relationships between ARGs and bacterial community (at the phylum level), housing and environmental variables, and MGEs. The bacterial community contained relative abundances of *Actinobacteria*, *Proteobacteria* and *Firmicutes* phyla; the housing variables contained floor height, number of habitants, cats and dogs; and the environmental variable contained population density. These factors were chosen according to significance test and variance inflation factors (VIFs) calculated during CCA. The percentage of variation explained by each axis was given in parenthesis. (b) Variation partitioning analysis (VPA) distinguishing relative contributions of the selected bacterial phyla, housing and environmental variables, and MGEs to alterations of ARGs across the samples. Individual effect of each component and joint effects of any two or three components were shown. Additionally, the interaction among all four components was zero through calculation.

genus *Kocuria*, the most abundant bacterial taxa in the examined dust, is usually regarded as typical microbiota of skin in humans and non-human mammals (Kolikonda et al., 2017). This population was recognized to be an indicator taxa of human skin herein (Table S3), though it has also been detected in other environments like soil, air and clinical specimen (Du et al., 2018; Lecomte et al., 2011). Besides, several other predominant bacterial taxa in the dust samples (Fig. 1a) have been reported to be associated with diverse environments such as skin of humans and nonhuman mammals (for *Paracoccus* and *Methylobacterium* genera), soil (for *Arthrobacter* genus), plants (for *Sphingomonas* and *Massilia* genera) and PM_{2.5} (for *Rubellimicrobium* genus) (Redford et al., 2010; Du et al., 2018; Leung et al., 2018; Ross et al., 2018). These findings were supported by the result of source tracking analysis which indicated that human skin, soil and plants were potential important sources of bacterial communities living in the surveyed homes.

The dust-related fungal communities were mostly dominated by molds (e.g., *Dothideomycetes* class within *Ascomycota* phyla) commonly found indoors, particularly in rooms with the presence of water source (e.g., bathroom) (Adams et al., 2013). It is worth noting that several genera affiliated with the saprotrophic *Dothideomycetes* class, like *Alternaria* and *Phoma* genera are considered to be potential allergens (Barberán et al., 2015a). Notably, *Alternaria* genus, the most dominant dust-related fungal taxa in this study, encompasses members (e.g., *A. alternata*) that are renowned triggers of allergic disease like asthma (Fernandez et al., 2011), and is often thought to be an outdoor allergen but also observed in indoor environments (Barberán et al., 2015a; Du et al., 2018). Intriguingly, these putative allergenic fungi were not only highly abundant but also prevalent across all the dust samples, suggesting the important role of fungal allergens in the household environment. Given that different species within these potential allergenic genera harbor distinct allergenicity, relationships between the proportions of these genera in the household dust and incidence of allergic disorders are still elusive and warrant further studies.

The cross-domain (bacteria vs. fungi) relationships in the dust-associated microbial communities were mostly detected as co-occurrence correlations, suggesting that bacterial populations mainly coexist with fungal populations in the household dust. Although correlation does not mean interaction, coexistence is a precondition for the occurrence of microbial interaction which has been found in diverse ecosystems (Pele et al., 2010). For example, Leung et al. (2016) have observed co-occurrence correlations between certain bacterial and fungal genera (e.g., *Pseudomonas* vs. *Candida*) in human skin microbiota, and proposed that those cross-domain relationships would aid interactive processes involving signal mechanisms, quorum and changing surrounding environments, and alterations in those correlation patterns could be linked with dermatopathology (Oh et al., 2013). Thus, the interactions arose from coexistence are likely to harness the growth and lifestyle features of bacterial and fungal assemblages inside homes. Additionally, it is noteworthy that the OTUs of a rare bacterial genus *Deinococcus* (average relative abundance < 0.1%, data not shown) exhibited a highly dense connectivity with their fungal neighbors (Fig. 4 and Table S6). This indicated that the relative abundances of microbial taxa were not necessarily linked with their connectivity levels in correlation patterns, and suggested the potential importance of exploring rare taxa within the microbiome (Shade et al., 2014).

4.2. Driving forces of geographical distribution patterns of bacterial and fungal communities in household dust

In the present study, there was no distinct geographical structure in the variations of bacterial communities found across all the dust samples, implying that geographical location was not a primary determinant of the bacterial community structure found in a given home. This implication was validated by the results of Mantel test and DistLM analysis which revealed that house-associated factors, especially human

occupancy, played a vital role in structuring bacterial communities living in the investigated homes (Tables 1 and S5). This was not surprising because human occupancy has been proved to greatly raise the concentration of indoor airborne microbiota via direct shedding and resuspension from surfaces (e.g., floor) (Hospodsky et al., 2012). Also, the result of source tracking analysis supported this finding. In addition to human occupants, indoor plants would also make a considerable contribution to the bacterial assemblages found inside homes given the occurrence of significant correlation between the dust-related bacterial community structure and the number of houseplants (Tables 1 and S5).

Taken together, we demonstrate the importance of occupancy and home design (e.g., introduction of houseplants) in determining the bacterial community structure inside homes, a finding that probably accounts for the differences in the relative abundances of the potential bacterial pathogens detected among all dust samples (Fig. S7). Although *Staphylococcus* genus are often found as part of the typical skin microbiota, *S. lugdunensis*, the predominant bacterial pathogen identified from the examined dust, is thought to be medically significant as it can cause many types of infection such as skin infections and endocarditis (Barberán et al., 2015a; Pereira et al., 2010). These suggested a possible relationship between occupancy and pathogenic bacteria inside homes. It has also been documented that building design like ventilation types and room temperature could affect the diversity and composition of indoor human pathogens (Kembel et al., 2012; Dannemiller et al., 2016). Additionally, introduction of appropriate houseplants has been recommended to be an energy-efficient approach to improve air quality in different indoor environments, by virtue of plants themselves and/or plant-associated microbiome (Brilli et al., 2018). Therefore houseplants may exert positive effects on human health although whether they could influence the indoor bacterial pathogens or not are yet unclear. Thus, homes are likely to be designed and manipulated (e.g., changing the types of occupants and introducing proper houseplants) to be inhospitable for bacterial pathogens and then suitable for people to live in. It should be noted that although the determination of potential bacterial pathogens in dust samples based on blasting against a database of 16S rRNA gene sequences of bacterial pathogens is an effective approach for pathogen identification, it still has some limitations such as inherent biases during PCR process (Haft and Tovchigrechko, 2012), and insufficient taxonomic resolution at species level due to relatively short sequence length (291 bp in the present study) (Segata et al., 2012). Therefore, future works on the basis of metagenomic technique should be conducted to provide a more accurate profile of bacterial pathogens in the household dust.

Unlike bacterial communities, geographical location might play a potential crucial part in shaping the fungal community living in a certain home considering a clearer geographical structure and stronger distance-decay relationship. This was confirmed by the results of Mantel test and DistLM analysis which demonstrated that georeferenced environmental variables, especially the diversity of vascular plants near the homes, determined the fungal community structure found in the examined homes (Tables 1 and S5). It has been proposed that the fungi detected inside homes were largely derived from outside the homes (Barberán et al., 2015a). In this study, the dominance of *Alternaria* and *Phoma* genera in the fungal communities across all the dust samples could support this assumption, because these two populations are usually detected as saprophytes and parasites in soil and plants (e.g., *Iris ensata* and *Medicago truncatula* which were present outdoors in this work), respectively, and thought to be outdoor allergens (Djebali, 2013; Barberán et al., 2015a; Luo et al., 2018). It has also been documented that fungal community in indoor air could be determined by that in outdoor air through dispersal either at a local or continental scale (Adams et al., 2013; Barberán et al., 2015a).

The differences in the geographical distribution patterns between bacterial and fungal communities living in the homes might be partially ascribed to the variations in dispersal abilities of different microbial populations and dispersal constraints, i.e., microbial populations that

can enter homes may be unevenly distributed across the geographical areas (Adams et al., 2013). Additionally, environmental variables (e.g., wind patterns and humidity) and housing characteristics (e.g., occupants' activities and the house itself) that select for specific microbial taxa in certain areas were also likely to explain such differences to some extent. It is important to note that many georeferenced environmental (e.g., wind direction and speed) and housing variables (e.g., ventilation type and rate) that were unavailable or not effectively obtained in this study might also contribute to some of the detected biogeographical patterns (Dannemiller et al., 2016; Li et al., 2019). Therefore, more explanatory variables should be taken into consideration in order to better understand the geographical distributions of microbes inside homes at various scales for future works.

4.3. ARGs in household dust

Surprisingly, microbial communities in the investigated dust samples harbored a great diversity and abundance of known ARGs, which were comparable to or even higher than those reported from soils added with sewage sludge (Chen et al., 2016), compost derived from pig manure (Zhou et al., 2019), and estuarine sediments (Zhu et al., 2017). These imply that household dust is an overlooked dissemination route and reservoir for antibiotic resistance. This implication was partially supported by several air-based studies which have proposed air transmission as an under-investigated pathway for spread of antibiotic resistance in a variety of environment contexts (Yooseph et al., 2013; Pal et al., 2016; Li et al., 2018). The explanations behind such a high diversity and abundance of ARGs in the household dust in the present study are probably resulted from the accumulation of antibiotic resistant bacteria coming from indoor and/or outdoor environments. On the one hand, it is assumed that some of the ARGs observed here were associated with home occupants and pets given the abundant clinical variant of the class 1 integron-integrase gene *intI1* across the dust samples (on average 0.06 copies per cell), because *intI1* gene abundance has been proposed to be an appropriate proxy for anthropogenic impacts, e.g., residents themselves and their domestic pets for this study (Gillings et al., 2015). The significant relationships between these occupancy-related factors and dust-associated antibiotic resistome (Fig. 6, see discussion below) further supported this assumption. On the other hand, atmospheric transport of bacterial taxa containing ARGs and even antibiotics from outdoor environments subjected to human-induced antibiotic exposure (e.g., clinic/hospitals and pharmaceutical companies) were also likely to account for the presence of ARGs in the household dust, as the aerial dispersal of bacterial populations carrying ARGs and veterinary antibiotics has already been found from large-scale cattle feed yards (McEachran et al., 2015) and also in urban air (Li et al., 2018). However, in the present study, the information related with antibiotics, for example, the history of antibiotic treatment for residents or pets before sampling and the presence of medical settings near the homes, is not available. Future studies should be performed taking these into consideration to better explain why the indoor environment in Beijing city encompasses such a great diversity and abundance of ARGs, and further reveal the origin of the household ARGs with the aid of more advanced technique such as metagenome sequencing analysis.

Genes conferring resistance to vancomycin and MLSB were highly abundant and prevalent for the surveyed dust, and significantly correlated with MGEs (Figs. 5; S9 and Table S8), indicating the possibility of spread and/or horizontal transfer of them through household dust. Both vancomycin and MLSB are effective in therapy of Gram-positive infectious diseases (Wang et al., 2014), and the former is perceived as one of the last resort for ward off bacterial pathogens like *Streptococcus pneumoniae* and *Enterococcus faecium* (Yang et al., 2015; Kakuta et al., 2020). Surprisingly, the normalized abundance of genes resistant to vancomycin observed here was about 1–2 orders of magnitude higher than that found in soils, composts and aquatic ecosystems (Su et al.,

2015; Chen et al., 2016; Zhu et al., 2017; Zhou et al., 2019). Such a high abundance and the consequently high possibility of dissemination and horizontal transfer of vancomycin resistance genes in the household environment of Beijing implied the potential requirement of policies regarding banning vancomycin and/or its analogues in mainland China, considering that an analogue of vancomycin, i.e., avoparcin, has been banned in some countries and regions since 1990s to mitigate the dispersal of vancomycin resistance (Zhu et al., 2017).

4.4. Bacterial community and housing factors shaping ARG profiles

Similar to bacterial community, it is unlikely that the geographical location of home drives the structure of ARGs inside homes because no clear geographical distribution patterns were detected in the ARG profiles across the sampling sites. The bacterial community consisted of *Actinobacteria*, *Firmicutes* and *Proteobacteria* phyla contributed most to the structural variance of dust-associated ARGs. *Actinobacteria* are renowned for antibiotic production, and considered to be a universal source phylum of ARGs as they usually carry diverse ARGs with a multitude of resistance mechanisms (Forsberg et al., 2014; Qian et al., 2016). Furthermore, members of *Firmicutes*, for instance, *Ureibacillus* genus, a dominant bacterial population observed here (Fig. 1a), is frequently related with antibiotic resistance (Su et al., 2015). It has also been demonstrated that *Proteobacteria* are prevalent carriers of integrons like *intI1* (Martinez et al., 2013). Therefore, it is reasonable that the shift of these bacterial phyla played a major part in forming the antibiotic resistome structure in the household dust. This finding was further supported by the result of network analysis as bacterial genera belonging to these three phyla were engaged in much larger numbers of robust correlations with specific ARGs than other phyla (Table S9). Moreover, we found housing factors, such as occupancy-related factors also played a nontrivial role in shaping the dust-associated antibiotic resistome. This might be partially explained by the fact that microbes shed from occupants (i.e., residents and pets) are likely to be putative ARG hosts and hence affect the structure of ARGs in the household environment. Another possible explanation could be that antibiotic selection pressure induced by anthropogenic activities, for example, antibiotic treatment for residents and pets, might directly influence the household resistome.

5. Conclusions

In summary, bacterial communities closely coexist with fungal communities in the household environment. Fungal communities living in homes showed a more distinct geographical distribution pattern than bacterial communities at the local scale. The bacterial communities found inside homes were most likely to be emanated from human skin, soil and plants, and mainly shaped by housing variables, notably human occupancy. By contrast, the fungal communities found inside homes were dominated by potential fungal allergens, and shaped by georeferenced environmental variables, especially vascular plant diversity. Moreover, the household microbes carried a considerably high diversity and abundance of ARGs. The profile of dust-related ARGs exhibited no clear geographical structure, and was primarily determined by certain bacterial phyla and occupancy-related factors. Taken together, these findings provide new insights into the local-scale distribution of microbiome and antibiotic resistome in the household environment, highlight the importance of anthropogenic influences and house location in shaping bacterial and fungal assemblages living in homes, respectively, and suggest the overlooked role of household dust in spread of antibiotic resistance.

Author contributions

Yong-Guan Zhu initiated the study, and participated in the experimental design. Long-Jun Ding designed the experiment, provided

sampling tools for volunteers, performed DNA extraction, analyzed the Illumina sequencing data. Xin-Yuan Zhou carried out the high-throughput quantitative PCR of antibiotic resistance genes and analyzed the data. Long-Jun Ding wrote the first draft of the manuscript, and all authors contributed to and approved the final manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank all the volunteers for their help in sample collection, and thank Mr. Liang Cheng for his help in plotting maps by ArcGIS software. This study was conducted with financial support by the National Natural Science Foundation, China (41977036 and 41601242), the National Key Research and Development Program, China (2017YFD0801502 and 2018YFC1801103), and Youth Innovation Promotion Association of Chinese Academy of Sciences, China.

Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.envint.2020.105702>.

