OPEN

Bacterial community analysis of floor dust and HEPA filters in air purifiers used in office rooms in ILAS, Beijing

Jianguo Guo[1,2], Yi Xiong[3], Taisheng Kang[1,2], Zhiguang Xiang[1,2] & Chuan Qin[1,2]*

Air purifiers with high-efficiency particulate air (HEPA) filters remove not only particulate matter but also airborne microorganisms in indoor environments. We investigated the bacterial community in HEPA filters (used for 1 year) and that in the floor dust of 12 office rooms in Beijing. We found that the viable bacteria proportion in the filter was significantly higher than that in the floor dust ($p < 0.001$). The Non-Metric Multi-Dimensional Scaling analysis showed that the bacterial communities in the filters and dust were significantly different ($p = 0.001$). The Chao1, Shannon–Wiener and phylogenetic diversity values in the filter were significantly higher than those in the dust ($p < 0.001$). The predominant bacterial classes in the filter were *Alphaproteobacteria* and *Actinobacteria*, whereas those in the dust were *Bacteroidia*, *Clostridia* and *Bacilli*. Human occupancy contributed more to the bacterial community in the filter than that in the dust. *Klebsiella* and *Alloprevotella* in the dust and filters positively correlated with the occupancy density. Soil bacteria contributed to a significantly higher proportion of the bacteria in the HEPA filter ($p < 0.001$). In contrast, human oral, indoor air and outdoor haze contributed to a higher proportion of the bacteria in the dust samples ($p < 0.001$, $p < 0.01$ and $p < 0.05$, respectively). As HEPA filters serve as an ecological niche for indoor bacteria, they should be carefully investigated during the assessment of indoor environmental health.

People spend the majority of their time indoors, wherein air pollution levels highly correlate with outdoor pollution levels[1–6]. In recent years, haze pollution, fine particulate matter (PM) in particular, has been reported to increase human mortality rates owing to the development of heart disease[7,8], stroke, respiratory disease[9] and lung cancer[10,11]. The level of air pollution in China is substantially higher than in many developed countries[12]. Moreover, its threat to public health has become a serious concern, particularly concerning haze events[13]. Thus, an effective technology is essential for the removal of PM from indoor environments during periods of haze. To this end, air purifiers with high-efficiency particulate air (HEPA) filters are widely used in the indoor environment, particularly in areas with heavy haze[14].

People are exposed to various microbes in indoor environments. Many factors could affect indoor microbial composition, including human occupancy, occupant behavior, pets and outdoor environment[15–18]. The air purifier was mainly used to remove indoor PM, partly comprising microorganisms. Microbes in the air could be transmitted to the HEPA filter in the air purifier. Suzuki *et al*. (2003) collected swab samples from the air purifier filters in the houses of seven healthy children with varicella. They found the varicella-zoster virus was detected in 29% of filters in air purifiers by day 3[19]. Furthermore, Maus *et al*. (2001) studied the effect of the filters on the viability of microbes. They reported microbes could keep alive for a long time in the filter[20]. Majchrzycka *et al*. (2016) proved that the survivability of microorganisms on filter materials depends on the amount of accumulated moisture and microorganism type[21]. For example, *Staphylococcus aureus* survived well on filter materials with particular condition[21]. Under the working condition, microbes in the HEPA filter could be inspired into the air. Gore *et al*. (2003) reported that the old HEPA-filter vacuum cleaner markedly increases the inspired cat allergen in operation compared with baseline[22]. Spores were also detected downstream of the filters after a longer period

[1]Institute of Laboratory Animal Science, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, 100021, China. [2]NHC Key Laboratory of Human Disease Comparative Medicine, Beijing, 100021, China. [3]Department of Food Science and Engineering, School of Chemistry and Chemical Engineering, Harbin Institute of Technology, Harbin, 150001, China. *email: qinchuan@pumc.edu.cn
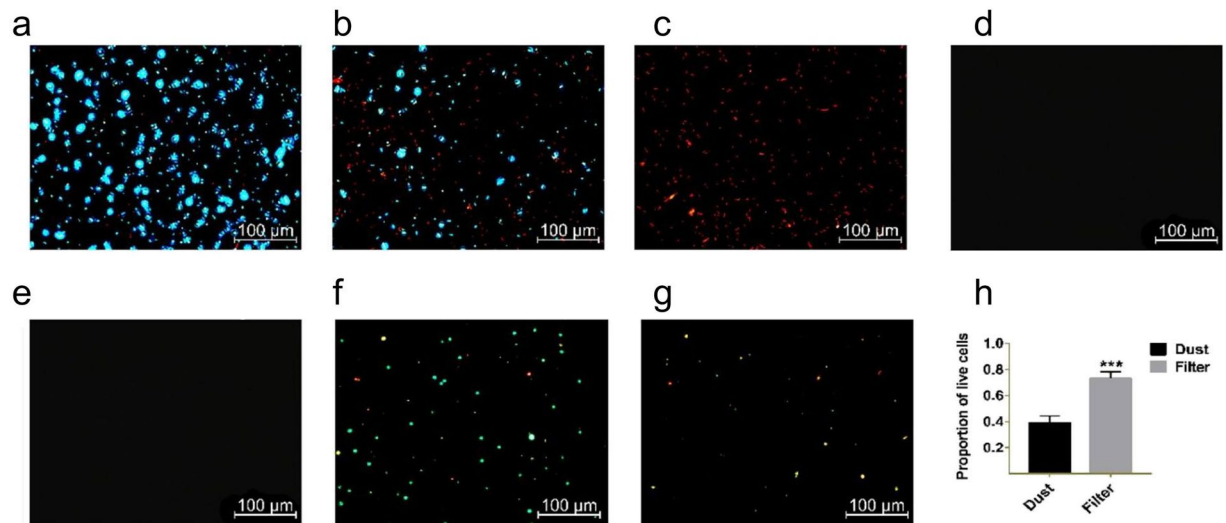
Figure 1. Proportion of the live cells in the dust and filter samples: 100% live cells of *E. coli* (**a**), 50% live cells of *E. coli* (**b**), 0% live cells of *E. coli* (**c**), filter sample without dyeing (**d**), dust sample without dyeing (**e**), filter sample after dyeing (**f**) dust sample after dyeing (**g**), and proportion of live cells in the dust and filter (**h**). Green, live cells; Red, dead cells. '*': $p < 0.05$, '**': $p < 0.01$, '***': $p < 0.001$.

of conditioning when ventilation was restarted[23]. These all indicated the microbes in the HEPA filter could be an important exposure source. Noris *et al.* reported the vast difference between HVAC filter and surface dust in an exploratory investigation in Austin[24]. However, the components of PM were distinctly different around the world. Surface PARTiculate mAtter Network collecting filter samples worldwide showed the chemical components of $PM_{2.5}$ varied by more than an order of magnitude between sites[25]. Besides, one research involving 19 cities around the world showed up to 100-fold in both bacterial and fungal levels among cities[26]. So, the microbial communities in the air purifier are regional. Moreover, limited studies were carried out on the microbial community in HEPA filter in an air purifier, while most focused on microbes in HVAC filter[24,27–31]. Knowledge of microbes in HEPA filter in office was also lacked, except the concentration and the size distribution of microbes[32]. Settled dust (SD), which well reflects the microbial composition in indoor air[33–35], was commonly used to explore the indoor microbial community[36]. It was reported that indoor dust could induce pulmonary inflammation[37,38]. Microorganisms in SD and HEPA filter in the purifiers both come from indoor air. Whether the microbial composition and source in the HEPA filter was similar to SD or should be considered as a new indoor ecological niche? Thus, developing a fundamental understanding of the origins and the characteristics of the microbial community in HEPA filter is a research priority in order to reduce the human health risk of exposure from air purifiers in indoor environments[39].

Bacteria are the main microorganisms in indoor environments[40], and indoor bacteria can be affected by differences in building design, ventilation, the environmental parameters[41–46]. Thus, in our study, we selected 12 office rooms with the same building design with natural ventilation. We investigated the characteristics of the bacterial communities, bacterial diversity and bacterial sources in the HEPA filters in air purifiers used in the indoor environment compared to floor dust collected in the corresponding rooms. Our results would help us to take appropriate measures to reduce the risk in the process of removing PM using air purifiers.

Results

Cell viability results of the bacteria in the filter and dust samples. Live bacteria were stained blue, whereas dead bacteria were stained red. The feasibility of this method was verified using different concentrations of control *E. coli* samples (Fig. 1a–c). Samples without dye showed no fluorescent signal (Fig. 1d,e), indicating that no PM in the samples showed a fluorescent signal. The survival rates of the bacteria in the filter samples were significantly higher than those in the dust samples (Fig. 1f–h).

Bacterial community structure and bacterial diversity in the filter and dust samples. A total of 836,091 and 526,674 raw reads were obtained separately from 24 dust and 12 filter samples, respectively. Chao1 and Good's Coverage both showed the sequencing data generally covered most of the species in the samples when they were rarefied to 20,754 sequences (Supplementary Fig. S1). NMDS analysis revealed that the bacterial community structures showed high similarity within filter samples or dust samples (Fig. 2). However, Jaccard, Bray–Curtis, Unweighted unifrac and Weighted unifrac distance analysis revealed significant differences between the filter and dust samples (Fig. 2). There were no significant differences between different rooms (adonis analysis: $R = 0.69$, $p = 0.50$). A similar bacterial community between rooms was possibly due to a large seasonal fluctuation in the levels of certain microbes, resulting in poor reproducibility[47,48]. In addition, the values of Chao1, Shannon index and phylogenetic diversity (PD) index in the filter samples were significantly greater than those in the dust samples ($p < 0.05$) (Fig. 3).
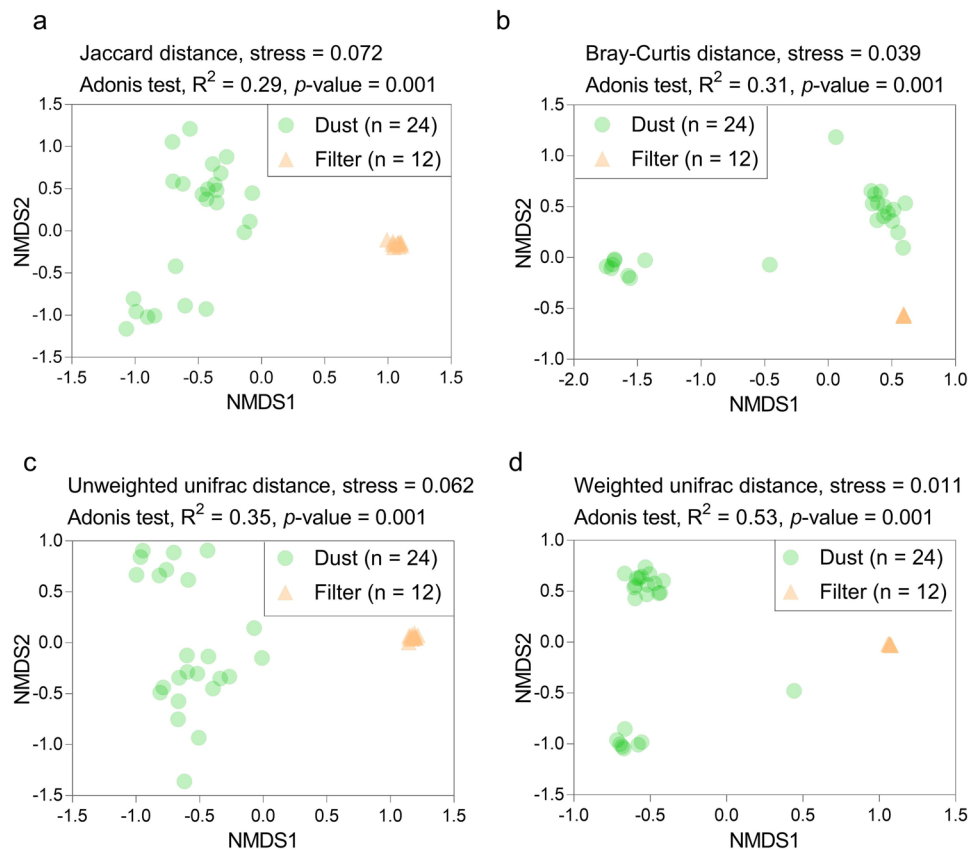
Figure 2. Bacterial community in the dust and filter samples. NMDS analysis based on Jaccard distance (**a**), Bray–Curtis distance (**b**), Unweighted unifrac distance (**c**) and Weighted unifrac distance (**d**). Adonis analysis between the groups was carried out. Stress and the *p*-values are shown in each figure.
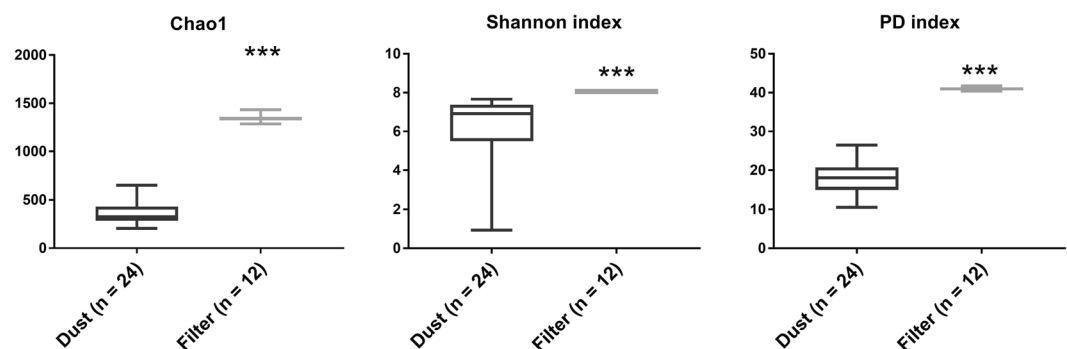

Figure 3. Bacterial diversity in the dust and filter samples—Chao1, Shannon and phylogenetic diversity (PD) index. *Mann–Whitney U test* was carried out. '*': $p < 0.05$, '**': $p < 0.01$, '***': $p < 0.001$.

Bacterial taxa in the filter and dust samples.   The abundance of taxa at the class and genus levels was compared between the filter and dust samples. The abundance of most classes in the filter samples was significantly different from that in the dust samples (Fig. 4a). The major classes in the filter samples were *Alphaproteobacteria* (51.8%) and *Actinobacteria* (17.2%), whereas those in the dust samples were *Bacteroidia* (25.6%), *Clostridia* (13.9%), *Bacilli* (15.9%), *Gammaproteobacteria* (11.7%) and *Alphaproteobacteria* (11.3%) (Fig. 4a). Furthermore, the filter and dust samples showed significant differences in taxa at the genus level. The major genera in the filter samples were *Sphingomonas* (4.7%), *Rubellimicrobium* (3.5%) and *Pseudonocardia* (3.1%), whereas those in the dust samples were *Streptococcus* (11.2%) and *Pantoea* (3.9%) (Fig. 4b).