References

- Abarenkov, K., Nilsson, R.H., Larsson, K.H., Alexander, I.J., Eberhardt, U., Erland, S., Hoiland, K., Kjoller, R., Larsson, E., Pennanen, T., Sen, R., Taylor, A.F.S., Tedersoo, L., Ursing, B.M., Vrålstad, T., Liimatainen, K., Peintner, U., Koljalg, U., 2010. The UNITE database for molecular identification of fungi—recent updates and future perspectives. *New Phytol.* 186 (2), 281–285.
- Adams, R.I., Miletto, M., Taylor, J.W., Bruns, T.D., 2013. Dispersal in microbes: fungi in indoor air are dominated by outdoor air and show dispersal limitation at short distances. *ISME J.* 7 (7), 1262–1273.
- Amend, A.S., Seifert, K.A., Samson, R., Bruns, T.D., 2010. Indoor fungal composition is geographically patterned and more diverse in temperate zones than in the tropics. *Proc. Natl. Acad. Sci. U. S. A.* 107 (31), 13748–13753.
- Barberán, A., Dunn, R.R., Reich, B.J., Pacifici, K., Laber, E.B., Menninger, H.L., Morton, J.M., Henley, J.B., Leff, J.W., Miller, S.L., Fierer, N., 2015a. The ecology of microscopic life in household dust. *Proc. R. Soc. B-Biol. Sci.* 282 (1814), 212–220.
- Barberán, A., Ladau, J., Leff, J.W., Pollard, K.S., Menninger, H.L., Dunn, R.R., Fierer, N., 2015b. Continental-scale distributions of dust-associated bacteria and fungi. *Proc. Natl. Acad. Sci. U. S. A.* 112 (18), 5756–5761.
- Bartram, A.K., Lynch, M.D.J., Stearns, J.C., Moreno-Hagelsieb, G., Neufeld, J.D., 2011. Generation of multimillion-sequence 16S rRNA gene libraries from complex microbial communities by assembling paired-end illumina reads. *Appl. Environ. Microbiol.* 77 (15), 3846–3852.
- Böbel, T.S., Hackl, S.B., Langgartner, D., Jarczok, M.N., Rohleder, N., Rook, G.A., Lowry, C.A., Gundel, H., Waller, C., Reber, S.O., 2018. Less immune activation following social stress in rural vs. urban participants raised with regular or no animal contact, respectively. *Proc. Natl. Acad. Sci. U. S. A.* 115 (20), 5259–5264.
- Bowers, R.M., McLetchie, S., Knight, R., Fierer, N., 2011. Spatial variability in airborne bacterial communities across land-use types and their relationship to the bacterial communities of potential source environments. *ISME J.* 5 (4), 601–612.
- Brilli, F., Fares, S., Ghirardo, A., de Visser, P., Calatayud, V., Muñoz, A., Annesi-Maesano, I., Sebastiani, F., Alivernini, F., Varriale, V., Menghini, F., 2018. Plants for sustainable improvement of indoor air quality. *Trends Plant Sci.* 23 (6), 507–512.
- Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Pena, A.G., Goodrich, J.K., Gordon, J.L., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Tumbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J., Knight, R., 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336.
- Chen, Q., An, X., Li, H., Su, J., Ma, Y., Zhu, Y.G., 2016. Long-term field application of sewage sludge increases the abundance of antibiotic resistance genes in soil. *Environ. Int.* 92–93, 1–10.
- Dannemiller, K.C., Gent, J.F., Leaderer, B.P., Peccia, J., 2016. Influence of housing characteristics on bacterial and fungal communities in homes of asthmatic children. *Indoor Air* 26 (2), 179–192.
- Ding, L.J., Su, J.Q., Li, H., Zhu, Y.G., Cao, Z.H., 2017. Bacterial succession along a long-term chronosequence of paddy soil in the Yangtze River Delta. *China. Soil Biol. Biochem.* 104, 59–67.
- Djebali, N., 2013. Aggressiveness and host range of *Phoma medicaginis* isolated from *Medicago* species growing in Tunisia. *Phytopathol. Mediterr.* 52 (1), 3–15.
- Du, P., Du, R., Ren, W., Lu, Z., Fu, P., 2018. Seasonal variation characteristic of inhalable microbial communities in PM_{2.5} in Beijing city, China. *Sci. Total Environ.* 610, 308–315.
- Ege, M.J., Mayer, M., Normand, A.C., Genuneit, J., Cookson, W.O.C.M., Braun-Fahrlander, C., Heederik, D., Piarroux, R., Mutius, E.V., 2011. Exposure to environmental microorganisms and childhood asthma. *N. Engl. J. Med.* 364, 701–709.
- Fernandez, C., Bevilacqua, E., Fernandez, N., Gajate, P., de la Camara, A.G., Garcimartin, M., Vives, R., Rodriguez, J., 2011. Asthma related to *Alternaria* sensitization: an analysis of skin-test and serum-specific IgE efficiency based on the bronchial provocation test. *Clin. Exp. Allergy* 41 (5), 649–656.
- Flores, G.E., Bates, S.T., Caporaso, J.G., Lauber, C.L., Leff, J.W., Knight, R., Fierer, N., 2013. Diversity, distribution and sources of bacteria in residential kitchens. *Environ. Microbiol.* 15 (2), 588–596.
- Forsberg, K.J., Patel, S., Gibson, M.K., Lauber, C.L., Knight, R., Fierer, N., Dantas, G., 2014. Bacterial phylogeny structures soil resistomes across habitats. *Nature* 509 (7502), 612–616.
- Gandara, A., Mota, L.C., Flores, C., Perez, H.R., Green, C.F., Gibbs, S.G., 2006. Isolation of *Staphylococcus aureus* and antibiotic resistant *Staphylococcus aureus* from residential indoor bioaerosols. *Environ. Health Perspect.* 114 (12), 1859–1864.
- Gilbert, J.A., Stephens, B., 2018. Microbiology of the built environment. *Nat. Rev. Microbiol.* 16, 661–670.
- Gillings, M.R., Gaze, W.H., Pruden, A., Smalla, K., Tiedje, J.M., Zhu, Y.G., 2015. Using the class 1 integron-integrase gene as a proxy for anthropogenic pollution. *ISME J.* 9 (6), 1269–1279.
- Gupta, S., 2017. Microbiome: puppy power. *Nature* 543 (7647), S48–S49.
- Haft, D.H., Tovchigrechko, A., 2012. High-speed microbial community profiling. *Nat. Methods* 9 (8), 793–794.
- Hartmann, E.M., Hickey, R., Hsu, T., Betancourt Román, C.M., Chen, J., Schwager, R., Kline, J., Brown, G.Z., Halden, R.U., Huttenhower, C., Green, J.L., 2016. Antimicrobial chemicals are associated with elevated antibiotic resistance genes in the indoor dust microbiome. *Environ. Sci. Technol.* 50 (18), 9807–9815.
- Hernandez, H., Martinez, L.R., 2018. Relationship of environmental disturbances and the infectious potential of fungi. *Microbiol. -SGM* 164 (3), 233–241.
- Hospodsky, D., Qian, J., Nazaroff, W.W., Yamamoto, N., Bibby, K., Rismani-Yazdi, H., Peccia, J., 2012. Human occupancy as a source of indoor airborne bacteria. *PLoS One* 7 (4), e34867.
- Kakuta, R., Nakano, R., Yano, H., Ozawa, D., Ohta, N., Matsuoka, T., Motoyoshi, N., Kawamoto, S., Saolo, Y., Katori, Y., Kaku, M., 2020. First two cases of infected aortic aneurysm caused by non-vaccine *Streptococcus pneumoniae* serotype 23A. *Ann. Lab. Med.* 40 (3), 270–273.
- Kemmel, S.W., Jones, E., Kline, J., Northcutt, D., Stenson, J., Womack, A.M., Bohannon, B.J.M., Brown, G.Z., Green, J.L., 2012. Architectural design influences the diversity and structure of the built environment microbiome. *ISME J.* 6 (8), 1469–1479.
- Knights, D., Kuczynski, J., Charlson, E.S., Zaneveld, J., Mozer, M.C., Collman, R.G., Bushman, F.D., Knight, R., Kelley, S.T., 2011. Bayesian community-wide culture-independent microbial source tracking. *Nat. Methods* 8 (9), 761–763.
- Kolikonda, M.K., Jayakumar, P., Sriramula, S., Lippmann, S., 2017. *Kocuria kristinae* infection during adalimumab treatment. *Postgrad. Med.* 129 (2), 296–298.
- Lecomte, J., St-Arnaud, M., Hijri, M., 2011. Isolation and identification of soil bacteria growing at the expense of arbuscular mycorrhizal fungi. *FEMS Microbiol. Lett.* 317 (1), 43–51.
- Leung, M.H.Y., Chan, C.K.K., Lee, P.K.H., 2016. Skin fungal community and its correlation with bacterial community of urban Chinese individuals. *Microbiome* 4, 46.
- Leung, M.H.Y., Tong, X., Wilkins, D., Cheung, H.H.L., Lee, P.K.H., 2018. Individual and household attributes influence the dynamics of the personal skin microbiota and its association network. *Microbiome* 6, 26.
- Li, H., Zhou, X.Y., Yang, X.R., Zhu, Y.G., Hong, Y.W., Su, J.Q., 2019. Spatial and seasonal variation of the airborne microbiome in a rapidly developing city of China. *Sci. Total Environ.* 665, 61–68.
- Li, J., Cao, J., Zhu, Y.G., Chen, Q.L., Shen, F., Wu, Y., Xu, S., Fan, H., Da, G., Huang, R.J., Wang, J., de Jesus, A.L., Morawska, L., Chan, C.K., Peccia, J., Yao, M., 2018. Global survey of antibiotic resistance genes in air. *Environ. Sci. Technol.* 52 (19), 10975–10984.
- Loof, T., Johnson, T.A., Allen, H.K., Bayles, D.O., Alt, D.P., Stedtfeld, R.D., Sul, W.J., Stedtfeld, T.M., Chai, B., Cole, J.R., Hashsham, S.A., Tiedje, J.M., Stanton, T.B., 2012. In-feed antibiotic effects on the swine intestinal microbiome. *Proc. Natl. Acad. Sci. U. S. A.* 109 (5), 1691–1696.
- Luo, H., Tao, Y.Q., Fan, X.Y., Oh, S.K., Lu, H.X., Deng, J.X., 2018. Identification and characterization of *Alternaria iridialstralis* causing leaf spot on *Iris ensata* in China. *Mycobiology* 46 (2), 168–171.
- Martin, L.J., Adams, R.L., Bateman, A., Bik, H.M., Hawks, J., Hird, S.M., Hughes, D., Kemmel, S.W., Kinney, K., Kolokotronis, S.O., Levy, G., McClain, C., Meadow, J.F., Medina, R.F., Mhuirreach, G., Moreau, C.S., Munshi-South, J., Nichols, L.M., Palmer, C., Popova, L., Schal, C., Taubel, M., Trautwein, M., Ugalde, J.A., Dunn, R.R., 2015. Evolution of the indoor biome. *Trends Ecol. Evol.* 30 (4), 223–232.
- Martinez, E., Djordjevic, S., Stokes, H., Roy Chowdhury, P., 2013. Mobilized integrons: team players in the spread of antibiotic resistance genes. In: Gophna, U. (Ed.), *Lateral Gene Transfer in Evolution*. Springer, New York, pp. 79–103.
- McEachran, A.D., Blackwell, B.R., Hanson, J.D., Wooten, K.J., Mayer, G.D., Cox, S.B., Smith, P.N., 2015. Antibiotics, bacteria, and antibiotic resistance genes: aerial transport from cattle feed yards via particulate matter. *Environ. Health Perspect.* 123 (4), 337–343.
- Mubanga, M., Byberg, L., Nowak, C., Egenvall, A., Magnusson, P.K., Ingelsson, E., Fall, T., 2017. Dog ownership and the risk of cardiovascular disease and death – a nationwide

- cohort study. *Sci. Rep.* 7, 15821.
- Oh, J., Freeman, A.F., Comparative Sequencing Program, N.I.S.C., Park, M., Sokolic, R., Candotti, F., Holland, S.M., Segre, J.A., Kong, H.H., 2013. The altered landscape of the human skin microbiome in patients with primary immunodeficiencies. *Genome Res.* 23 (12), 2103–2114.
- Pal, C., Bengtsson-Palme, J., Kristiansson, E., Larsson, D.G.J., 2016. The structure and diversity of human, animal and environmental resistomes. *Microbiome* 4, 54.
- Peleg, A.Y., Hogan, D.A., Mylonakis, E., 2010. Medically important bacterial-fungal interactions. *Nat. Rev. Microbiol.* 8 (5), 340–349.
- Pereira, E.M., Oliveira, F.L.F., Schuenck, R.P., Zoletti, G.O., dos Santos, K.R.N., 2010. Detection of *Staphylococcus lugdunensis* by a new species-specific PCR based on the *fbi* gene. *FEMS Immunol. Med. Microbiol.* 58 (2), 295–298.
- Qian, X., Sun, W., Gu, J., Wang, X.J., Sun, J.J., Yin, Y.N., Duan, M.L., 2016. Variable effects of oxytetracycline on antibiotic resistance gene abundance and the bacterial community during aerobic composting of cow manure. *J. Hazard. Mater.* 315, 61–69.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glockner, F.O., 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids. Res.* 41 (D1), D590–D596.
- Redford, A., Bowers, R., Knight, R., Linhart, Y., Fierer, N., 2010. The ecology of the phyllosphere: Geographic and phylogenetic variability in the distribution of bacteria on tree leaves. *Environ. Microbiol.* 12 (11), 2885–2893.
- Rintala, H., Pitkaranta, M., Taubel, M., 2012. Microbial communities associated with house dust. *Adv. Appl. Microbiol.* 78, 75–120.
- Rook, G.A., Raison, C.L., Lowry, C.A., 2013. Childhood microbial experience, immunoregulation, inflammation and adult susceptibility to psychosocial stressors and depression in rich and poor countries. *Evol. Med. Public Health* 1, 14–17.
- Ross, A.A., Müller, K.M., Weese, J.S., Neufeld, J.D., 2018. Comprehensive skin microbiome analysis reveals the uniqueness of human skin and evidence for phyllosymbiosis within the class Mammalia. *Proc. Natl. Acad. Sci. U. S. A.* 115 (25), E5786–E5795.
- Segata, N., Waldron, L., Ballarini, A., Narasimhan, V., Jousson, O., Huttenhower, C., 2012. Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat. Methods* 9 (8), 811–814.
- Shade, A., Jones, S.E., Caporaso, J.G., Handelsman, J., Knight, R., Fierer, N., Gilbert, J.A., 2014. Conditionally rare taxa disproportionately contribute to temporal changes in microbial diversity. *mBio* 5 (4), e01371–e01374.
- Shi, Y., Li, Y., Xiang, X., Sun, R., Yang, T., He, D., Zhang, K., Ni, Y., Zhu, Y.G., Adams, J.M., Chu, H., 2018. Spatial scale affects the relative role of stochasticity versus determinism in soil bacterial communities in wheat fields across the North China Plain. *Microbiome* 6, 27.
- Su, J.Q., An, X.L., Li, B., Chen, Q.L., Gillings, M.R., Chen, H., Zhang, T., Zhu, Y.G., 2017. Metagenomics of urban sewage identifies an extensively shared antibiotic resistome in China. *Microbiome* 5, 84.
- Su, J.Q., Wei, B., Ou-Yang, W.Y., Huang, F.Y., Zhao, Y., Xu, H.J., Zhu, Y.G., 2015. Antibiotic resistome and its association with bacterial communities during sewage sludge composting. *Environ. Sci. Technol.* 49 (12), 7356–7363.
- van Hoek, A.H., Mevius, D., Guerra, B., Mullany, P., Roberts, A.P., Aarts, H.J.M., 2011. Acquired antibiotic resistance genes: an overview. *Front. Microbiol.* 2, 203.
- Wang, J., Wang, P., Wang, X., Zheng, Y., Xiao, Y., 2014. Use and prescription of antibiotics in primary health care settings in china. *JAMA Intern. Med.* 174 (12), 1914–1920.
- World Health Organization, 2014. Antimicrobial Resistance: Global Report on Surveillance. <http://www.who.int/drugresistance/documents/surveillancereport/en/>.
- Xiang, Q., Zhu, D., Giles, M., Neilson, R., Yang, X.R., Qiao, M., Chen, Q.L., 2020. Agricultural activities affect the pattern of the resistome within the phyllosphere microbiome in peri-urban environments. *J. Hazard. Mater.* 382, 121068.
- Yang, J.X., Li, T., Ning, Y.Z., Shao, D.H., Liu, J., Wang, S.Q., Liang, G.W., 2015. Molecular characterization of resistance, virulence and clonality in vancomycin-resistant *Enterococcus faecium* and *Enterococcus faecalis*: a hospital-based study in Beijing China. *Infect. Genet. Evol.* 33, 253–260.
- Yooseph, S., Andrews-Pfannkoch, C., Tenney, A., McQuaid, J., Williamson, S., Thiagarajan, M., Bami, D., Zeigler-Allen, L., Hoffman, J., Goll, J.B., Fadrosh, D., Glass, J., Adams, M.D., Friedman, R., Venter, J.C., 2013. A metagenomic framework for the study of airborne microbial communities. *PLoS One* 8 (12), e81862.
- Zhou, X., Qiao, M., Su, J.Q., Wang, Y., Cao, Z.H., Cheng, W.D., Zhu, Y.G., 2019. Turing pig manure into biochar can effectively mitigate antibiotic resistance genes as organic fertilizer. *Sci. Total Environ.* 649, 902–908.
- Zhu, Y.G., Zhao, Y., Li, B., Huang, C.L., Zhang, S.Y., Yu, S., Chen, Y.S., Zhang, T., Gillings, M.R., Su, J.Q., 2017. Continental-scale pollution of estuaries with antibiotic resistance genes. *Nat. Microbiol.* 2 (4), 16270.

Supplementary material for

Microbiome and antibiotic resistome in household dust from Beijing, China

Long-Jun Ding^a, Xin-Yuan Zhou^{b,c}, Yong-Guan Zhu^{a,b,c,*}

^aState Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China

^bUniversity of Chinese Academy of Sciences, Beijing 100049, China

^cKey Laboratory of Urban Environment and Health, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen 361021, China

*Correspondence should be addressed to [Yong-Guan Zhu](#) at

^aState Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, 18 Shuangqing Road, Haidian District, Beijing 100085, China
E-mail: ygzhu@rcees.ac.cn; Phone: 86-10-62936940.

This SM file contains 38 pages of Supplementary Methods, Supplementary Results, 9 Tables (Table S1 to Table S9), 11 Figures (Figure S1 to Figure S11), and Supplementary References.

Supplementary Methods

Sample collection and characterization

The households examined in this study were selected to cover most of the administrative districts of Beijing city. Each participant was provided with a sampling protocol and a sampling kit containing sterile cotton-tipped swabs (Shandong Longhai Medical Devices Co., Ltd., China). They were instructed to sample dust when there was little external disturbance (e.g., entertaining friends) in the house. Approximately a 100-cm² area of interior windowsill surface was vigorously rubbed for 30 seconds in back-and-forth motions, using the sterile swabs pre-moistened with a sterile solution of 0.15 M NaCl and 0.1% Tween 20 (Leung et al., 2018). The participants wore sterile latex gloves and surgical face masks throughout the sampling period to preclude pollution while talking or handling swabs. After sampling, the swabs were rapidly transported to the lab within 12 hours, and stored at –80 °C for DNA extraction. Meanwhile, un-used swabs in the sampling kit are regarded as control samples, and they were subjected to the same molecular analyses (e.g., DNA extraction and amplification) as used swabs.

Geographical coordinates (latitude and longitude) were derived from house location information (i.e., home address). Coordinates were then used to obtain information of georeferenced environmental variables for each household, including population density, vascular plant diversity, percentage of woody vegetation cover, hog density, elevation, distance to city center, main vascular plant species outdoors, and climatic factors like mean annual temperature (MAT), mean annual precipitation (MAP), and mean annual concentrations of fine particulate matter (PM_{2.5}) and inhalable particles (PM₁₀) (see details in [Table S1](#)). The data of these environmental variables except elevation, distance to city center, MAT and MAP were collected or calculated from the Beijing Regional Statistical Yearbook (2016). The data of elevation and distance to city center were acquired based on the

geographical coordinates of the sampled house. The data of MAT and MAP were averaged for the years 1970-2000, and downloaded from World Clim-Global Climate Data (<http://www.worldclim.org/>). The information of the home design and occupancy was obtained through the returned questionnaire. The list of housing variables is also shown in [Table S1](#).

Illumina sequencing

For bacterial analysis, we employed a primer pair 515F (GTGCCAGCMGCCGCGGTAA) and 806R (GGACTACHVGGGTWTCTAAT) to target the V4 hypervariable region of the 16S rRNA gene. For fungal analysis, the internal transcribed spacer (ITS) region 1 of the rRNA operon was sequenced using a primer pair ITS1F (CTTGGTCATTTAGAGGAAGTAA) and ITS2 (GCTGCGTTCTTCATCGATGC) (Barberán et al., 2015). The primers contained appropriate Illumina adapters and 12-bp barcodes for sample identification. For each sample, PCR was performed in 20- μ l triplicate reactions, using 10 ng of template DNA, 0.2 μ M of each primer, 0.25 mM of each dNTP, 4 μ l of 5 $\times$ FastPfu Buffer, 0.4 μ l of FastPfu Polymerase, and 0.2 μ l of bull serum albumin (BSA; 20 mg ml⁻¹). PCR thermal procedures were as follows with 27 and 35 cycles for amplification of bacteria and fungi, respectively: initial denaturation for 5 min at 95°C; 27 or 35 cycles of 30 s at 95°C, 30 s at 55°C, and 45 s at 72°C; and a final extension of 10 min at 72°C. Moreover, negative controls including no-template (three replications) and template from un-used swabs (three replications) were also subjected to amplification. Consequently, neither the bacterial 16S rRNA sequences nor fungal ITS sequences of these negative controls could be amplified, indicating that both PCR reagents and sample DNA were not polluted. The PCR products successfully amplified were purified with the AxyPrep DNA Gel Extraction Kit (Axygen Biosciences, Union, CA, USA), and subsequently quantified with the Pico Green using the QuantiFluor™-ST (Promega, Madison, WI, USA). The purified PCR products were

pooled in equimolar concentration for paired-end sequencing on an Illumina MiSeq PE300 platform by Majorbio Bio-Pharm Technology Co., Ltd. (Shanghai, China).

Analyses of source tracking and bacterial source indicators

In order to discern the relative contribution of different putative sources to the bacterial communities living in the investigated homes, two complementary methods were used, including Bayesian source tracking and indicator taxa analysis (Flores et al., 2013). Six putative source environments were chosen for both methods, including soil, plants, human skin, human oral cavity, human/livestock stool and insects, as they are regarded as potential important sources of indoor microbes (Barberán et al., 2015). The Bayesian source tracking approach was carried out based on the SourceTracker algorithm in QIIME (Knights et al., 2011), in which samples from different putative source environments are compared with those in various examined habitats (home environment herein) (Dunn et al., 2013). In brief, soil source taxa were obtained from Griffiths et al. (2011) and Leff et al. (2015), plant source taxa were obtained from several tree species and vegetables (Redford et al., 2010; Flores et al., 2013), data of human-related sources were from Costello et al. (2009), and taxa of exclusive symbionts and/or pathogens of insect source were from Feldhaar (2011) and Engel and Moran (2013). Meanwhile, to identify potential bacterial populations indicative of these particular source environments, indicator taxa analysis was carried out with bacterial taxa grouped at the genus level, using packages of vegan 2.3-5, indicpecies 1.7.6, and labdsv 1.8-0 in R 3.1.3. Taxa with an indicator value > 0.6 and $P < 0.01$ were chosen to be indicator taxa as they were thought to be highly related with a specific source environment (Dunn et al., 2013). The list of bacterial indicator taxa of different putative source environments identified in this study is presented in [Table S3](#).

Statistical and network analyses

For all statistical analyses, significance level was set as $P < 0.05$ or with listed P -values.

SPSS V20.0 (IBM, USA) software was used to implement standard statistical tests such as nonparametric Mann-Whitney test for significance and correlation analysis.

For data analyses associated with dust-related microbial community, the collinearity of all the examined candidate explanatory factors was reduced by discerning strongly correlated factors through principle component analysis (PCA) and computing variance inflation factors (VIFs) (Barberán et al., 2015). After collinearity analysis, 20 independent explanatory variables (including environmental and housing variables) were left for subsequent analyses (listed in [Table S1](#), except information of main vascular plant species outdoors and houseplant species). The relationships between dust-associated microbial community structure and the remaining explanatory factors were investigated using Mantel test. The contribution of each explanatory factor to the variations of bacterial or fungal community structure in dust samples was then calculated by DistLM forward3 ($P < 0.05$, permutation = 999) (Shi et al., 2018).

For data analyses associated with dust-related ARGs, canonical correspondence analysis (CCA) was performed to reveal relationships between the structural variations of ARGs and four components including bacterial community (at the phylum level), environmental and housing variables, and MGEs. The percentage of variance explained by each component was further distinguished by variation partitioning analysis (VPA).

In addition, the dissimilarities of microbial (i.e., bacterial and fungal) communities and ARGs among all dust samples were analyzed by a non-metric multidimensional scaling (NMDS) ordination based on a Bray-Curtis distance matrix. Diversity index, PCA, Mantel test, CCA, VPA and NMDS mentioned above were performed with packages of vegan 2.3-5 and ecodist 2.0.1 in R. Furthermore, distance-decay analyses of microbial communities and ARG profiles were carried out with packages of labdsv 1.8-0, vegan 2.3-5, and geosphere 1.5-10 in R. ArcGIS V10.0 software was employed to map the NMDS scores and the proportional abundances of microbial taxa at the spatial scale examined (Su et al., 2017).

To uncover co-occurrence and co-exclusion patterns between bacterial and fungal communities (at the OTU level) in dust samples, a correlation matrix (based on Spearman's rank correlation) was built up as described by Su et al. (2017), using R package pheatmap 1.0.12 with P -value adjusted by the false discovery rate (FDR) method to decrease false-positive results (Chen et al., 2016). A correlation between any two OTUs with correlation coefficient (ρ) ≥ 0.7 and $P < 0.01$ was denoted as a statistically strong co-occurrence relationship, and with $\rho \leq -0.7$ and $P < 0.01$ as a strong co-exclusion relationship (Barberán et al., 2012). The consequent correlation matrix was then visualized by a network plot using Gephi 0.9.2. The co-occurrence pattern between dust-associated bacterial community (at the genus level) and ARGs was also surveyed in a similar way with $\rho \geq 0.7$ and $P < 0.01$ being considered as a strong co-occurrence relationship (Su et al., 2017).

Supplementary Results

Overview of sequencing results of bacterial and fungal communities in household dust

Overall, a sum of 2,648,020 and 2,720,186 qualified bacterial and fungal sequences were clustered into 5,738 and 2,461 bacterial and fungal OTUs, respectively, with bacterial OTUs ranging from 315 to 2,554 per sample (mean = 1,511) and fungal OTUs ranging from 46 to 900 (mean = 372) (Table S4). Among these OTUs detected, a set of 45 bacterial and 9 fungal OTUs being shared by all samples, contributing to 37.1% and 58.3% of the bacterial 16S rRNA gene and fungal ITS sequences, respectively. In addition, 425 and 422 rare bacterial and fungal OTUs appeared only in one sample, occupying 0.21% and 0.13% of the 16S rRNA gene and ITS sequences, respectively, an extremely minor proportion of the microbial communities found across the samples.

Table S1 Information of geographic location, environmental variables and housing variables for sampling sites in Beijing city.

See Attached Excel File.

Table S2 Information of 296 primers for high-throughput quantitative PCR.

See Attached Excel File.

Table S3 Bacterial populations identified as indicator taxa of different potential source environments at the genus level.¹

Potential environment	source	Bacterial indicator taxa ²
Human skin		<i>Kocuria</i> (n = 82), <i>Micrococcus</i> (n = 82), <i>Friedmanniella</i> (n = 82), <i>Propionibacterium</i> (n = 76), <i>Salinicoccus</i> (n = 56), <i>Staphylococcus</i> (n = 63), <i>Alloiococcus</i> (n = 47), <i>Anaerococcus</i> (n = 68), <i>Finegoldia</i> (n = 57), <i>Peptoniphilus</i> (n = 53), <i>Rubellimicrobium</i> (n = 82), <i>Roseomonas</i> (n = 79), <i>Enhydrobacter</i> (n = 78), <i>Deinococcus</i> (n = 82)
Soil		<i>Arthrobacter</i> (n = 82), <i>Candidatus Koribacter</i> (n = 10), <i>Candidatus Solibacter</i> (n = 45), <i>Phycococcus</i> (n = 72), <i>Mycobacterium</i> (n = 82), <i>Streptomyces</i> (n = 81), <i>Rubrobacter</i> (n = 81), <i>Sporocytophaga</i> (n = 46), <i>Flavobacterium</i> (n = 73), <i>Chitinophaga</i> (n = 23), <i>Niastella</i> (n = 8), <i>Sediminibacterium</i> (n = 25), <i>Anoxybacillus</i> (n = 15), <i>Cohnella</i> (n = 72), <i>Gemmatimonas</i> (n = 39), <i>JG37-AG-70</i> (n = 3), <i>Nitrospira</i> (n = 70), <i>Gemmata</i> (n = 67), <i>A17</i> (n = 32), <i>Pirellula</i> (n = 46), <i>Planctomyces</i> (n = 73), <i>Phenylobacterium</i> (n = 77), <i>Bradyrhizobium</i> (n = 3), <i>Devosia</i> (n = 81), <i>Hyphomicrobium</i> (n = 55), <i>Achromobacter</i> (n = 67), <i>Burkholderia</i> (n = 79), <i>Limnohabitans</i> (n = 23), <i>Methylibium</i> (n = 58), <i>Ramlibacter</i> (n = 26), <i>Rubrivivax</i> (n = 14), <i>Janthinobacterium</i> (n = 75), <i>Bdellovibrio</i> (n = 80), <i>Geobacter</i> (n = 14), <i>Plesiocystis</i> (n = 31), <i>Aquicella</i> (n = 18), <i>Legionella</i> (n = 27), <i>Steroidobacter</i> (n = 73), <i>Dokdonella</i> (n = 54), <i>Lysobacter</i> (n = 79), <i>Rhodanobacter</i> (n = 11), <i>Pedosphaera</i> (n = 1), <i>Candidatus Xiphinematobacter</i> (n = 60), <i>Chthoniobacter</i> (n = 22), <i>DA101</i> (n = 50), <i>Ellin506</i> (n = 49), <i>OR-59</i> (n = 8)
Plants		<i>Alkanindiges</i> (n = 54), <i>Aneurinibacillus</i> (n = 44), <i>Bosea</i> (n = 71), <i>Burkholderia-Paraburkholderia</i> (n = 80), <i>Curtobacterium</i> (n = 82), <i>Delftia</i> (n = 77), <i>Massilia</i> (n = 82), <i>Mesorhizobium</i> (n = 75), <i>Novosphingobium</i> (n = 80), <i>Paenibacillus</i> (n = 79), <i>Pedobacter</i> (n = 74), <i>Rhizobium</i> (n = 81), <i>Stenotrophomonas</i> (n = 75), <i>Sphingomonas</i> (n = 82)
Insects		<i>Arsenophonus</i> (n = 5), <i>Brevibacterium</i> (n = 77), <i>Cellulomonas</i> (n = 82), <i>Xylanimicrobium</i> (n = 37), <i>Candidatus Azobacteroides</i> (n = 1), <i>Dysgonomonas</i> (n = 10), <i>Paludibacter</i> (n = 17), <i>Tannerella</i> (n = 6), <i>Rikenella</i> (n = 1), <i>Blattabacterium</i> (n = 1), <i>Mucispirillum</i> (n = 2), <i>Lysinibacillus</i> (n = 69), <i>Enterococcus</i> (n = 63), <i>Vagococcus</i> (n = 16), <i>Lactobacillus</i> (n = 77), <i>Leuconostoc</i> (n = 60), <i>Weissella</i> (n = 20), <i>Lactococcus</i> (n = 52), <i>Turicibacter</i> (n = 81), <i>Oxobacter</i> (n = 40), <i>Dehalobacterium</i> (n = 11), <i>Pseudoramibacter</i> <i>Eubacterium</i> (n = 2), <i>Dehalobacter</i> <i>Syntrophobotulus</i> (n = 3), <i>Desulfosporosinus</i> (n = 49), <i>Syntrophomonas</i> (n = 32), <i>Sporomusa</i> (n = 31), <i>Anaerovorax</i> (n = 13), <i>Cetobacterium</i> (n = 13), <i>Wolbachia</i> (n = 3), <i>Dechloromonas</i> (n = 56), <i>Dok59</i> (n = 1), <i>Propionivibrio</i> (n = 6), <i>Desulfobulbus</i> (n = 2), <i>Acholeplasma</i> (n = 1), <i>Desulfovibrio</i> (n = 7), <i>Arcobacter</i> (n = 12), <i>Citrobacter</i> (n = 6), <i>Erwinia</i> (n = 29), <i>Treponema</i> (n = 29)
Human oral cavity		<i>Rothia</i> (n = 60), <i>Atopobium</i> (n = 15), <i>Porphyromonas</i> (n = 46), <i>Prevotella</i> (n = 65), <i>Granulicatella</i> (n = 22), <i>Streptococcus</i> (n = 78), <i>Catonella</i> (n = 5), <i>Moryella</i> (n = 9), <i>Oribacterium</i> (n = 7), <i>Veillonella</i> (n = 31), <i>Bulleidia</i> (n = 16), <i>Leptotrichia</i> (n = 35), <i>Neisseria</i> (n = 50), <i>Campylobacter</i> (n = 14), <i>Actinobacillus</i> (n = 1), <i>Aggregatibacter</i> (n = 21), <i>Haemophilus</i> (n = 41)

Human/livestock stool	<i>Collinsella</i> (n = 57), <i>Paraprevotella</i> (n = 33), <i>Lachnospira</i> (n = 33), <i>Faecalibacterium</i> (n = 40), <i>Dialister</i> (n = 38), <i>Sutterella</i> (n = 31), <i>Bilophila</i> (n = 9)
-----------------------	---

¹Indicator taxa analysis was performed with bacterial taxa classified at the genus level.