The results of the operational taxonomic units (OTU) distribution showed that the number of unique OTUs in the filter samples was more than twice as much as that in the dust samples (Fig. 5a). Furthermore, most OTU species of *Cytophagia, Thermoleophilia, Sphingobacteriia, Gemmatimonadetes, Acidimicrobiia, Deltaproteobacteria, Cyanobacteria* and *Acidobacteria* were observed in the filter samples. In contrast, most OTU species of
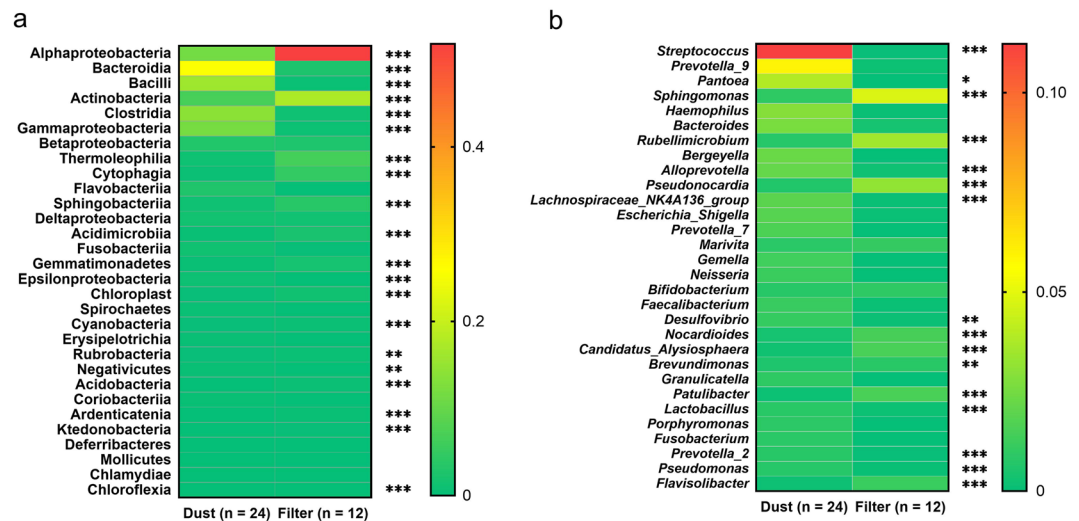
Figure 4. Abundance of top 30 taxa at the class (**a**) and genus (**b**) levels. *Mann–Whitney U test* was carried out in each taxon between dust and filter. The colour represent the mean abundance of each taxon. '*': $p < 0.05$, '**': $p < 0.01$, '***': $p < 0.001$.
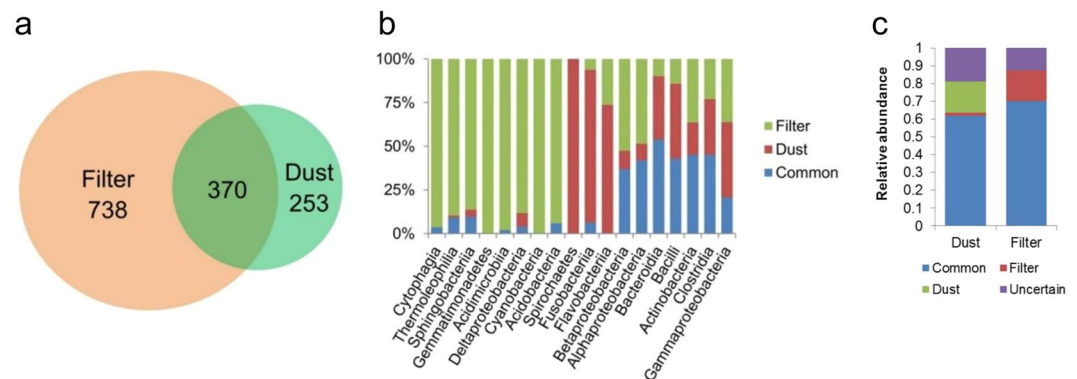

Figure 5. Distribution of operational taxonomic units (OTUs) in the dust and filter. The number of unique and common OTUs in the dust and filter (**a**); the relative proportion of unique and common OTUs in each class (**b**); the abundance of unique and common OTUs in the dust and filter (**c**).

Spirochaetes, Fusobacteria and *Flavobacteriia* were observed in the dust samples (Fig. 5b). The common OTUs made up approximately 60% of total reads, and unique OTUs made up 17.1% and 17.6% of total reads in the filter and dust samples, respectively (Fig. 5c).

Effect of occupancy density on the bacterial community in the filter and dust samples.
Redundancy analysis (RDA) analysis revealed that 1.6% of the bacterial community in the dust samples and 6.4% of the bacterial community in the filter samples attributed to occupancy density of humans in that room (Fig. 6). Although these values are low, the occupancy density had a greater impact on the bacterial community in the filter samples than that in the dust samples. Furthermore, we found that the abundances of *Bacteroidia* and *Ardenticatenia* in the filter samples significantly positively correlated with occupancy density ($p < 0.05$), whereas *Coriobacteriia* in the dust samples was found to be significantly positively correlated the occupancy density (Table 1). Analysis of the relationship between bacterial composition and occupancy density at the genus level revealed that the abundance of *Klebsiella and Alloprevotella* in the dust and filter samples both significantly positively correlated with occupancy density (Table 1). The taxa at the class and genus levels that were significantly negatively correlated with occupancy density were different in the dust and filter samples (Table 1).

Bacterial source.
The in-house bacterial community data of outdoor air, soil, indoor air, human oral and outdoor haze (relative works was performed but not published) were used for the principal coordinates analysis (PCoA) analysis. We found that the Bray–Curtis distance between the filter and soil was less than that between the filter and other groups. The bacterial communities of the dust were clustered into two groups: one was similar to outdoor air samples and the other one was similar to human oral bacteria (Fig. 7a). Source tracking was conducted in the R environment based on the Knight's method[49]. It was found that soil bacteria contributed to
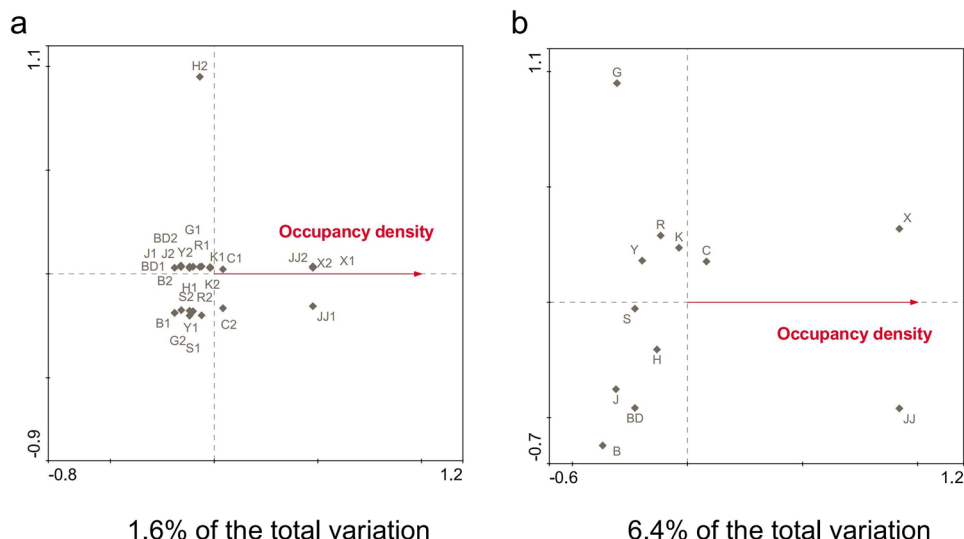
1.6% of the total variation   6.4% of the total variation

Figure 6. Redundancy analysis of genus assemblage data and occupancy density (red lines) in the dust (**a**) and filter (**b**).
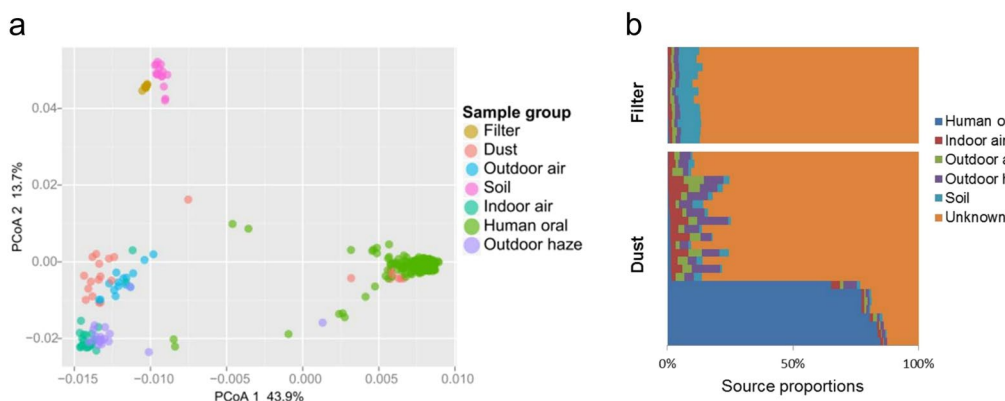

Figure 7. Bacterial source analysis in the dust and filter. Principal coordinates analysis (PCoA) of samples from different sources based on Bray–Curtis distance of bacterial genera (**a**); source tracking using Knight's method (**b**).

		Dust			Filter		
		rho	p	Mean abundance	rho	p	Mean abundance
Class	*Bacteroidia*	—	—	—	0.79	0.002**	0.024
	Ardenticatenia	—	—	—	0.61	0.036*	0.002
	Coriobacteriia	0.62	0.032*	0.003	—	—	—
	Actinobacteria	−0.64	0.025*	0.064	—	—	—
Genus	*Alloprevotella*	0.62	0.03*	0.022	0.64	0.025*	0.0019
	Klebsiella	0.66	0.02*	0.006	0.64	0.026*	0.0001
	Lentzea	−0.67	0.018*	0.001	—	—	—

Table 1. Spearman correlation analysis between occupancy density and the taxa at class level and genus level in the dust and filter. The taxa with mean abundance greater than 0.001 in either group were selected for correlation analysis. Spearman correlation analysis was carried out in stats packages in R environment.

a significantly higher proportion of the bacteria in the filter samples than to that in the dust samples ($p < 0.001$); on the other hand, bacteria from human oral, indoor air and outdoor haze contributed to a significantly higher proportion of the bacteria in the dust samples than to that in the filter samples ($p < 0.001$, $p < 0.01$ and $p < 0.05$, respectively) (Fig. 7b). The higher proportion of human oral source as well as indoor air in the separated samples

in the dust also suggested that there were two main sources of dust bacteria (Fig. 7b). This is consistent with the findings shown in Fig. 7a.

Discussion.   Bacteria in the environment exist primarily by attaching to living and inanimate surfaces[50]. Two typical inanimate surfaces were analyzed in this research. After attachment to the surfaces, sufficient moisture and nutrients are needed for microbial growth[51]. Floor dust collects the airborne microorganisms around the sampling site by natural settlement, and thus, it can accurately reflect the microbial composition and the microbial levels in indoor air[52]. Indoor bacteria that originate from various sources coexist with PM[40,53,54]. Airborne microorganisms, along with PM, moisture and organic matter from the whole room, are pumped into the air purifier compulsively and intercepted on the filter in an air purifier via its high dust-holding capacity[55,56]. Indoor bacteria are trapped in the fine porous structures of the HEPA filters[55], but the trapped bacteria are not killed[57]. Atmospheric dust deposited in air filters may serve as a nutrient for microbes[20]. Accompanied by the absorption of organic matter to the filter, the filter changes gradually, forming a more favorable habitat for the microorganisms[57]. The more nutritious habitat might explain the higher proportion of viable bacteria in the filter (Fig. 1).

Two distinct bacterial community types were formed consistent with the sample types (HEPA filter and dust) (Fig. 2). Indoor microbes might be influenced by the outdoor environment owing to the continual flow of outdoor microbes into homes[43,58] or because indoor environments are incompletely buffered from outdoor microbes[59]. During the usage of the air purifier, indoor air is purified, and particles are enriched on the HEPA film. Meanwhile, a large number of outdoor particles are continuously entering the room. This process enables a large number of outdoor air microorganisms to accumulate on the HEPA film. The types of bacteria in outdoor air can vary according to the function of surrounding soil conditions[60]. Thus, more soil microorganisms are enriched on the HEPA membrane than on the dust because of the high volume of filtered air. The enriched particulate matter membrane helps these microorganisms to survive or reproduce, forming microbial communities similar to the soil environment (Fig. 7a). The bacterial community in floor dust similar to indoor air, outdoor air, outdoor haze and human oral were consistent with main source types[36,43,58,61,62]. Human oral emissions are the largest sources of some floor dust via coughing, sneezing, talking and breathing[63,64]. Among the top 30 genera, more than 10 genera are human-associated bacteria, whereas 5 genera are environmental-associated bacteria. Sources of dominant genera also indicate that humans and the outdoor environment are important bacterial sources in indoor environments[1,65]. The abundance of several human-associated genera was significantly higher in dust samples, including *Streptococcus* (salivary microbiome[66–69]), *Lachnospiraceae* spp. (human and mammal gut microbiota[70]), *Prevotella* spp. and *Lactobacillus* spp. (vaginal microbiota[71,72]) (Fig. 4). *Sphingomonas* spp. and *Brevundimonas* spp. mainly habitat in HEPA filter, which were ubiquitous in the land habitats[73] and contaminated environments[74] (Fig. 4). These were consistent with Fig. 7b. Occupancy density is an important factor affecting the dynamics of the airborne microbial community in the built environment[46]. Its effect on HEPA filter was greater than dust, possibly because of HEPA filter trapped bacteria from indoor air for a much larger volume with a larger amount of human bacteria.

Human oral, indoor air and outdoor haze were critical sources of bacteria in floor dust. In contrast, bacteria present in soil contributed to a significantly higher population in the filter (Fig. 7). The filter in the air purifier and other air-handling equipment could be identified as sources of air contaminants[56,75]. The HEPA filter in the air purifier should be considered as a new ecological niche owing to its enrichment of particles and quite different bacterial sources and bacterial community structure. Air purifiers remove PM by pumping air, and the environmental bacteria are trapped by the HEPA filters[55,57]. Thus, various bacteria possibly tend to keep alive on the filter. Human activities such as walking could greatly increase the emission rate of biological aerosol particles that originated from the surrounding environment[76]. These aerosols may cause adverse effects, including respiratory infections and hypersensitivities in humans[77–79]. Thus, the alive bacteria in the HEPA filter should be considered during the assessment of indoor environmental health.

Conclusions.   Air purifiers are widely used in indoor environments in haze days. Our study showed that the HEPA filters in air purifiers used for a long time were with a higher proportion of viable bacteria compared with the dust. Considering the significantly different bacterial communities and bacterial diversity between the filter and dust samples, HEPA filter in the air purifier should be recognized as a new ecological niche in indoor environment. Source tracking analysis explained the formation of the environment in the filter and dust. The important sources of bacteria were soil and human oral, indoor air and outdoor haze in the case of filter and dust, respectively. The number of unique OTUs in the filter samples was more than twice as much as that in the dust samples. The unique OTUs made up 17.1% and 17.6% of the total reads in the filter and dust samples, respectively. Our study suggested that air purifiers should be included in the evaluation of indoor environmental health risks.