²The number of home dust samples where each bacterial indicator taxon was present in parenthesis.

Table S4 Overview of Illumina sequencing results targeting bacterial and fungal communities in all dust samples.

Sample ID	Number of qualified reads		Number of OTUs	
	Bacteria	Fungi	Bacteria	Fungi
HJC_DX	26145	35726	2258	636
LY_TZ	36680	30699	1548	258
WXH_DX	28755	33233	1983	661
JRY_DX	37997	37775	1835	337
DX_9_DX	30198	36184	1923	440
WJY_DX	31070	39534	315	117
HC_DX	34187	NA ¹	510	NA
BLJ_DX	14685	35053	931	71
HY_FT	26261	42153	497	137
DY_FT	39532	42146	1852	445
LX_FT	40530	44458	1955	298
HT_FT	28138	29903	2212	214
RWY_FT	35146	32599	563	116
LXM_FT	39581	39345	2149	565
DWC_CY	42337	31363	1171	86
WXJ_4_FT	33013	36675	826	216
DX_13_DC	31407	33937	781	88
DX_15_HD	35628	38500	1546	188
ZCX_SJS	32085	34393	2426	737
LP_DC	16342	40631	879	83
DX_17_HD	30596	39820	1745	385
DX_5_XC	33680	33186	716	167
DX_2_CY	38791	41006	2280	359
DX_10_XC	29757	38415	1623	402
DX_16_CY	39365	31918	1991	330
PZ_HD	37238	31248	1881	553
CC_HD	29987	43388	1625	418
DX_19_HD	39624	40981	1507	57
WQ_HD	27921	41237	1970	612
DX_6_HD	38956	37163	2118	348
DX_3_HD	37347	31796	605	190
LJ_HD	32255	36434	1848	525
BFQ_DC	28866	41864	591	81
CXJ_HD	33740	NA	1549	NA
DX_18_HD	36430	NA	688	NA
LXX_CY	27257	40220	761	73
DX_4_HD	28614	34008	1253	72
WXJ_1_CY	28609	34996	1714	544
WXJ_7_CY	32141	34824	1546	549
WXJ_10_CY	25372	31237	1411	442
CYS_HD	34270	30471	1726	341
DGL_CY	32729	41003	1797	288
ZYT_HD	31273	29945	1734	446
DX_11_HD	30035	32275	1726	46
HH_CY	24732	39700	1034	629
DX_7_HD	34985	31947	1344	295
LH_CY	35912	33519	1716	166
BR_CY	30693	43597	2085	738
XZF_CY	32104	40474	2334	732
CQ_CY	35110	31679	702	115
LHQ_CY	31929	33127	1986	900

Table S4 Overview of Illumina sequencing results targeting bacterial and fungal communities in all dust samples (*Continued*).

Sample ID	Number of qualified reads		Number of OTUs	
	Bacteria	Fungi	Bacteria	Fungi
WXJ_8_CY	35302	32163	1651	674
JY_HD	35353	31774	844	201
GSH_CY	30754	44280	852	247
ZYG_CY	30425	39384	1190	746
QM_HD	29061	NA	1005	NA
WPP_HD	32075	NA	663	NA
GGX_CY	36253	31714	2254	717
DX_14_HD	28073	32610	860	223
RJG_CY	32668	37085	1858	375
LXY_CY	32681	32730	1796	488
BL_HD	33992	44144	2432	330
XF_CP	30777	31070	1522	628
SGX_QH	35088	34139	2377	808
ZRL_HD	37099	39685	1354	240
WXJ_5_CY	41108	37179	533	144
YYG_CP	36996	38899	1370	214
QJL_HD	36952	34918	1705	470
HLL_CY	30623	34557	2493	687
SZX_HD	33075	29615	2041	153
SGX_WQ	26717	37614	1972	798
DX_HD	27275	39242	1816	301
HY_CP	40134	NA	653	NA
LYR_HD	30696	35544	856	71
DX_12_CP	17546	34275	401	65
DX_1_CP	29837	37575	986	151
ZLM_CP	37665	37475	2554	647
SJP_CP	24268	31314	2076	615
WXJ_6_SY	32171	NA	1634	NA
ZY_CP	28097	36805	2295	332
LX_CP	35059	42181	2304	343
DLJ_CP	30165	36430	1813	710

¹The samples marked by NA were removed from sequencing due to failed amplification.

Table S5 Relationships between explanatory variables and structure of bacterial or fungal community found in dust samples.

Class	Variable ¹	Bacterial community		Fungal community	
		r_M	P	r_M	P
Environmental variables	Population density	0.1285	0.049*	0.1828	0.011*
	Vascular plant diversity	0.1091	0.074	0.2571	0.002*
	Woody vegetation cover	0.0581	0.214	0.1426	0.025*
	PM _{2.5} concentration	0.0346	0.221	0.1685	0.019*
	PM ₁₀ concentration	0.0317	0.285	0.1031	0.057
	Hog density	0.0151	0.333	0.0999	0.068
	MAT	-0.0701	0.821	-0.0692	0.830
	MAP	-0.0509	0.817	-0.0645	0.836
	Elevation	-0.0234	0.654	0.0875	0.091
	Distance to city center	0.0360	0.293	0.0171	0.390
Housing variables	Number of habitants	0.3746	0.001*	0.0714	0.054
	Number of dogs	0.1965	0.028*	0.0708	0.058
	Number of cats	0.1645	0.035*	0.0385	0.272
	Female : Male ratio	0.1564	0.043*	0.0356	0.321
	Number of houseplants	0.1312	0.047*	0.1234	0.055
	Floor height	0.0673	0.150	-0.0480	0.787
	House space	-0.0779	0.879	-0.1303	0.980
	Living room space	-0.0413	0.712	-0.0190	0.607
	Number of carpets	-0.1226	0.910	-0.1621	0.996
	Construction age	0.0294	0.355	-0.0859	0.860

¹Woody vegetation cover: percentage of woody vegetation cover; PM_{2.5}: mean annual concentration of fine particulate matter; PM₁₀: mean annual concentration of inhalable particles; MAT: mean annual temperature; MAP: mean annual precipitation.

* $P < 0.05$.

Table S6 Significant correlations between OTUs of fungal and bacterial origins obtained from dust samples.

See Attached Excel File.

Table S7 Potential bacterial pathogens identified from the dust samples.

Genus	Species/Strains	NCBI reference sequence
<i>Staphylococcus</i>	<i>Staphylococcus lugdunensis</i> HKU09-01	NC_013893.1
	<i>Staphylococcus aureus</i> RF122	NC_007622
	<i>Staphylococcus carnosus</i> subsp. <i>carnosus</i> TM300	NC_012121.1
	<i>Staphylococcus epidermidis</i> isolate M99	AY_699287.1
	<i>Staphylococcus pseudintermedius</i> HKU10-03	NC_014925.1
	<i>Staphylococcus aprophyticus</i>	L_37596.1
<i>Bacillus</i>	<i>Bacillus megaterium</i> QM B1551	NC_014019.1
	<i>Bacillus cellulosilyticus</i> DSM 2522	NC_014829.1
	<i>Bacillus cereus</i> biovar <i>anthracis</i> str. CI	NC_014335.1
	<i>Bacillus clausii</i> KSM-K16	NC_006582.1
	<i>Bacillus cytotoxicus</i> NVH 391-98	NC_009674.1
	<i>Bacillus halodurans</i> C-125	NC_002570.2
	<i>Bacillus licheniformis</i> ATCC 14580	NC_006270.3
	<i>Bacillus pseudofirmus</i> OF4	NC_013791.2
	<i>Bacillus pumilus</i> SAFR-032	NC_009848.1
	<i>Bacillus subtilis</i> BSn5	NC_014976.1
	<i>Bacillus subtilis</i> subsp. <i>spizizenii</i> str. W23	NC_014479.1
<i>Pseudomonas</i>	<i>Bacillus thuringiensis</i> BMB171	NC_014171.1
	<i>Pseudomonas putida</i> GB-1	NC_010322.1
	<i>Pseudomonas putida</i> W619	NC_010501.1
	<i>Pseudomonas aeruginosa</i> UCBPP-PA14	NC_008463
	<i>Pseudomonas entomophila</i> str. L48	NC_008027.1
	<i>Pseudomonas fluorescens</i> Pf0-1	NC_007492
	<i>Pseudomonas fluorescens</i> SBW25	NC_012660
	<i>Pseudomonas mendocina</i> ymp	NC_009439.1
	<i>Pseudomonas protegens</i> Pf-5	NC_004129.6
	<i>Pseudomonas stutzeri</i> A1501	NC_009434.1
	<i>Pseudomonas syringae</i> pv. <i>phaseolicola</i> 1448A	NC_005773.3
<i>Streptococcus</i>	<i>Pseudomonas syringae</i> pv. <i>syringae</i> B728a	NC_007005.1
	<i>Streptococcus agalactiae</i> A909	NC_007432
	<i>Streptococcus gallolyticus</i> str. ATCC BAA-2069	NC_015215.1
	<i>Streptococcus gordonii</i> str. Challis substr. CH1	NC_009785
	<i>Streptococcus mitis</i> B6	NC_013853.1
	<i>Streptococcus mutans</i> UA159	NC_004350
	<i>Streptococcus pneumoniae</i> 670-6B	NC_014498.1
	<i>Streptococcus pyogenes</i> NZ131	NC_011375.1
	<i>Streptococcus sanguinis</i> SK36	NC_009009
	<i>Streptococcus suis</i> P1/7	NC_012925.1
	<i>Streptococcus thermophilus</i> LMD-9	NC_008532.1
<i>Listeria</i>	<i>Streptococcus uberis</i> 0140J	NC_012004.1
	<i>Listeria monocytogenes</i> str. ATCC 19117	NC_018584.1
<i>Burkholderia</i>	<i>Burkholderia cenocepacia</i> J2315	NC_011000.1
	<i>Burkholderia glumae</i> BGR1	NC_012721.2
	<i>Burkholderia rhizoxinica</i> HKI 454	NC_014722.1
	<i>Burkholderia</i> sp. CCGE1001	NC_015136.1
	<i>Burkholderia</i> sp. CCGE1002	NC_014117.1
<i>Klebsiella</i>	<i>Klebsiella pneumoniae</i> KCTC 2242	NC_017540
	<i>Klebsiella pneumoniae</i> subsp. <i>pneumoniae</i> 1084	NC_018522
<i>Clostridium</i>	<i>Clostridium beijerinckii</i> NCIMB 8052	NC_009617.1
	<i>Clostridium botulinum</i> Ba4 str. 657	NC_012658.1
	<i>Clostridium botulinum</i> E3 str. Alaska E43	NC_010723.1
	<i>Clostridium cellulolyticum</i> H10	NC_011898.1
	<i>Clostridium cellulovorans</i> 743B	NC_014393.1
	<i>Clostridium difficile</i> CD196	NC_013315.1

Table S7 Potential bacterial pathogens identified from the dust samples (*Continued*).

Genus	Species/Strains	NCBI reference sequence
<i>Clostridium</i>	<i>Clostridium kluyveri</i> NBRC 12016	NC_011837.1
	<i>Clostridium lentocellum</i> DSM 5427	NC_015275.1
	<i>Clostridium perfringens</i> ATCC 13124	NC_008261.1
	<i>Clostridium saccharolyticum</i> WM1	NC_014376.1
	<i>Clostridium sticklandii</i> str. DSM 519	NC_014614.1
	<i>Clostridium tetani</i> E88	NC_004557
	<i>Clostridium thermocellum</i> ATCC 27405	NC_009012.1
<i>Bordetella</i>	<i>Bordetella avium</i> 197N	NC_010645
	<i>Bordetella petrii</i> str. DSM 12804	NC_010170.1
<i>Corynebacterium</i>	<i>Corynebacterium aurimucosum</i> ATCC 700975	NC_012590.1
	<i>Corynebacterium diphtheriae</i> str. NCTC11397	X_84248.1
	<i>Corynebacterium glutamicum</i> R	NC_009342.1
	<i>Corynebacterium jeikeium</i> K411	NC_007164
	<i>Corynebacterium kroppenstedtii</i> DSM 44385	NC_012704.1
	<i>Corynebacterium pseudotuberculosis</i> FRC41	NC_014329.1
<i>Achromobacter</i>	<i>Achromobacter xylosoxidans</i> NH44784-1996	NC_021285
<i>Acinetobacter</i>	<i>Acinetobacter baumannii</i> SDF	NC_010400
	<i>Acinetobacter oleivorans</i> DR1	NC_014259.1
	<i>Acinetobacter</i> sp. ADP1	NC_005966.1
<i>Aeromonas</i>	<i>Aeromonas salmonicida</i> subsp. <i>salmonicida</i> A449	NC_009348.1
<i>Aggregatibacter</i>	<i>Aggregatibacter actinomycetemcomitans</i> D7S-1	NC_017846
<i>Bacteroides</i>	<i>Bacteroides fragilis</i> 638R	NC_016776
	<i>Bacteroides thetaiotaomicron</i> VPI-5482	NC_004663
<i>Bordetella</i>	<i>Bordetella avium</i> 197N	NC_010645
	<i>Bordetella petrii</i> str. DSM 12804	NC_010170.1
<i>Brucella</i>	<i>Brucella microti</i> CCM 4915	NC_013118.1
<i>Campylobacter</i>	<i>Campylobacter concisus</i> 13826	NC_009802.1
	<i>Campylobacter hominis</i> ATCC BAA-381	NC_009714.1
	<i>Campylobacter lari</i> RM2100	NC_012039.1
<i>Coxiella</i>	<i>Coxiella burnetii</i> CbuK Q154	NC_011528.1
<i>Enterococcus</i>	<i>Enterococcus faecium</i> NRRL B-2354	NC_020207
<i>Fusobacterium</i>	<i>Fusobacterium nucleatum</i> 3-1-36A2	NC_022196
	<i>Fusobacterium nucleatum</i> 4-8	NC_021281
	<i>Fusobacterium nucleatum</i> ATCC 25586	NC_003454
<i>Haemophilus</i>	<i>Haemophilus influenzae</i> PittEE	NC_009566
	<i>Haemophilus influenzae</i> F3047	NC_014922.1
<i>Mycobacterium</i>	<i>Mycobacterium abscessus</i>	NC_010397.1
	<i>Mycobacterium gilvum</i> Spyr1	NC_014814.1
	<i>Mycobacterium sinense</i> str. JDM601	NC_015576.1
	<i>Mycobacterium</i> sp. JLS	NC_009077.1
<i>Mycoplasmas</i>	<i>Mycoplasmas penetrans</i> HF-2	NC_004432
<i>Neisseria</i>	<i>Neisseria meningitidis</i> 8013	NC_017501.1
<i>Rickettsia</i>	<i>Rickettsia prowazekii</i> Rp22	NC_017560.1
	<i>Rickettsia typhi</i> str. Wilmington	NC_006142
<i>Salmonella</i>	<i>Salmonella enterica</i> str. 287/91	NC_011274.1
<i>Shigella</i>	<i>Shigella dysenteriae</i> Sd197	NC_007606
	<i>Shigella flexneri</i> 301	NC_004337
<i>Vibrio</i>	<i>Vibrio fischeri</i> MJ11	NC_011184.1
	<i>Vibrio harveyi</i> ATCC BAA-1116	NC_009784.1
	<i>Vibrio</i> sp. Ex25	NC_013456.1
	<i>Vibrio splendidus</i> LGP32	NC_011753.2
	<i>Vibrio vulnificus</i> CMCP6	NC_004459.3
<i>Yersinia</i>	<i>Yersinia pestis</i> Angola	NC_010159

Table S8 Pearson correlations between the normalized abundance of ARGs and MGEs.