Methods

Site description and sample collection.
Twelve office rooms were selected as the sampling rooms in the Institute of Laboratory Animal Science (ILAS) in Beijing, China. These rooms were used as administrative offices that were not related to laboratory and animal houses. These rooms were distributed in three different buildings, and each room was independent of each other. Each room had openable windows. Ventilation was carried out by opening windows in no haze day. All the rooms were covered with tiles. Different types of house dust samples are widely used to assess the health effects of indoor microbes representing airborne inhalation exposure. Although it was reported a better reproducibility of microbial determinations has been found for mattress dusts[80,81], settled dust was suitable to study the microbes in the office room. Settled dust reflects the microbial composition of indoor air and responds similarly to environmental determinants[36]. In March 2016, two randomly different areas on these room floors, which were larger than 2 × 2 cm² with little or no direct contact with the occupants, were disinfected with bromogeramine and marked as floor dust sampling sites, and a new HEPA filter was set

in the air purifier in each of these rooms. The flow rate of the air purifier was 116 m³/h. The air purifiers were applied in haze day in which a 24-hour average concentration was greater than 75 μg/m³. The sampling rooms and the occupant numbers in each room are shown in Supplementary Table S1. Twenty-four dust sampling sites were left undisturbed, and 12 air purifiers were used without replacing the filter for 1 year. The air purifier used days in each room were different, owing to differed office plans and arrangements. The range was 121–143 days, except for holidays. In March 2017, dust samples were collected with disposable sterile gloves and masks from the 24 dust sampling sites. A sterile cotton swab was used with 100 μL sterile physiological saline (salt content: 0.9%) to smear the floor dust and kept in 1.5 mL sterile Eppendorf tubes at −80 °C before using them for DNA isolation. The filter samples were collected with disposable sterile gloves and masks from the 12 air purifiers (Supplementary Fig. S2). The concentrations of PM$_{2.5}$ and PM$_{10}$ in Beijing from March 2016 to March 2017 are shown in Supplementary Fig. S3. Maestre *et al.* found that vacuum and swab samples of HVAC filters were repeatable and generally superior to cut samples[28]. We thought it was easy to handle the filter and control operation consistency by cut. Furthermore, the dust recovery methodologies did not affect the contribution of environmental and human sources to the bacterial and fungal communities in HVAC filter[28]. Thus, we selected to sample by cut. A 2 × 2 cm² area of the outer membrane filter was sheared and cut with sterile surgical scissors. The filter fragments were placed in 5 mL of saline and vortexed for 20 min. The supernatants were pipetted out and centrifuged at 13,000 rpm for 2 min. Subsequently, the supernatants were removed, and the pellets were maintained in 1.5 mL sterile Eppendorf tubes at −80 °C before DNA isolation.

Samples from outdoor air, soil, indoor air, human oral and outdoor haze were used to study the source of bacteria in the filter and dust. The detailed information of the samples is presented in Table 2. Sample sequencing steps were similar to those performed for the filter and dust samples.

Cell viability analysis of the dust and filter samples.

The cell viability of the samples was determined by Invitrogen™ LIVE/DEAD® BacLight™ Bacterial Viability kits, which was used to evaluate the cell membrane integrity[82]. These kits utilize mixtures of the SYTO® 9 green-fluorescent nucleic acid stain and the red-fluorescent nucleic acid stain, propidium iodide. Bacterial suspensions from dust and filter samples were obtained by mixing the bacterial pellet with 100 μL saline. The *Escherichia coli* (*E. coli*) suspension with 100%, 50% and 0% live cells was prepared as positive controls, and samples without dyeing were prepared as a negative control. The remaining steps were performed as directed by the manufacturer.

DNA extraction and Illumina sequencing of 16S rRNA genes.

Genomic DNA was extracted from the dust and filter samples by using the PowerSoil® DNA Isolation Kit (MO BIO Laboratories, Inc., Carlsbad, CA, USA) according to the manufacturer's recommendations. The quality and quantity of DNA were verified by NanoDrop and agarose gel electrophoresis (Table S1). There was no significant difference in DNA content between HEPA filter and floor dust (Fig. S4). The extracted DNA was diluted to a concentration of 1 ng/μL and stored at −20 °C until further processing. The diluted DNA was used as the template for PCR amplification (26 cycles) of bacterial 16 S rRNA genes with barcoded primers and HiFi Hot Start Ready Mix (KAPA). For the bacterial diversity analysis, the V3–V4 variable region of the 16 S rRNA genes was amplified with universal primers (343 F 5′-TACGGRAGGCAGCAG-3′, 798 R 5′-AGGGTATCTAATCCT-3′)[83]. The length of the PCR product was about 445 bp. Negative control was used in the same amplification system by using sterile deionized water as the template to monitor the contamination of the amplification system. The amplicon quality was visualized by gel electrophoresis, followed by the purification of PCR product with AMPure XP beads. Amplification for another round of PCR (7 cycles) was conducted, and the product was purified again. In each batch, we ran negative controls. We monitored the PCR systems to avoid contamination of the PCR product with undesirable DNA. If no 400 bp band was detected in the negative control, the samples in the same batch were processed to the sequencing steps. The final amplicon concentration was then quantified using the Qubit™ dsDNA Assay Kit. Equal amounts of the purified amplicons were pooled for subsequent sequencing with MiSeq Reagent Kit v2 (Illumina, Inc., San Diego, CA, USA) using the MiSeq Sequencing System (Illumina, Inc., San Diego, CA, USA). The PE300 sequencing model was used, and paired ends were applied.

Binning sequences into species.

For further taxonomy analysis, the community-by-operational taxonomic unit (OTU) matrix was generated as follows. The raw sequencing data after Illumina sequencing were in the FASTQ format. The paired-end reads were then pre-processed using the Trimmomatic software to detect and cut off ambiguous bases (N)[84]. It also cuts off low-quality sequences with average quality scores below 20 by using the sliding window trimming approach. After trimming, the paired-end reads were assembled using the FLASH software[85]. The parameters of the assembly were as follows: 10 bp of minimal overlapping, 200 bp of maximum overlapping, and 20% of the maximum mismatch rate. Further denoising of the sequences was performed as follows: reads with ambiguous sequences, homologous sequences, or sequences below 200 bp were abandoned. Reads with 75% of the bases above Q20 were retained. Then, reads with the chimera were detected and removed[86]. These two steps were achieved using the QIIME software (version 1.8.0)[87].

Clean reads were subjected to primer sequence removal and clustering to generate OTUs using the UPARSE software with 97% similarity cutoff[88]. The representative read of each OTU was selected using the QIIME package. All the representative reads were annotated and blasted against the Silva database Version 123 (16 S rDNA) using the RDP classifier (confidence threshold was 70%)[89]. Each sample was rarefied to 20,754 sequences.

Bacterial diversity and other statistical analyses.

To ensure sampling depth, the Chao1 and Good's Coverage Estimator was calculated to analyze the species diversity in each sample. The α-diversity[90] of the bacterial community was analyzed on the basis of Chao1, Shannon–Wiener and Faith's phylogenetic diversity (PD)[91] index using the new OTU matrix file. The phylogenetic tree was built using approximately maximum likelihood

Group	Source	Sampling method	Sampling process	Sampling time	Sampling sites
Indoor air	indoor air	Natural sedimentation	Place uncovered, disposable, sterile cell culture dishes (φ90 mm) at a height of 1.5 m above the ground and left them for 12 h, Samples were then collected by smearing the inner surface of each dish with a sterile swab that had been dipped in 100 μL of sterile normal saline, and the swabs were used for DNA extraction.	December 9 to 27, 2016	Bei Jing
Outdoor haze	outdoor haze	Natural sedimentation	At 8:00 and 20:00, 200 ml deionized water was used for mouthwashes collection. The samples were centrifuged with 12000 g for 10 min, and the deposit was used for DNA extraction.		
Human oral	human oral	Mouthwashes			
Outdoor air	outdoor air	Air impact method	3000 L air was collected with Planktonic airborne bacteria sampler FKC1 with flux of 100 L/min onto the surface of sterile cell culture dishes covered with 200 ul 20% glycerol. Samples were then collected by smearing the inner surface of each dish with a sterile swab that had been dipped in 100 μL of sterile normal saline, and the swabs were used for DNA extraction.	September 12 to 13, 2017	Shi Jiazhuang
Soil	soil	Digger blade	Remove 5 cm of topsoil with disposable sterile shovel, and collect 20 g soil with sterile 50 ml tube. The soil was used for DNA extraction		

Table 2. The sample information used in source tracking analysis.

based on the evolutionary distance calculated by Jukes and Cantor model with QIIME. The phylogenetic tree was used for PD index, Unweighted unifrac and Weighted unifrac distance analysis. We performed non-metric multi-dimensional scaling (NMDS) analysis[92] to compare the bacterial community structure in the filter and dust based on Jaccard, Bray–Curtis, Weighted and Unweighted unifrac distances 92. Permutational multivariate analysis of variance (Adonis) was also carried out using the package 'vegan' in the R environment. All other statistical analyses were performed using the package 'state' in the R environment.

The dominant 30 taxa at the class and genus level were used to examine the differences of bacterial composition between the dust and filter. Mann–Whitney U test was performed, and *p-values* less than 0.05 were considered statistically significantly different. A unique OTU distribution was also analyzed. The OTU that existed in more than three samples in dust and less than three samples in the filter was considered as unique in the dust. The OTU that existed in more than three samples in the filter and less than three samples in dust was considered as unique in the filter.

Human occupants are an important source of indoor microorganisms. We performed a Spearman correlation analysis between DNA content and room area, the number of people and occupancy density and found no correlation (Supplementary Table S2). We explored the Spearman correlation analysis between taxa (comprising > 0.1%) at the class and genus level and occupancy density. To compare the differences in the effect of occupancy density on the bacterial community in filter and dust, RDA was performed for dust and filter samples using CANOCO 4.5 software (http://canoco.software.informer.com/4.5/). In this study, dominant genera (comprising > 0.1%) were used to analyze the correlations with occupancy density by RDA. A PCoA was performed based on the Bray–Curtis distance of genera among samples from different sources. Source tracking was carried out in the R environment based on the Knights' method[49].

All bacterial 16 S rRNA gene sequences generated for the present study were deposited in the National Centre for Biotechnology Information Sequence Read Archive (http://www.ncbi.nlm.nih.gov/sra) under the accession number SRP160378.

Data availability

All bacterial 16 S rRNA gene sequences generated for the present study have been deposited in the National Centre for Biotechnology Information Sequence Read Archive (http://www.ncbi.nlm.nih.gov/sra) under the accession number SRP160378.

References

1. Hospodsky, D. *et al*. Human Occupancy as a Source of Indoor Airborne Bacteria. *PLoS One* **7**, 10, https://doi.org/10.1371/journal.pone.0034867 (2012).
2. Fujiyoshi, S., Tanaka, D. & Maruyama, F. Transmission of Airborne Bacteria across Built Environments and Its Measurement Standards: A Review. *Frontiers in Microbiology* **8**, https://doi.org/10.3389/fmicb.2017.02336 (2017).
3. Adams, R. I., Bateman, A. C., Bik, H. M. & Meadow, J. F. Microbiota of the indoor environment: a meta-analysis. *Microbiome* **3**, 18, https://doi.org/10.1186/s40168-015-0108-3 (2015).
4. Han, Y. *et al*. Influences of ambient air PM2.5 concentration and meteorological condition on the indoor PM2.5 concentrations in a residential apartment in Beijing using a new approach. *Environ. Pollut.* **205**, 307–314, https://doi.org/10.1016/j.envpol.2015.04.026 (2015).
5. Thatcher, T. L. & Layton, D. W. DEPOSITION, RESUSPENSION, AND PENETRATION OF PARTICLES WITHIN A RESIDENCE. *Atmospheric Environment* **29**, 1487–1497, https://doi.org/10.1016/1352-2310(95)00016-r (1995).
6. Baek, S. O., Kim, Y. S. & Perry, R. Indoor air quality in homes, offices and restaurants in Korean urban areas - Indoor/outdoor relationships. *Atmospheric Environment* **31**, 529–544, https://doi.org/10.1016/s1352-2310(96)00215-4 (1997).

7. Pope, C. A. III *et al.* Cardiovascular Mortality and Exposure to Airborne Fine Particulate Matter and Cigarette Smoke Shape of the Exposure-Response Relationship. *Circulation* **120**, 941–948, https://doi.org/10.1161/circulationaha.109.857888 (2009).

8. Pope, C. A. *et al.* Cardiovascular mortality and long-term exposure to particulate air pollution - Epidemiological evidence of general pathophysiological pathways of disease. *Circulation* **109**, 71–77, https://doi.org/10.1161/01.cir.0000108927.80044.7f (2004).

9. Zhang, Q. L., Qiu, Z. M., Chung, K. F. & Huang, S. K. Link between environmental air pollution and allergic asthma: East meets West. *Journal of Thoracic Disease* **7**, 14–22, https://doi.org/10.3978/j.issn.2072-1439.2014.12.07 (2015).

10. Pope, C. A. III *et al.* Lung Cancer and Cardiovascular Disease Mortality Associated with Ambient Air Pollution and Cigarette Smoke: Shape of the Exposure-Response Relationships. *Environmental Health Perspectives* **119**, 1616–1621, https://doi.org/10.1289/ehp.1103639 (2011).

11. Lepeule, J., Laden, F., Dockery, D. & Schwartz, J. Chronic Exposure to Fine Particles and Mortality: An Extended Follow-up of the Harvard Six Cities Study from 1974 to 2009. *Environmental Health Perspectives* **120**, 965–970, https://doi.org/10.1289/ehp.1104660 (2012).

12. Zhao, J. H. *et al.* Spatiotemporal Trend Analysis of PM2.5 Concentration in China, 1999-2016. *Atmosphere* **10**, https://doi.org/10.3390/atmos10080461 (2019).

13. Guan, W. J., Zheng, X. Y., Chung, K. F. & Zhong, N. S. Impact of air pollution on the burden of chronic respiratory diseases in China: time for urgent action. *Lancet* **388**, 1939–1951, https://doi.org/10.1016/S0140-6736(16)31597-5 (2016).

14. Peck, R. L. *et al.* Efficiency of portable HEPA air purifiers against traffic related combustion particles. **98**, 21–29 (2016).

15. Casas, L. *et al.* Endotoxin, extracellular polysaccharides, and (1-3)-glucan concentrations in dust and their determinants in four European birth cohorts: results from the HITEA project. *Indoor air* **23**, 208–218, https://doi.org/10.1111/ina.12017 (2013).

16. Sordillo, J. E., Alwis, U. K., Hoffman, E., Gold, D. R. & Milton, D. K. Home Characteristics as Predictors of Bacterial and Fungal Microbial Biomarkers in House Dust. *Environmental Health Perspectives* **119**, 189–195, https://doi.org/10.1289/ehp.1002004 (2011).

17. Johansson, E. *et al.* Streptomycetes in house dust: associations with housing characteristics and endotoxin. *Indoor air* **21**, 300–310, https://doi.org/10.1111/j.1600-0668.2010.00702.x (2011).

18. Frankel, M. *et al.* Seasonal Variations of Indoor Microbial Exposures and Their Relation to Temperature, Relative Humidity, and Air Exchange Rate. *Applied and Environmental Microbiology* **78**, 8289–8297, https://doi.org/10.1128/aem.02069-12 (2012).

19. Suzuki, K. *et al.* Spread of varicella-zoster virus DNA to the environment from varicella patients who were treated with oral acyclovir. *Pediatrics international: official journal of the Japan Pediatric Society* **45**, 458–460, https://doi.org/10.1046/j.1442-200x.2003.01746.x (2003).

20. Maus, R., Goppelsroder, A. & Umhauer, H. Survival of bacterial and mold spores in air filter media. *Atmospheric Environment* **35**, 105–113, https://doi.org/10.1016/s1352-2310(00)00280-6 (2001).

21. Majchrzycka, K., Okrasa, M., Skora, J. & Gutarowska, B. Evaluation of the Survivability of Microorganisms Deposited on Filtering Respiratory Protective Devices under Varying Conditions of Humidity. *International Journal of Environmental Research and Public Health* **13**, https://doi.org/10.3390/ijerph13010098 (2016).