	Total ARGs	MGEs	Amino-glycoside	Beta-Lactamase	FCA	MLSB	Multidrug	Sulfonamide	Tetracycline	Vancomycin	Others	Transposase	Integrase
Total ARGs	1												
MGEs	0.608**	1											
Aminoglycoside	-0.155	0.260*	1										
Beta-Lactamase	0.577**	0.240*	0.003	1									
FCA	-0.020	-0.014	0.179	-0.015	1								
MLSB	0.983**	0.575**	-0.139	0.505**	-0.082	1							
Multidrug	0.104	0.234*	-0.051	0.001	0.743**	0.096	1						
Sulfonamide	0.942**	0.709**	-0.144	0.631**	-0.082	0.902**	-0.072	1					
Tetracycline	0.724**	0.625**	-0.081	0.078	-0.062	0.725**	-0.078	0.677**	1				
Vancomycin	0.993**	0.594**	-0.182	0.602**	-0.097	0.978**	0.083	0.945**	0.692**	1			
Others	-0.099	0.109	0.532**	-0.001	0.289**	-0.090	-0.021	-0.091	-0.056	-0.115	1		
Transposase	-0.185	0.318**	0.780**	-0.057	0.120	0.276*	0.137	-0.180	-0.105	0.262*	0.414**	1	
Integrase	0.718**	0.915**	-0.058	0.276*	-0.066	0.679**	0.271*	0.821**	0.701**	0.710**	-0.061	-0.091	1

FCA: (flor)/(chlor)/(am)phenicol resistance genes; MLSB: Macrolide-Lincosamide-Streptogramin B resistance genes; Others: ARGs conferring resistance to other antibiotic classes like bacitracin, fosfomycin, nitroimidazole, pyrazinamide and streptothricin.

*, $P < 0.05$; **, $P < 0.01$.

Table S9 Co-occurrence patterns between specific ARGs and bacterial taxa.

Phylum	Genus	ARG	ARG type	Spearman's ρ	Adjusted P-value
<i>Acidobacteria</i>	<i>Edaphobacter</i>	cphA-02	Beta-Lactamase	0.99	1.46E-06
	<i>Paludibaculum</i>	tetB-02	Tetracycline	0.78	0.0027
<i>Actinobacteria</i>	<i>Actinomyces</i>	ermC	MLSB	0.85	0.0005
		tet(34)	Tetracycline	0.80	0.0017
	<i>Actinotignum</i>	fox5	Beta-Lactamase	0.95	5.08E-06
	<i>Alloscardovia</i>	oleC	MLSB	0.86	0.0003
	<i>Brachybacterium</i>	cphA-02	Beta-Lactamase	0.77	0.0036
	<i>Corynebacterium</i>	cmlA1-01	FCA	0.71	0.0091
		pmrA	Multidrug	0.71	0.0091
	<i>Corynebacterium_1</i>	cmlA1-01	FCA	0.73	0.0073
		pmrA	Multidrug	0.73	0.0073
	<i>Dactylosporangium</i>	dfrA1	Sulfonamide	0.85	0.0004
	<i>Dermabacter</i>	aphA1(aka kanR)	Aminoglycoside	0.73	0.0068
		cmlA1-01	FCA	0.87	0.0002
		cmlA1-02	FCA	0.81	0.0014
		ermA	MLSB	0.80	0.0018
		ermA_ermTR	MLSB	0.75	0.0049
		pmrA	Multidrug	0.87	0.0002
		tetG-02	Tetracycline	0.99	1.22E-06
		aadA5-01	Aminoglycoside	0.73	0.0075
		aadA5-02	Aminoglycoside	0.73	0.0071
		lnuB-01	MLSB	0.85	0.0005
	<i>Glutamicibacter</i>	lnuB-02	MLSB	0.72	0.0005
		cmr	Multidrug	0.85	0.0084
		tetH	Tetracycline	0.85	0.0005
		tetL-02	Tetracycline	0.71	0.0091
		tetR-02	Tetracycline	0.74	0.0061
		tetS	Tetracycline	0.85	0.0005
		mecA	Beta-Lactamase	0.88	0.0002
		tetK	Tetracycline	0.84	0.0007
		blaTEM	Beta-Lactamase	0.87	0.0002
		vatE-02	MLSB	0.87	0.0002
	<i>Planomonospora</i>	blaCTX-M-05	Beta-Lactamase	0.80	0.0018
	<i>Propionibacterium</i>	aphA1(aka kanR)	Aminoglycoside	0.75	0.0048
		blaZ	Beta-Lactamase	0.72	0.0078
		cmlA1-01	FCA	0.87	0.0003
		cmlA1-02	FCA	0.81	0.0014
		ermA	MLSB	0.79	0.0019
		ermA_ermTR	MLSB	0.87	0.0002
		pmrA	Multidrug	0.87	0.0003
		sat4	Others	0.81	0.0014
		tetM-01	Tetracycline	0.75	0.0053
		fox5	Beta-Lactamase	0.80	0.0017
	<i>Rhodococcus</i>	vanSB	Vancomycin	0.86	0.0003
		vanYD-01	Vancomycin	0.79	0.0024
	<i>Rothia</i>	aadA1	Aminoglycoside	0.71	0.0097
		erm(36)	MLSB	0.86	0.0003
	<i>Streptosporangium</i>	blaCTX-M-05	Beta-Lactamase	0.73	0.0073
	<i>Tsukamurella</i>	cphA-02	Beta-Lactamase	0.85	0.0005
<i>Armatimonadetes</i>	<i>Chthonomonas</i>	oleC	MLSB	0.73	0.0074
<i>Bacteroidetes</i>	<i>Arcticibacter</i>	ampC-02	Beta-Lactamase	0.90	4.51E-05
		ceoA	Multidrug	0.88	0.0001
	<i>Bergeyella</i>	fox5	Beta-Lactamase	0.78	0.0030

Table S9 Co-occurrence patterns between specific ARGs and bacterial taxa (*Continued*).

Phylum	Genus	ARG	ARG type	Spearman's ρ	Adjusted <i>P</i> -value
<i>Bacteroidetes</i>	<i>Capnocytophaga</i>	ermC	MLSB	0.90	5.39E-05
		tet(34)	Tetracycline	0.87	0.0003
	<i>Chitinophaga</i>	fox5	Beta-Lactamase	0.84	0.0007
	<i>Cnuella</i>	yidY/mdtL-02	Multidrug	0.74	0.0056
	<i>Crenotalea</i>	lnuA-01	MLSB	0.77	0.0032
	<i>Filimonas</i>	fox5	Beta-Lactamase	0.94	5.68E-06
	<i>Macellibacteroides</i>	ermC	MLSB	0.73	0.0067
		tet(34)	Tetracycline	0.77	0.0032
	<i>Moheibacter</i>	oleC	MLSB	0.90	5.80E-05
	<i>Myroides</i>	oleC	MLSB	0.90	4.35E-05
	<i>Nibribacter</i>	ampC-02	Beta-Lactamase	0.95	4.92E-06
		ceoA	Multidrug	0.93	8.22E-06
	<i>Olivibacter</i>	oleC	MLSB	0.92	2.21E-05
	<i>Parapedobacter</i>	oleC	MLSB	0.91	3.21E-05
	<i>Pedobacter</i>	ampC-02	Beta-Lactamase	0.95	4.31E-06
		ceoA	Multidrug	0.94	5.70E-06
	<i>Prevotellaceae_UCG-001</i>	cphA-02	Beta-Lactamase	0.89	0.0001
	<i>Pricia</i>	lnuA-01	MLSB	0.71	0.0093
	<i>Pseudopedobacter</i>	ampC-02	Beta-Lactamase	1.00	0
		ceoA	Multidrug	0.99	1.62E-06
	<i>Ruminofilibacter</i>	aadA9-02	Aminoglycoside	0.87	0.0002
		matA_mel	MLSB	0.93	1.66E-05
	<i>Runella</i>	fox5	Beta-Lactamase	0.91	2.99E-05
	<i>Salinimicrobium</i>	cphA-02	Beta-Lactamase	0.82	0.0011
	<i>Spirosoma</i>	emrD	Multidrug	0.81	0.0013
		tolC-03	Multidrug	0.98	1.69E-06
	<i>Tunicatimonas</i>	blaCTX-M-05	Beta-Lactamase	0.79	0.0025
	<i>Ulvibacter</i>	cphA-02	Beta-Lactamase	0.98	1.75E-06
	<i>Vitellibacter</i>	cfr	FCA	0.99	1.27E-06
		tnpA-02	MGEs	0.82	0.0012
		mexE	Multidrug	0.99	1.18E-06
		yceL/mdtH-02	Multidrug	1.00	1.09E-06
		bacA-01	Others	1.00	1.09E-06
<i>Chlamydiae</i>	<i>Candidatus_Proteochlamydia</i>	lnuA-01	MLSB	0.77	0.0033
<i>Chloroflexi</i>	<i>Nitrolancea</i>	lnuA-01	MLSB	0.75	0.0049
<i>Deinococcus-Thermus</i>	<i>Thermus</i>	vanSB	Vancomycin	0.73	0.0072
		vanYD-01	Vancomycin	0.76	0.004
	<i>[Bacteroides]_pectinophilus_group</i>	mecA	Beta-Lactamase	0.77	0.0032
		tetK	Tetracycline	0.74	0.0064
	<i>[Clostridium]_innocuum_group</i>	cphA-02	Beta-Lactamase	0.92	2.22E-05
	<i>[Eubacterium]_hallii_group</i>	cphA-02	Beta-Lactamase	0.80	0.0017
	<i>Aerococcus</i>	aac(6')-Ib(aka aacA4)-02	Aminoglycoside	0.76	0.0039
		catB3	FCA	0.82	0.0009
		tnpA-03	MGEs	0.79	0.0020
		ermB	MLSB	0.97	3.49E-06
		dfrA1	Sulfonamide	0.77	0.0037
	<i>Ammoniibacillus</i>	acrA-02	Multidrug	0.89	6.32E-05
		qacH-01	Multidrug	0.95	4.88E-06
		yidY/mdtL-01	Multidrug	0.97	2.97E-06
		tetA-02	Tetracycline	0.97	2.97E-06
	<i>Atopostipes</i>	cphA-02	Beta-Lactamase	0.98	2.11E-06

Table S9 Co-occurrence patterns between specific ARGs and bacterial taxa (*Continued*).

Phylum	Genus	ARG	ARG type	Spearman's ρ	Adjusted P-value
<i>Firmicutes</i>	<i>Brochothrix</i>	cfr	FCA	0.97	3.56E-06
		tnpA-02	MGEs	0.79	0.0022
		mexE	Multidrug	0.97	3.52E-06
		yceL/mdtH-02	Multidrug	0.97	2.91E-06
		bacA-01	Others	0.97	2.91E-06
	<i>Butyricicoccus</i>	cphA-02	Beta-Lactamase	0.92	1.73E-05
	<i>Candidatus_Arthromitus</i>	aac(6')-II	Aminoglycoside	0.73	0.0072
	<i>Candidatus_Soleaferrea</i>	qacH-01	Multidrug	0.71	0.0094
		yidY/mdtL-01	Multidrug	0.72	0.0087
		tetA-02	Tetracycline	0.72	0.0087
	<i>Clostridium_sensu_stricto_13</i>	cfr	FCA	0.84	0.0005
		mexE	Multidrug	0.85	0.0005
		yceL/mdtH-02	Multidrug	0.85	0.0004
		bacA-01	Others	0.85	0.0002
	<i>Clostridium_sensu_stricto_7</i>	cfr	FCA	0.86	0.0003
		mexE	Multidrug	0.87	0.0003
		yceL/mdtH-02	Multidrug	0.87	0.0002
		bacA-01	Others	0.87	0.0002
	<i>Dolosicoccus</i>	cphA-02	Beta-Lactamase	0.88	0.0002
	<i>Dolosigranulum</i>	aph(2')-Id-01	Aminoglycoside	0.94	5.73E-06
		cfxA	Beta-Lactamase	0.78	0.0028
		pbp2x	Beta-Lactamase	0.94	5.73E-06
		Tp614	MGEs	0.92	2.14E-05
		fabK	Others	0.94	5.73E-06
		tetO-01	Tetracycline	0.71	0.0093
		tetPB-03	Tetracycline	0.84	0.0007
		aadA5-01	Aminoglycoside	0.81	0.0012
		aadA5-02	Aminoglycoside	0.78	0.0027
		aadE	Aminoglycoside	0.75	0.0053
	<i>Eremococcus</i>	lnuB-01	MLSB	0.93	8.84E-06
		lnuB-02	MLSB	0.82	0.0011
		cmr	Multidrug	0.93	8.84E-06
		tetH	Tetracycline	0.93	8.84E-06
		tetR-02	Tetracycline	0.79	0.0025
		tetS	Tetracycline	0.93	8.84E-06
		strA	Aminoglycoside	0.95	4.69E-06
		ampC-01	Beta-Lactamase	0.95	4.69E-06
		ampC-04	Beta-Lactamase	0.95	4.69E-06
		ampC-05	Beta-Lactamase	0.95	4.69E-06
	<i>Exiguobacterium</i>	ampC-06	Beta-Lactamase	0.95	4.69E-06
		blaCMY2-01	Beta-Lactamase	0.95	4.69E-06
		catA1	FCA	0.95	4.69E-06
		acrR-01	Multidrug	0.95	4.69E-06
		ttgA	Multidrug	0.95	5.09E-06
		ttgB	Multidrug	0.86	0.0003
		dfrA1	Sulfonamide	0.82	0.0012
		tet(32)	Tetracycline	0.70	0.0099
		aphA1(aka kanR)	Aminoglycoside	0.74	0.0059
		cmlA1-01	FCA	0.91	2.74E-05
	<i>Faecalitalea</i>	cmlA1-02	FCA	0.85	0.0005
		ermA	MLSB	0.83	0.0007
		ermA_ermTR	MLSB	0.75	0.0048
		pmrA	Multidrug	0.91	2.74E-05
	<i>Family_XIII_AD3011_group</i>				
	<i>Finegoldia</i>				

Table S9 Co-occurrence patterns between specific ARGs and bacterial taxa (*Continued*).

Phylum	Genus	ARG	ARG type	Spearman's ρ	Adjusted P-value
<i>Firmicutes</i>	<i>Gottschalkia</i>	aadA5-01	Aminoglycoside	0.70	0.0098
		aadA5-02	Aminoglycoside	0.73	0.0075
		lnuB-01	MLSB	0.87	0.0003
		lnuB-02	MLSB	0.74	0.0056
		cmr	Multidrug	0.87	0.0003
		tetH	Tetracycline	0.87	0.0003
		tetR-02	Tetracycline	0.71	0.0097
		tetS	Tetracycline	0.87	0.0003
	<i>Halobacillus</i>	acrA-02	Multidrug	0.92	2.12E-05
		qacH-01	Multidrug	0.98	1.98E-06
		yidY/mdtL-01	Multidrug	1.00	1.06E-06
		tetA-02	Tetracycline	1.00	1.06E-06
	<i>Halocella</i>	acrA-02	Multidrug	0.87	0.0002
		qacH-01	Multidrug	0.93	1.51E-05
		yidY/mdtL-01	Multidrug	0.95	5.03E-06
		tetA-02	Tetracycline	0.95	5.03E-06
	<i>Helcococcus</i>	ermC	MLSB	0.75	0.0046
		tet(34)	Tetracycline	0.79	0.0023
	<i>Holdemanella</i>	ermC	MLSB	0.81	0.0015
		tet(34)	Tetracycline	0.84	0.0006
	<i>Jeotgalicoccus</i>	cphA-02	Beta-Lactamase	0.86	0.0004
	<i>Lachnoclostridium</i>	blaCTX-M-05	Beta-Lactamase	0.81	0.0016
	<i>Lachnospira</i>	cphA-02	Beta-Lactamase	0.95	4.20E-06
	<i>Lachnospiraceae_ND3007_group</i>	cphA-02	Beta-Lactamase	0.94	6.26E-06
	<i>Lachnospiraceae_NK4A136_group</i>	mecA	Beta-Lactamase	0.90	4.11E-05
		tetK	Tetracycline	0.86	0.0003
	<i>Lactobacillus</i>	fox5	Beta-Lactamase	0.79	0.0021
	<i>Macrococcus</i>	lnuA-01	MLSB	0.76	0.0039
		cfr	FCA	0.94	6.21E-06
		tnpA-02	MGEs	0.77	0.0037
		mexE	Multidrug	0.94	5.91E-06
		yceL/mdtH-02	Multidrug	0.94	5.86E-06
		bacA-01	Others	0.94	5.86E-06
		ermC	MLSB	0.96	3.72E-06
		tet(34)	Tetracycline	0.98	1.80E-06
		ermC	MLSB	0.84	0.0005
		tet(34)	Tetracycline	0.83	0.0009
	<i>Phascolarctobacterium</i>	vanXD	Vancomycin	0.74	0.0064
		vanYD-01	Vancomycin	0.79	0.0019
	<i>Piscibacillus</i>	cphA-02	Beta-Lactamase	0.99	1.12E-06
	<i>Pseudobacteroides</i>	acrA-02	Multidrug	0.75	0.0048
		qacH-01	Multidrug	0.80	0.0016
		yidY/mdtL-01	Multidrug	0.83	0.0009
		tetA-02	Tetracycline	0.83	0.0009
		cphA-02	Beta-Lactamase	0.78	0.0026
	<i>Ruminococcaceae_UCG-014</i>	aadA-02	Aminoglycoside	0.80	0.0019
	<i>Saccharibacillus</i>	acrA-02	Multidrug	0.93	1.69E-05
	<i>Selenomonas</i>	qacH-01	Multidrug	0.98	1.72E-06
		yidY/mdtL-01	Multidrug	1.00	1.06E-06
		tetA-02	Tetracycline	1.00	1.06E-06
		ermC	MLSB	0.87	0.0002
	<i>Selenomonas_3</i>	tet(34)	Tetracycline	0.92	2.27E-05
	<i>Tissierella</i>	cphA-02	Beta-Lactamase	0.92	2.13E-05

Table S9 Co-occurrence patterns between specific ARGs and bacterial taxa (*Continued*).