22. Gore, R. B. *et al.* High-efficiency particulate arrest–filter vacuum cleaners increase personal cat allergen exposure in homes with cats. *Journal of Allergy and Clinical Immunology* **111**, 784–787, https://doi.org/10.1067/mai.2003.1378 (2003).

23. Forthomme, A. *et al.* Microbial aerosol filtration: Growth and release of a bacteria-fungi consortium collected by fibrous filters in different operating conditions. *Journal of Aerosol Science* **72**, 32–46, https://doi.org/10.1016/j.jaerosci.2014.02.004 (2014).

24. Noris, F., Siegel, J. A. & Kinney, K. A. Evaluation of HVAC filters as a sampling mechanism for indoor microbial communities. *Atmospheric Environment* **45**, 338–346, https://doi.org/10.1016/j.atmosenv.2010.10.017 (2011).

25. Snider, G. *et al.* Variation in global chemical composition of PM2.5: emerging results from SPARTAN. *Atmospheric Chemistry and Physics* **16**, 9629–9653, https://doi.org/10.5194/acp-16-9629-2016 (2016).

26. Li, J. *et al.* Differing toxicity of ambient particulate matter (PM) in global cities. *Atmospheric Environment* **212**, 305–315, https://doi.org/10.1016/j.atmosenv.2019.05.048 (2019).

27. Liu, Z. J. *et al.* Effect of environmental parameters on culturability and viability of dust accumulated fungi in different HVAC segments. *Sustainable Cities and Society* **48**, https://doi.org/10.1016/j.scs.2019.101538 (2019).

28. Maestre, J. P. *et al.* Filter forensics: microbiota recovery from residential HVAC filters. *Microbiome* **6**, https://doi.org/10.1186/s40168-018-0407-6 (2018).

29. Liu, Z. J., Ma, S. Y., Cao, G. Q., Meng, C. & He, B. J. Distribution characteristics, growth, reproduction and transmission modes and control strategies for microbial contamination in HVAC systems: A literature review. *Energy and Buildings* **177**, 77–95, https://doi.org/10.1016/j.enbuild.2018.07.050 (2018).

30. Liu, Z. J., Zhu, Z. Q., Zhu, Y. X., Xu, W. & Li, H. Investigation of dust loading and culturable microorganisms of HVAC systems in 24 office buildings in Beijing. *Energy and Buildings* **103**, 166–174, https://doi.org/10.1016/j.enbuild.2015.06.056 (2015).

31. Noris, F., Siegel, J. A., Kinney, K. A. & Ashrae. In *Ashrae Transactions 2009*, Vol 115, Pt 2 Vol. 115 ASHRAE Transactions 484-491 (2009).

32. Golofit-Szymczak, M. & Gorny, R. L. Microbiological air quality in office buildings equipped with different ventilation systems. *Indoor air* **28**, 792–805, https://doi.org/10.1111/ina.12495 (2018).

33. Frankel, M., Timm, M., Hansen, E. W. & Madsen, A. M. Comparison of sampling methods for the assessment of indoor microbial exposure. *Indoor air* **22**, 405–414, https://doi.org/10.1111/j.1600-0668.2012.00770.x (2012).

34. Noss, I. *et al.* Evaluation of a low-cost electrostatic dust fall collector for indoor air endotoxin exposure assessment. *Applied and Environmental Microbiology* **74**, 5621–5627, https://doi.org/10.1128/aem.00619-08 (2008).

35. Kilburg-Basnyat, B., Peters, T. M., Perry, S. S. & Thorne, P. S. Electrostatic dust collectors compared to inhalable samplers for measuring endotoxin concentrations in farm homes. *Indoor air* **26**, 724–733, https://doi.org/10.1111/ina.12243 (2016).

36. Leppanen, H. K. *et al.* Quantitative assessment of microbes from samples of indoor air and dust. *Journal of Exposure Science and Environmental Epidemiology* **28**, 231–241, https://doi.org/10.1038/jes.2017.24 (2018).

37. Kim, Y. S. *et al.* Extracellular vesicles, especially derived from Gram-negative bacteria, in indoor dust induce neutrophilic pulmonary inflammation associated with both Th1 and Th17 cell responses. *Clinical and Experimental Allergy* **43**, 443–454, https://doi.org/10.1111/cea.12085 (2013).

38. Shan, Y. F., Wu, W. D., Fan, W., Haahtela, T. & Zhang, G. C. House dust microbiome and human health risks. *International Microbiology* **22**, 297–304, https://doi.org/10.1007/s10123-019-00057-5 (2019).

39. Sarnet, J. M. The perspective of the National Research Council's Committee on Research Priorities for Airborne Particulate Matter. *J. Toxicol. Env. Health Part A* **68**, 1063–1067, https://doi.org/10.1080/15287390590935905 (2005).

40. Zhai, Y. *et al.* A review on airborne microorganisms in particulate matters: Composition, characteristics and influence factors. *Environment International* **113**, 74–90, https://doi.org/10.1016/j.envint.2018.01.007 (2018).

41. Dannemiller, K. C., Gent, J. F., Leaderer, B. P. & Peccia, J. Influence of housing characteristics on bacterial and fungal communities in homes of asthmatic children. *Indoor air* **26**, 179–192, https://doi.org/10.1111/ina.12205 (2016).

42. Kettleson, E. M. *et al.* Key determinants of the fungal and bacterial microbiomes in homes. *Environmental Research* **138**, 130–135, https://doi.org/10.1016/j.envres.2015.02.003 (2015).

43. Kodama, A. M. & McGee, R. I. Airborne Microbial Contaminants In Indoor Environments - Naturally Ventilated And Air-Conditioned Homes. *Archives of Environmental Health* **41**, 306–311, https://doi.org/10.1080/00039896.1986.9936702 (1986).

44. Amend, A. S., Seifert, K. A., Samson, R. & Bruns, T. D. Indoor fungal composition is geographically patterned and more diverse in temperate zones than in the tropics. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 13748–13753, https://doi.org/10.1073/pnas.1000454107 (2010).

45. Kembel, S. W. *et al*. Architectural design influences the diversity and structure of the built environment microbiome. *Isme Journal* **6**, 1469–1479, https://doi.org/10.1038/ismej.2011.211 (2012).

46. Meadow, J. F. *et al*. Indoor airborne bacterial communities are influenced by ventilation, occupancy, and outdoor air source. *Indoor air* **24**, 41–48, https://doi.org/10.1111/ina.12047 (2014).

47. Kaarakainen, P. *et al*. Microbial content of house dust samples determined with qPCR. *Science of the Total Environment* **407**, 4673–4680, https://doi.org/10.1016/j.scitotenv.2009.04.046 (2009).

48. Leppanen, H. K. *et al*. Determinants, reproducibility, and seasonal variation of ergosterol levels in house dust. *Indoor air* **24**, 248–259, https://doi.org/10.1111/ina.12078 (2014).

49. Knights, D. *et al*. Bayesian community-wide culture-independent microbial source tracking. *Nature methods* **8**, 761–763, https://doi.org/10.1038/nmeth.1650 (2011).

50. Donlan, R. M. Biofilm Formation: A Clinically Relevant Microbiological Process. *Clin. Infect. Dis.* **33**, 1387–1392 (2001).

51. Stetzenbach, L. D. *Airborne Bacteria in Indoor Environments*. (John Wiley & Sons, Inc., 2006).

52. Leppä¤nen, H. K. *et al*. Quantitative assessment of microbes from samples of indoor air and dust. 28 (2017).

53. Maki, T. *et al*. Variations in the structure of airborne bacterial communities in a downwind area during an Asian dust (Kosa) event. *Sci. Total Environ.* **488-489**, 75–84 (2014).