Phylum	Genus	ARG	ARG type	Spearman's ρ	Adjusted <i>P</i> -value
<i>Firmicutes</i>	<i>Trichococcus</i>	blaTEM	Beta-Lactamase	0.90	6.11E-05
		vatE-02	MLSB	0.90	6.11E-05
	<i>Tyzzarella_3</i>	blaCTX-M-05	Beta-Lactamase	0.80	0.0017
	<i>Tyzzarella_4</i>	cphA-02	Beta-Lactamase	0.88	0.0001
	<i>Ureibacillus</i>	qacH-01	Multidrug	0.71	0.0097
		yidY/mdtL-01	Multidrug	0.73	0.0072
		tetA-02	Tetracycline	0.73	0.0072
<i>Fusobacteria</i>	<i>Cetobacterium</i>	lnuA-01	MLSB	0.77	0.0032
	<i>Sneathia</i>	ampC-01	Beta-Lactamase	0.95	5.43E-06
		ampC-04	Beta-Lactamase	0.95	5.43E-06
		ampC-05	Beta-Lactamase	0.95	5.43E-06
		ampC-06	Beta-Lactamase	0.95	5.43E-06
		blaCMY2-01	Beta-Lactamase	0.95	5.43E-06
		catA1	FCA	0.95	5.43E-06
		acrR-01	Multidrug	0.95	5.43E-06
		ttgA	Multidrug	0.93	8.29E-06
		ttgB	Multidrug	0.85	0.0005
<i>Gemmatimonadetes</i>	<i>Longimicrobium</i>	ermK-01	MLSB	0.98	1.67E-06
<i>Planctomycetes</i>	<i>AKYG587</i>	mecA	Beta-Lactamase	0.76	1.92E-06
		tetK	Tetracycline	0.72	5.74E-06
	<i>Algisphaera</i>	fox5	Beta-Lactamase	0.86	0.0004
	<i>Isosphaera</i>	yidY/mdtL-02	Multidrug	0.92	1.91E-05
	<i>SM1A02</i>	fox5	Beta-Lactamase	0.86	0.0004
	<i>Zavarzinella</i>	dfrA1	Sulfonamide	0.74	0.0056
<i>Proteobacteria</i>	<i>Achromobacter</i>	vanSB	Vancomycin	0.85	0.0005
		vanYD-01	Vancomycin	0.70	0.0098
	<i>Acinetobacter</i>	strA	Aminoglycoside	0.81	0.0015
		ampC-01	Beta-Lactamase	0.81	0.0015
		ampC-04	Beta-Lactamase	0.81	0.0015
		ampC-05	Beta-Lactamase	0.81	0.0015
		ampC-06	Beta-Lactamase	0.81	0.0015
		blaCMY2-01	Beta-Lactamase	0.81	0.0015
		catA1	FCA	0.81	0.0015
		acrR-01	Multidrug	0.81	0.0015
		ttgA	Multidrug	0.84	0.0007
		ttgB	Multidrug	0.75	0.0051
		lnuA-01	MLSB	0.77	0.0032
		ermK-01	MLSB	0.88	0.0001
		lnuA-01	MLSB	0.78	0.0028
	<i>Alcanivorax</i>	lnuA-01	MLSB	0.73	0.0072
	<i>Aliidiomarina</i>	lnuA-01	MLSB	0.77	0.0033
	<i>Aquamicrobium</i>	lnuA-01	MLSB	0.77	0.0032
	<i>Aquicella</i>	lnuA-01	MLSB	0.77	0.0032
	<i>Azospirillum</i>	ermK-01	MLSB	0.72	0.0083
	<i>Bartonella</i>	mecA	Beta-Lactamase	0.74	0.0056
		tetK	Tetracycline	0.72	0.0087
	<i>BD1-7_clade</i>	cphA-02	Beta-Lactamase	0.99	1.12E-06
	<i>Bosea</i>	mepA	Multidrug	0.72	0.0081
	<i>Buchnera</i>	ermC	MLSB	0.97	3.62E-06
		tet(34)	Tetracycline	1.00	1.06E-06
	<i>Burkholderia</i>	vanSB	Vancomycin	0.71	0.0090
	<i>Candidatus_Alysiosphaera</i>	mecA	Beta-Lactamase	0.76	0.0039
		tetK	Tetracycline	0.72	0.0085

Table S9 Co-occurrence patterns between specific ARGs and bacterial taxa (*Continued*).

Phylum	Genus	ARG	ARG type	Spearman's ρ	Adjusted <i>P</i> -value
<i>Proteobacteria</i>	<i>Candidatus_Odyssella</i>	mepA	Multidrug	0.82	0.0011
	<i>Chitinimonas</i>	fox5	Beta-Lactamase	0.91	2.34E-05
	<i>Comamonas</i>	ermC	MLSB	0.87	0.0002
		tet(34)	Tetracycline	0.90	5.25E-05
	<i>Deefgea</i>	lnuA-01	MLSB	0.77	0.0032
	<i>Delftia</i>	cfr	FCA	0.77	0.0035
		mexE	Multidrug	0.77	0.0035
		yceL/mdtH-02	Multidrug	0.78	0.0030
		bacA-01	Others	0.78	0.0030
	<i>Devosia</i>	ermK-01	MLSB	0.79	0.0025
	<i>Dokdonella</i>	fox5	Beta-Lactamase	0.95	4.99E-06
	<i>Dongia</i>	lnuA-01	MLSB	0.72	0.0082
	<i>Endosymbionts8</i>	aacA_apHD	Aminoglycoside	0.78	0.0026
		aacC1	Aminoglycoside	0.98	1.71E-06
		tnpA-01	MGEs	0.87	0.0002
		tnpA-07	MGEs	0.98	2.28E-06
		msrA-01	MLSB	0.74	0.0058
		sulA/folP-01	Sulfonamide	0.76	0.0043
	<i>Enhydrobacter</i>	erm(36)	MLSB	0.80	0.0016
	<i>Shigella</i>	vanSB	Vancomycin	0.77	0.0032
	<i>Geoalkalibacter</i>	mecA	Beta-Lactamase	0.94	5.63E-06
		tetK	Tetracycline	0.90	4.30E-05
	<i>Gluconobacter</i>	ermC	MLSB	0.97	3.62E-06
		tet(34)	Tetracycline	1.00	1.06E-06
	<i>Haemophilus</i>	aadA-02	Aminoglycoside	0.72	0.0077
	<i>Hydrogenophaga</i>	cfr	FCA	0.72	0.0078
		mexE	Multidrug	0.72	0.008
		yceL/mdtH-02	Multidrug	0.72	0.0076
		bacA-01	Others	0.72	0.0076
	<i>Inhella</i>	fox5	Beta-Lactamase	0.95	5.29E-06
	<i>Jannaschia</i>	blaCTX-M-05	Beta-Lactamase	0.73	0.0066
	<i>Kushneria</i>	intI3	MGEs	0.90	4.24E-05
	<i>Limnobacter</i>	cphA-01	Beta-Lactamase	0.72	0.0079
	<i>Methyloparacoccus</i>	lnuA-01	MLSB	0.79	0.0024
	<i>Mizugakiibacter</i>	lnuA-01	MLSB	0.78	0.0028
	<i>MWH</i>	lnuA-01	MLSB	0.79	0.0024
	<i>Novispirillum</i>	ermC	MLSB	0.89	7.42E-05
		tet(34)	Tetracycline	0.93	1.40E-05
	<i>Panacagrimonas</i>	aadA-02	Aminoglycoside	0.74	0.0061
	<i>Parasutterella</i>	cphA-02	Beta-Lactamase	0.77	0.0031
	<i>Pelagibacterium</i>	oleC	MLSB	0.78	0.0027
	<i>Perlucidibaca</i>	vanSB	Vancomycin	0.74	0.0057
	<i>Pigmentiphaga</i>	tolC-03	Multidrug	0.71	0.0095
		pncA	Others	0.76	0.0041
	<i>Polynucleobacter</i>	strA	Aminoglycoside	0.97	3.37E-06
		ampC-01	Beta-Lactamase	0.97	3.37E-06
		ampC-04	Beta-Lactamase	0.97	3.37E-06
		ampC-05	Beta-Lactamase	0.97	3.37E-06
		ampC-06	Beta-Lactamase	0.97	3.37E-06
		blaCMY2-01	Beta-Lactamase	0.97	3.37E-06

Table S9 Co-occurrence patterns between specific ARGs and bacterial taxa (*Continued*).

Phylum	Genus	ARG	ARG type	Spearman's ρ	Adjusted <i>P</i> -value
<i>Proteobacteria</i>	<i>Polynucleobacter</i>	catA1	FCA	0.97	3.37E-06
		acrR-01	Multidrug	0.97	3.37E-06
		ttgA	Multidrug	0.96	4.11E-06
		ttgB	Multidrug	0.87	0.0002
	<i>Pseudochelatococcus</i>	cphA-02	Beta-Lactamase	0.99	1.49E-06
	<i>Pseudoxanthomonas</i>	mepA	Multidrug	0.92	2.02E-05
	<i>Pusillimonas</i>	cphA-02	Beta-Lactamase	0.83	0.0009
	<i>Rheinheimera</i>	cfr	FCA	0.98	1.73E-06
		tnpA-02	MGEs	0.81	0.0015
		mexE	Multidrug	0.98	1.69E-06
		yceL/mdtH-02	Multidrug	0.99	0.0015
		bacA-01	Others	0.99	1.59E-06
	<i>Rhizobium</i>	mepA	Multidrug	0.91	3.92E-05
	<i>Rhodobacter</i>	lnuA-01	MLSB	0.76	0.0039
	<i>Rickettsia</i>	strA	Aminoglycoside	0.91	2.37E-05
		ampC-01	Beta-Lactamase	0.91	2.37E-05
		ampC-04	Beta-Lactamase	0.91	2.37E-05
		ampC-05	Beta-Lactamase	0.91	2.37E-05
		ampC-06	Beta-Lactamase	0.91	2.37E-05
		blaCMY2-01	Beta-Lactamase	0.91	2.37E-05
		catA1	FCA	0.91	2.37E-05
		acrR-01	Multidrug	0.91	2.37E-05
		ttgA	Multidrug	0.90	4.63E-05
		ttgB	Multidrug	0.82	0.0011
		aph(2')-Id-01	Aminoglycoside	0.71	0.0095
		pbp2x	Beta-Lactamase	0.71	0.0095
		fabK	Others	0.71	0.0095
	<i>Roseomonas</i>	cphA-02	Beta-Lactamase	0.96	3.68E-06
	<i>Roseovarius</i>	lnuA-01	MLSB	0.73	0.0065
	<i>Sandaracinobacter</i>	tolC-03	Multidrug	0.81	0.0015
	<i>Skermanella</i>	aph(2')-Id-01	Aminoglycoside	0.71	0.0091
	<i>Sphingomonas</i>	pbp2x	Beta-Lactamase	0.71	0.0091
		fabK	Others	0.71	0.0091
		mepA	Multidrug	0.79	0.0023
	<i>Stenotrophomonas</i>	lnuA-01	MLSB	0.71	0.0093
	<i>Sulfitobacter</i>	lnuA-01	MLSB	0.78	0.0026
	<i>Sva0081_sediment_group</i>	fox5	Beta-Lactamase	0.88	0.0002
	<i>Tahibacter</i>	cphA-01	Beta-Lactamase	0.90	4.29E-05
	<i>Telmatospirillum</i>	mepA	Multidrug	0.82	0.0009
	<i>Thauera</i>	lnuA-01	MLSB	0.79	0.0025
	<i>Vogesella</i>	strA	Aminoglycoside	0.89	8.96E-05
	<i>Wolbachia</i>	ampC-01	Beta-Lactamase	0.89	8.96E-05
		ampC-04	Beta-Lactamase	0.89	8.96E-05
		ampC-05	Beta-Lactamase	0.89	8.96E-05
		ampC-06	Beta-Lactamase	0.89	8.96E-05
		blaCMY2-01	Beta-Lactamase	0.89	8.96E-05
		catA1	FCA	0.89	8.96E-05
		acrR-01	Multidrug	0.89	8.96E-05
		ttgA	Multidrug	0.92	1.96E-05
		ttgB	Multidrug	0.84	0.0005
		ermK-01	MLSB	0.95	5.44E-06
	<i>Zymomonas</i>				
<i>Spirochaetae</i>	<i>Brevinema</i>	lnuA-01	MLSB	0.70	0.0099
	<i>Treponema_2</i>	vanSB	Vancomycin	0.71	0.0096

<i>Tenericutes</i>	<i>Mycoplasma</i>	lnuA-01	MLSB	0.76	0.0045
--------------------	-------------------	---------	------	------	--------

MLSB: Macrolide-Lincosamide-Streptogramin B resistance genes; FCA: (flor)/(chlor)/(am)phenicol resistance genes; Others: ARGs conferring resistance to other antibiotic classes like bacitracin, fosfomycin, nitroimidazole, pyrazinamide and streptothricin.

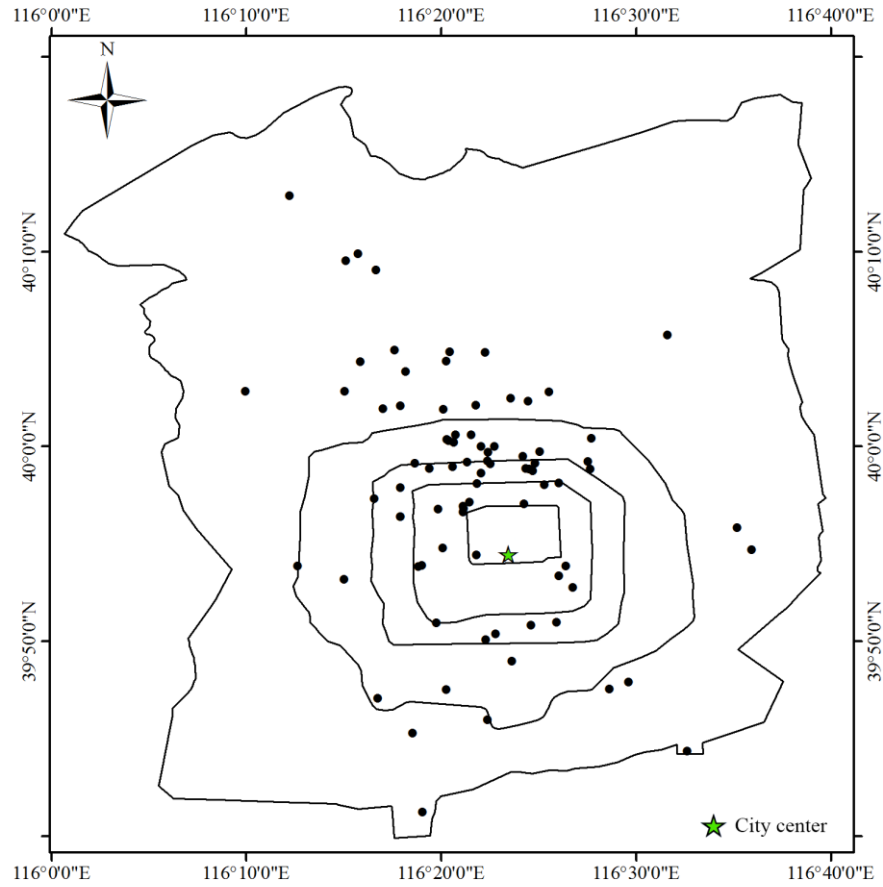


Fig. S1. ArcGIS map of sampling house locations in Beijing city. The boundaries around the city center denote the second, third, fourth, fifth and sixth ring road in sequence.

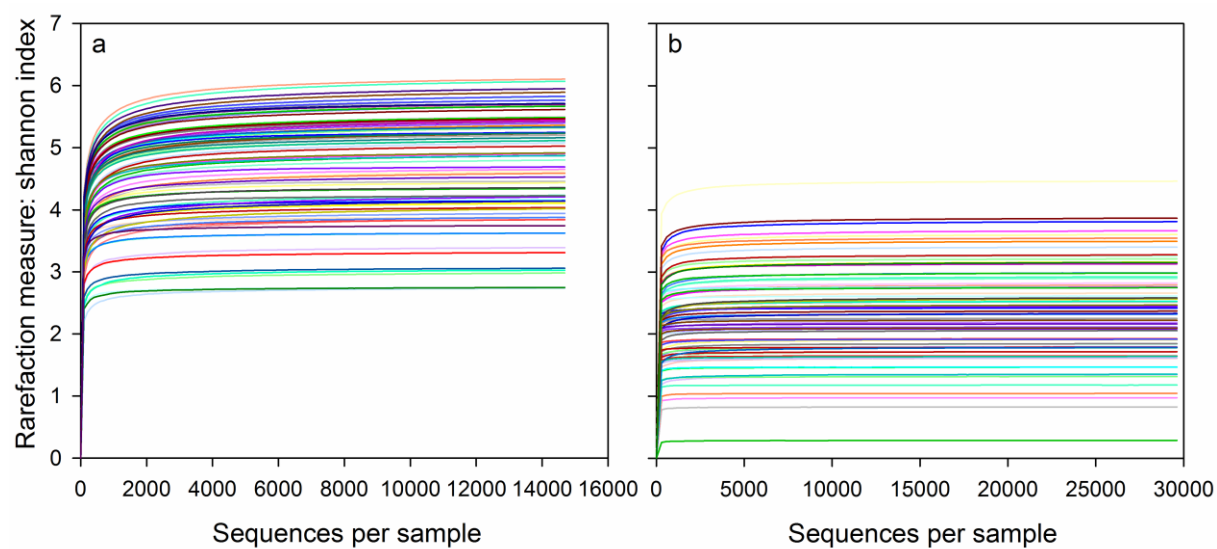


Fig. S2. Rarefaction curve of Shannon index for bacterial (a) and fungal communities (b) across dust samples. The index was calculated using 0.03 distance OTUs with sequences normalized to 14,685 and 29,615 per sample for bacteria and fungi, respectively.

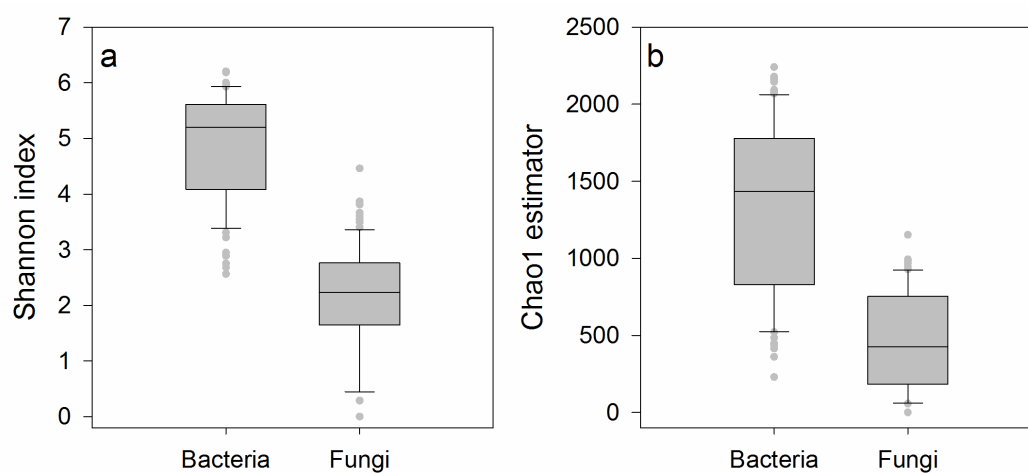


Fig. S3. Shannon index (a) and Chao1 estimator (b) of bacterial and fungal communities across dust samples.

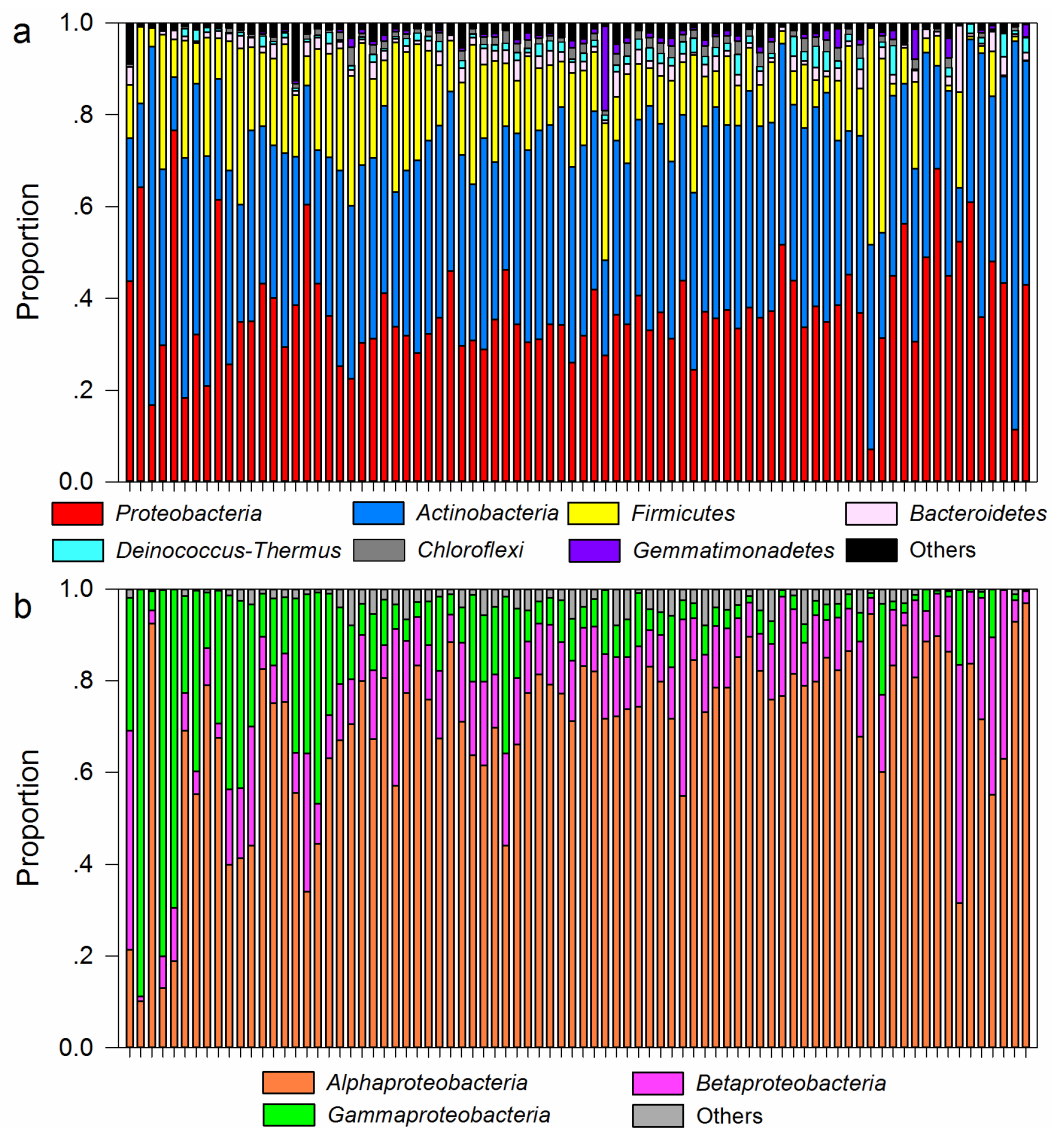


Fig. S4. Bacterial community composition at the phylum level (a) and within the *Proteobacteria* phylum (b). The proportional abundances of bacterial taxa in each dust sample were shown. Samples were ordered by their scores along the bacterial non-metric multidimensional scaling (NMDS) ordination axis. “Others” refers to the unclassified taxa and other taxa with proportions below 0.01 in more than 90% of the samples.

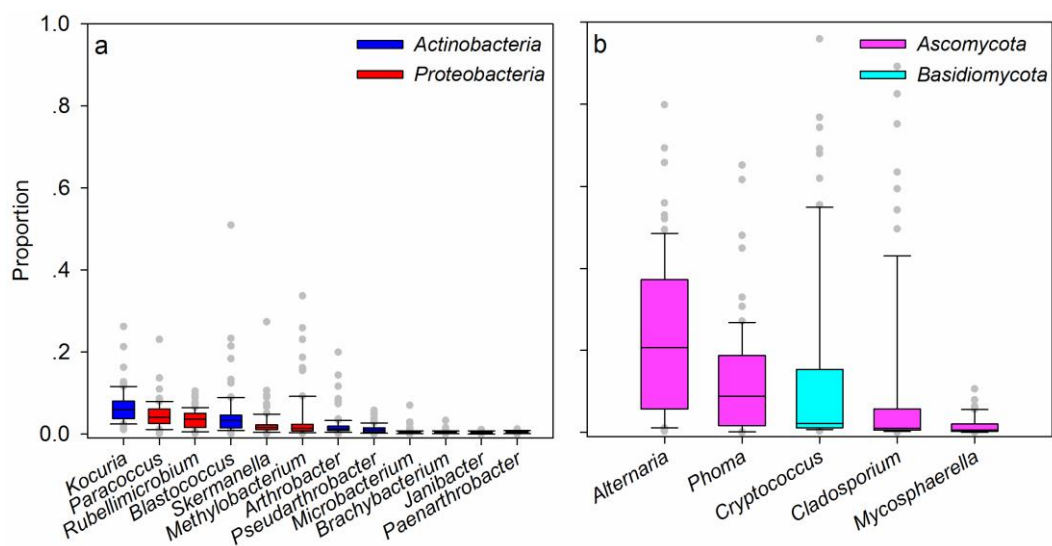


Fig. S5. Proportions of the shared and classified bacterial (a) and fungal genera (b) across dust samples. Genera were colored by their respective phylum.

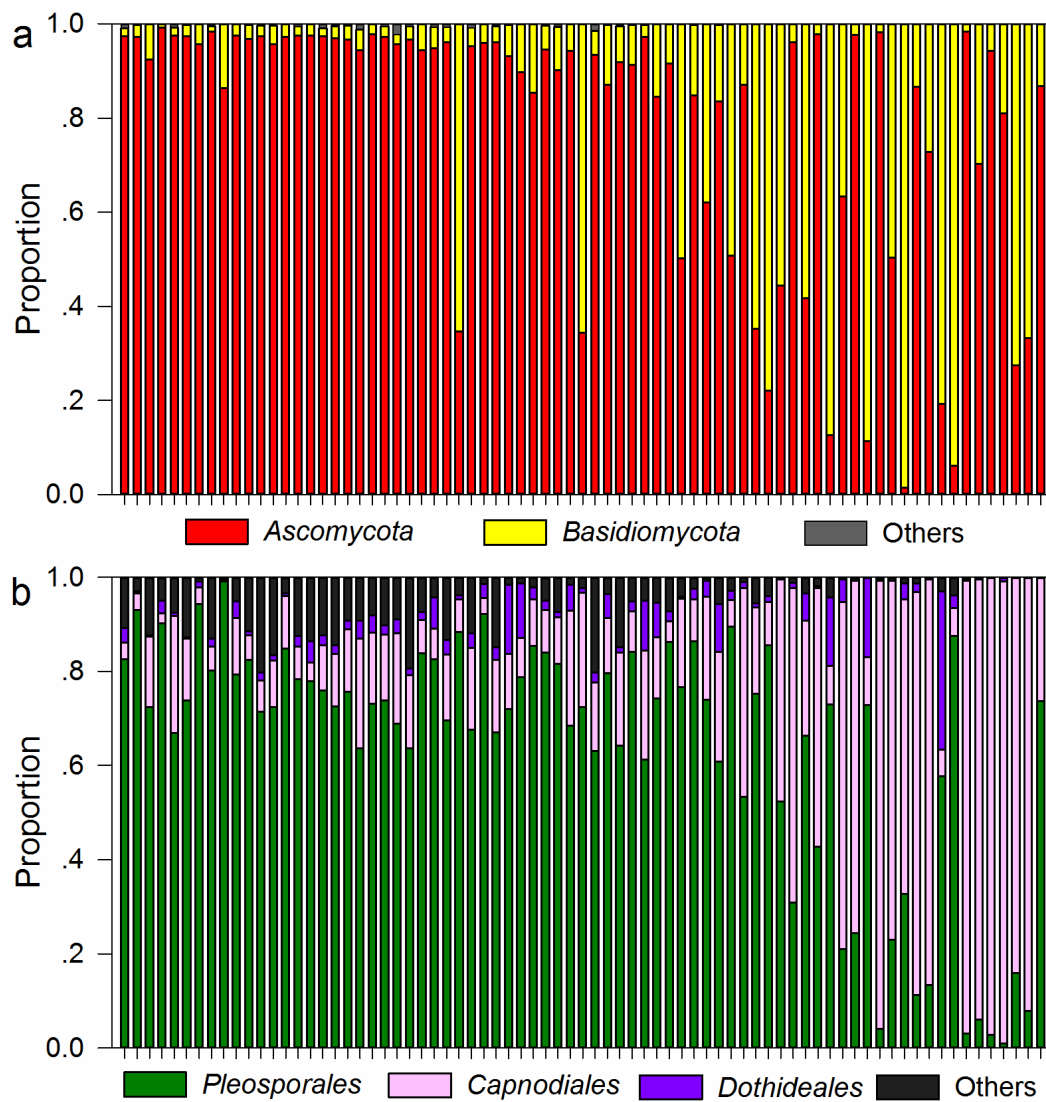


Fig. S6. Fungal community composition at the phylum level (a) and within the *Dothideomycetes* class (b). The *Dothideomycetes* class was affiliated with *Ascomycota* phylum. The proportional abundances of fungal taxa in each sample were shown. Samples were ordered by their scores along the fungal non-metric multidimensional scaling (NMDS) ordination axis. “Others” refers to the unclassified taxa and other taxa with proportions below 0.01 in more than 90% of the samples.

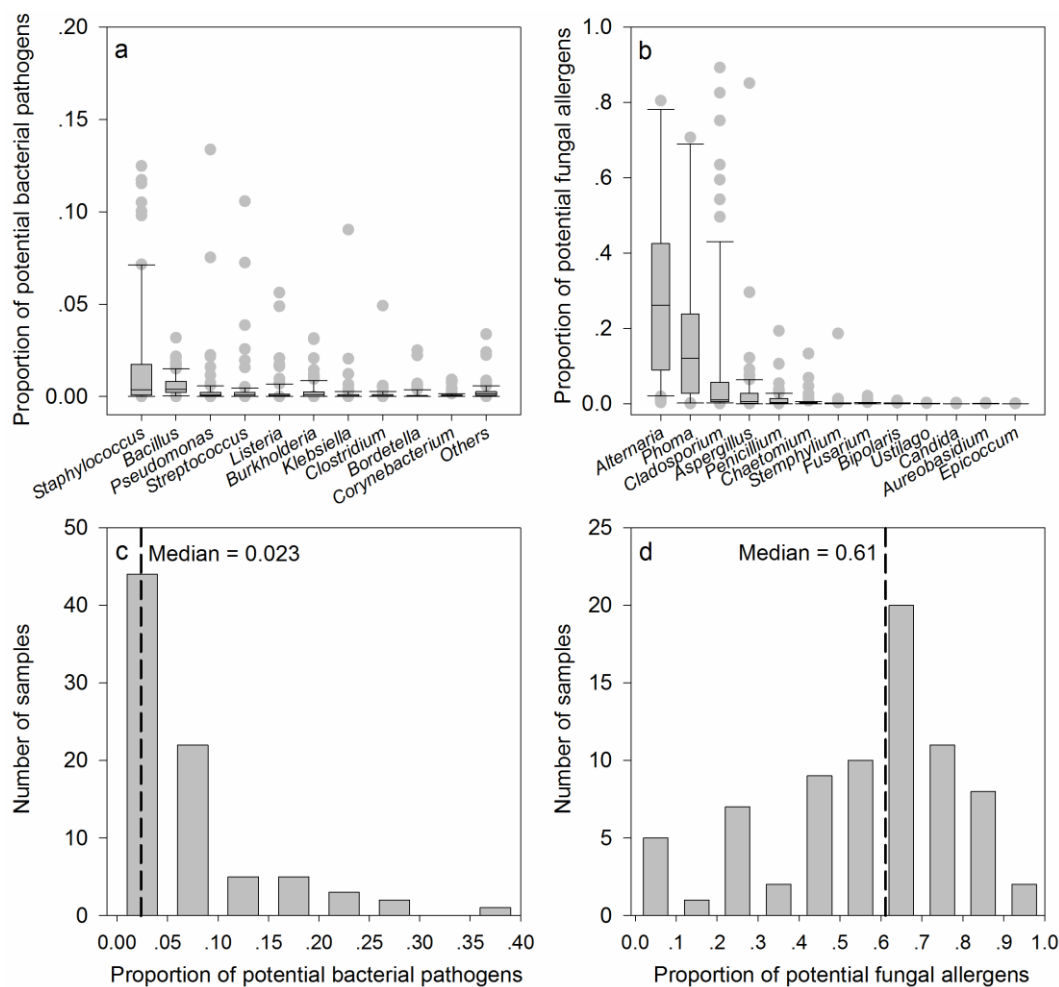


Fig. S7. Profiles of potential bacterial pathogens and fungal allergens in dust samples. **a** and **b.** Proportions of potential bacterial pathogens (a) and fungal allergens (b) at the genus level. “Others” refers to taxa with proportions below 0.01 in all samples. Detailed information of the potential bacterial pathogens identified is shown in [Table S7](#). **c** and **d.** Distribution of proportions of total potential bacterial pathogens (c) and total potential fungal allergens (d) across the samples. Vertical dash lines denote median value.

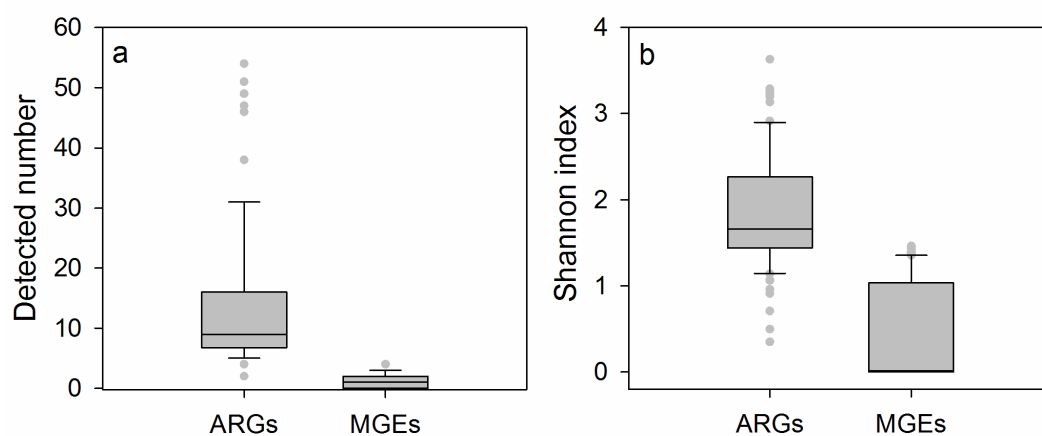


Fig. S8. Detected number (a) and Shannon index (b) of ARGs and MGEs for all dust samples.

ARGs: antibiotic resistance genes; MGEs: mobile genetic elements.

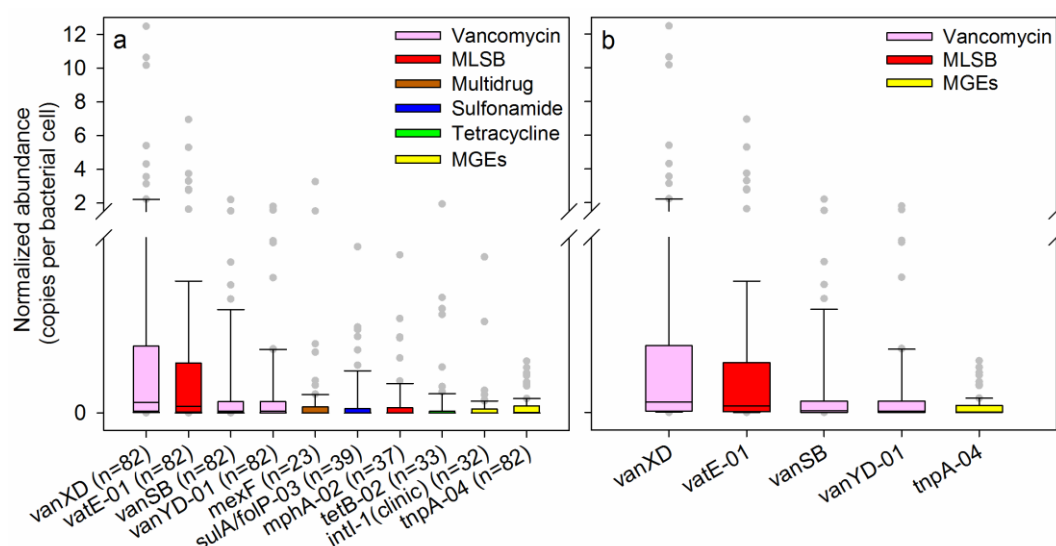


Fig. S9. Normalized abundances (copies per bacterial cell) of ARGs and MGEs across dust samples. (a) The normalized abundances of the top 10 abundant ARGs and MGEs. Number in parenthesis represents the frequency of occurrence of corresponding ARG or MGE across the samples. ARGs were colored by their respective types. MLSB: Macrolide-Lincosamide-Streptogramin B resistance genes. (b) The normalized abundances of the shared ARGs and MGEs.

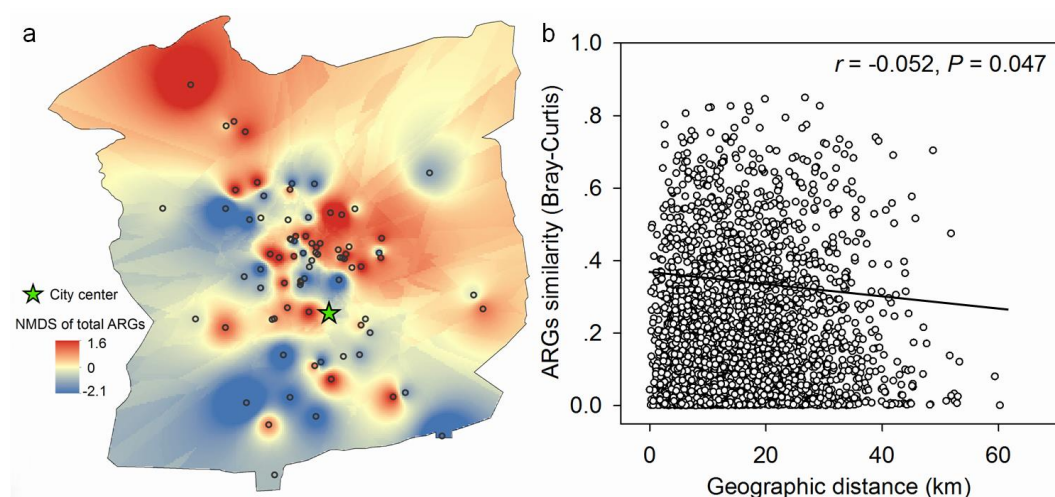


Fig. S10. Geographical distribution pattern of ARG profiles across dust samples. (a) ArcGIS map showing the variation of ARGs according to scores of non-metric multidimensional scaling (NMDS) ordination axis. NMDS analysis was performed based on Bray-Curtis distance of the normalized abundances (copies per bacterial cell) of ARGs across all the samples. Different colors indicate that antibiotic resistome is more distinct in its structure. (b) Relationship between Bray-Curtis similarity of ARGs and geographical distance.

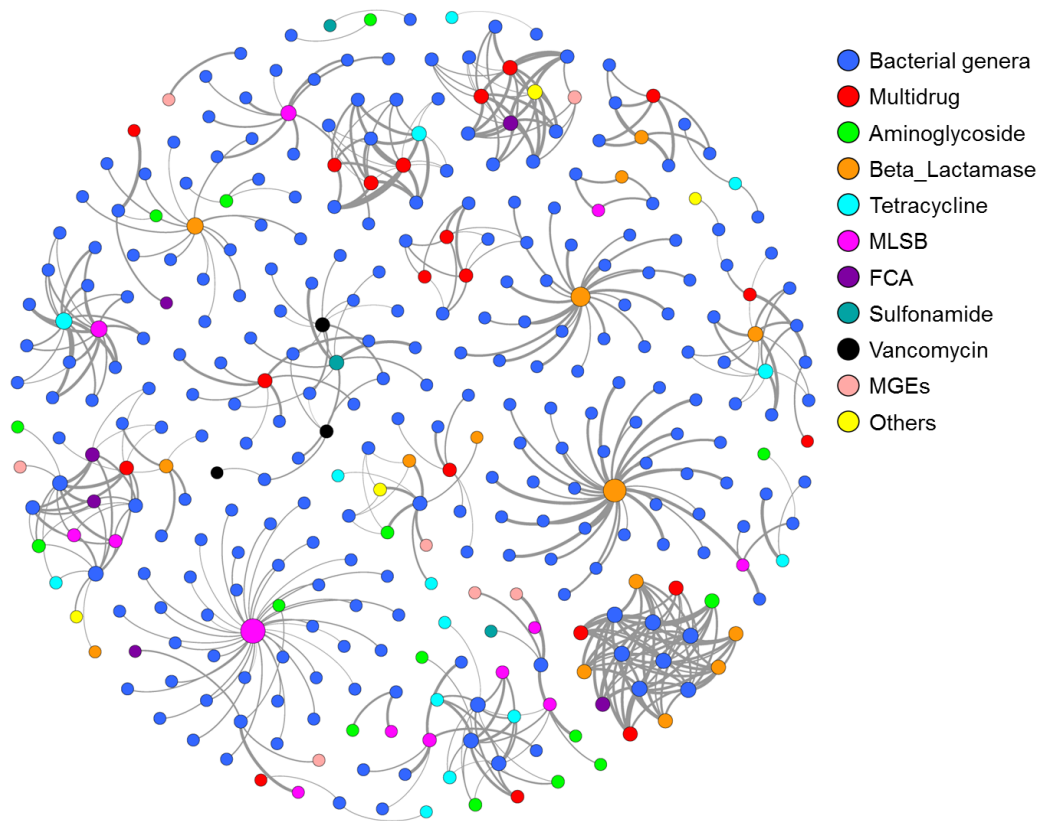


Fig. S11. Co-occurrence patterns between ARGs and bacterial taxa at the genus level visualized by network analysis. Only significant correlations ($P < 0.01$) with a correlation coefficient magnitude (i.e., Spearman's ρ) ≥ 0.7 are shown to concentrate on merely strong correlations. The detailed information of these strong and significant correlations is shown in [Table S9](#). Nodes (circles) are colored according to ARG types and bacterial taxa, and the size of them is proportioned to the number of connections. The width of the edges (i.e., lines connecting the nodes) is weighted based on the correlation coefficient magnitude. MLSB: Macrolide-Lincosamide-Streptogramin B resistance genes; FCA: (flor)/(chlor)/(am)phenicol resistance genes; Others: ARGs conferring resistance to other antibiotic classes like bacitracin, fosfomycin, nitroimidazole, pyrazinamide and streptothricin.

Supplementary References

- Barberán, A., Bates, S.T., Casamayor, E.O., Fierer, N., 2012. Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME J.* 6 (2), 343–351.
- Barberán, A., Ladau, J., Leff, J.W., Pollard, K.S., Menninger, H.L., Dunn, R.R., Fierer, N., 2015. Continental-scale distributions of dust-associated bacteria and fungi. *Proc. Natl. Acad. Sci. U. S. A.* 112 (18), 5756–5761.
- Beijing Municipal Bureau of Statistics, Survey Office of the National Bureau of Statistics in Beijing, 2016. *Beijing Regional Statistical Yearbook*, China Statistics Press, Beijing.
- Chen, Q., An, X., Li, H., Su, J., Ma, Y., Zhu, Y.G., 2016. Long-term field application of sewage sludge increases the abundance of antibiotic resistance genes in soil. *Environ. Int.* 92–93, 1–10.
- Costello, E.K., Lauber, C.L., Hamady, M., Fierer, N., Gordon, J.I., Knight, R., 2009. Bacterial community variation in human body habitats across space and time. *Science* 326, 1694–1697.
- Dunn, R.R., Fierer, N., Henley, J.B., Leff, J.W., Menninger, H.L., 2013. Home life: factors structuring the bacterial diversity found within and between homes. *PLoS One* 8, e64133.
- Engel, P., Moran, N.A., 2013. The gut microbiota of insects-diversity in structure and function. *FEMS Microbiol. Rev.* 37, 699–735.
- Feldhaar, H., 2011. Bacterial symbionts as mediators of ecologically important traits of insect hosts. *Ecol. Entomol.* 36, 533–543.
- Flores, G.E., Bates, S.T., Caporaso, J.G., Lauber, C.L., Leff, J.W., Knight, R., Fierer, N., 2013. Diversity, distribution and sources of bacteria in residential kitchens. *Environ. Microbiol.* 15 (2), 588–596.
- Griffiths, R.I., Thomson, B.C., James, P., Bell, T., Bailey, M., Whiteley, A.S., 2011. The

- bacterial biogeography of British soils. *Environ. Microbiol.* 13 (6), 1642–1654.
- Knights, D., Kuczynski, J., Charlson, E.S., Zaneveld, J., Mozer, M.C., Collman, R.G., Bushman, F.D., Knight, R., Kelley, S.T., 2011. Bayesian community-wide culture-independent microbial source tracking. *Nat. Methods* 8 (9), 761–763.
- Leff, J.W., Jones, S.E., Prober, S.M., Barberán, A., Borer, E.T., Firn, J.L., Harpole, W.S., Hobbie, S.E., Hofmockel, K.S., Knops, J.M.H., McCulley, R.L., La Pierre, K., Risch, A.C., Seabloom, E.W., Schutz, M., Steenbock, C., Stevens, C.J., Fierer, N., 2015. Consistent responses of soil microbial communities to elevated nutrient inputs in grasslands across the globe. *Proc. Natl. Acad. Sci. U. S. A.* 112 (35), 10967–10972.
- Leung, M.H.Y., Tong, X., Wilkins, D., Cheung, H.H.L., Lee, P.K.H., 2018. Individual and household attributes influence the dynamics of the personal skin microbiota and its association network. *Microbiome* 6, 26.
- Redford, A., Bowers, R., Knight, R., Linhart, Y., Fierer, N., 2010. The ecology of the phyllosphere: Geographic and phylogenetic variability in the distribution of bacteria on tree leaves. *Environ. Microbiol.* 12 (11), 2885–2893.
- Shi, Y., Li, Y., Xiang, X., Sun, R., Yang, T., He, D., Zhang, K., Ni, Y., Zhu, Y.G., Adams, J.M., Chu, H. 2018. Spatial scale affects the relative role of stochasticity versus determinism in soil bacterial communities in wheat fields across the North China Plain. *Microbiome* 6, 27.
- Su, J.Q., An, X.L., Li, B., Chen, Q.L., Gillings, M.R., Chen, H., Zhang, T., Zhu, Y.G., 2017. Metagenomics of urban sewage identifies an extensively shared antibiotic resistome in China. *Microbiome* 5, 84.