54. Smets, W., Moretti, S., Denys, S. & Lebeer, S. Airborne bacteria in the atmosphere: Presence, purpose, and potential. *Atmospheric Environment* **139**, 214–221, https://doi.org/10.1016/j.atmosenv.2016.05.038 (2016).

55. Stocks, G. W., O'Connor, D. P., Self, S. D., Marcek, G. A. & Thompson, B. L. Directed Air Flow to Reduce Airborne Particulate and Bacterial Contamination in the Surgical Field During Total Hip Arthroplasty. *Journal of Arthroplasty* **26**, 771–776, https://doi.org/10.1016/j.arth.2010.07.001 (2011).

56. Korves, T. M. *et al*. Bacterial communities in commercial aircraft high-efficiency particulate air (HEPA) filters assessed by PhyloChip analysis. *Indoor air* **23**, 50–61, https://doi.org/10.1111/j.1600-0668.2012.00787.x (2013).

57. Pokhum, C. *et al*. A facile and cost-effective method for removal of indoor airborne psychrotrophic bacterial and fungal flora based on silver and zinc oxide nanoparticles decorated on fibrous air filter. *Atmos. Pollut. Res.* **9**, 172–177, https://doi.org/10.1016/j.apr.2017.08.005 (2018).

58. Adams, R. I., Miletto, M., Taylor, J. W. & Bruns, T. D. Dispersal in microbes: fungi in indoor air are dominated by outdoor air and show dispersal limitation at short distances. *Isme Journal* **7**, 1262–1273, https://doi.org/10.1038/ismej.2013.28 (2013).

59. Martin, L. J. *et al*. Evolution of the indoor biome. *Trends in Ecology & Evolution* **30**, 223–232, https://doi.org/10.1016/j.tree.2015.02.001 (2015).

60. Bowers, R. M., McLetchie, S., Knight, R. & Fierer, N. Spatial variability in airborne bacterial communities across land-use types and their relationship to the bacterial communities of potential source environments. *Isme Journal* **5**, 601–612, https://doi.org/10.1038/ismej.2010.167 (2011).

61. Barberan, A. *et al*. The ecology of microscopic life in household dust. *Proceedings of the Royal Society B-Biological Sciences* **282**, 212–220, https://doi.org/10.1098/rspb.2015.1139 (2015).

62. Rintala, H., Pitkaranta, M. & Taubel, M. In *Advances in Applied Microbiology, Vol 78 Vol. 78 Advances in Applied Microbiology* (eds A. I. Laskin, S. Sariaslani, & G. M. Gadd) 75-120 (Elsevier Academic Press Inc, 2012).

63. Xie, X., Li, Y., Sun, H. & Liu, L. Exhaled droplets due to talking and coughing. *J. R. Soc. Interface* **6**, S703–S714, https://doi.org/10.1098/rsif.2009.0388.focus (2009).

64. Nicas, M., Nazaroff, W. W. & Hubbard, A. Toward understanding the risk of secondary airborne infection: Emission of respirable pathogens. *Journal of Occupational and Environmental Hygiene* **2**, 143–154, https://doi.org/10.1080/15459620590918466 (2005).

65. Adams, R. I. *et al*. Chamber Bioaerosol Study: Outdoor Air and Human Occupants as Sources of Indoor Airborne Microbes. *PLoS One* **10**, 18, https://doi.org/10.1371/journal.pone.0128022 (2015).

66. Xu, J. *et al*. Evolution of symbiotic bacteria in the distal human intestine. *PLoS. Biol.* **5**, 1574–1586, https://doi.org/10.1371/journal.pbio.0050156 (2007).

67. Wang, K. *et al*. Preliminary analysis of salivary microbiome and their potential roles in oral lichen planus. *Sci Rep* **6**, https://doi.org/10.1038/srep22943 (2016).

68. Chotai, S. *et al*. Brain Abscess Caused by Gemella Morbillorum: Case Report and Review of the Literature. *Turk. Neurosurg.* **22**, 374–377, https://doi.org/10.5137/1019-5149.jtn.3634-10.0 (2012).

69. Mor, G. & Kwon, J.-Y. Trophoblast-microbiome interaction: a new paradigm on immune regulation. *American Journal of Obstetrics and Gynecology* **213**, S131–S137, https://doi.org/10.1016/j.ajog.2015.06.039 (2015).

70. Ruan, J. Bergey's Manual of Systematic Bacteriology (second edition) Volume 5 and the study of Actinomycetes systematic in China. *Wei sheng wu xue bao = Acta microbiologica Sinica* **53**, 521–530 (2013).

71. Sagrista, M. *et al*. Inguinal syndrome secondary to Prevotella bivia after accidental bite in orogenital sex. *Sex. Transm. Infect.* **88**, 250–251, https://doi.org/10.1136/sextrans-2011-050348 (2012).

72. Fettweis, J. M. *et al*. Differences in vaginal microbiome in African American women versus women of European ancestry. *Microbiology-(UK)* **160**, 2272–2282, https://doi.org/10.1099/mic.0.081034-0 (2014).

73. Shi, P. L. *et al*. The response of soil bacterial communities to mining subsidence in the west China aeolian sand area. *Appl. Soil Ecol.* **121**, 1–10, https://doi.org/10.1016/j.apsoil.2017.09.020 (2017).

74. Li, X. H. *et al*. Diversity of chlorpyrifos-degrading bacteria isolated from chlorpyrifos-contaminated samples. *Int. Biodeterior. Biodegrad.* **62**, 331–335, https://doi.org/10.1016/j.ibiod.2008.03.001 (2008).

75. Tringe, S. G. *et al*. The airborne metagenome in an indoor urban environment. *PLoS One* **3**, e1862 (2008).

76. Bhangar, S. *et al*. Chamber bioaerosol study: human emissions of size-resolved fluorescent biological aerosol particles. *Indoor air* **26**, 193–206, https://doi.org/10.1111/ina.12195 (2016).

77. Burge, H. Bioaerosols - Prevalence And Health-Effects In The Indoor Environment. *J. Allergy Clin. Immunol.* **86**, 687–701, https://doi.org/10.1016/s0091-6749(05)80170-8 (1990).

78. Clark, S., Rylander, R. & Larsson, L. Airborne bacteria, endotoxin and fungi in dust in poultry and swine confinement buildings. *Am. Ind. Hyg. Assoc. J.* **44**, 537–541, doi:10.1202/0002-8894(1983)044<0537:abeafi>2.3.co;2 (1983).

79. Dales, R. E. *et al*. Influence of ambient fungal spores on emergency visits for asthma to a regional children's hospital. *Am. J. Respir. Crit. Care Med.* **162**, 2087–2090, https://doi.org/10.1164/ajrccm.162.6.2001020 (2000).

80. Hyvarinen, A. *et al*. Dust sampling methods for endotoxin - an essential, but underestimated issue. *Indoor air* **16**, 20–27, https://doi.org/10.1111/j.1600-0668.2005.00392.x (2006).

81. Park, J. H. *et al*. Longitudinal study of dust and airborne endotoxin in the home. *Environmental Health Perspectives* **108**, 1023–1028, https://doi.org/10.1289/ehp.001081023 (2000).

82. Manteca, A., Fernandez, M. & Sanchez, J. J. M. A death round affecting a young compartmentalized mycelium precedes aerial mycelium dismantling in confluent surface cultures of Streptomyces antibioticus. **151**, 3689–3697 (2005).

83. Simeng, L., Zhuangzhuang, W. & Guoqiang, L. Degradation kinetics of toilet paper fiber during wastewater treatment: Effects of solid retention time and microbial community. *Chemosphere* **225**, 915–926, https://doi.org/10.1016/j.chemosphere.2019.03.097 (2019).

84. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
85. Reyon, D. *et al.* FLASH assembly of TALENs for high-throughput genome editing. *Nature Biotechnology* **30**, 460 (2012).
86. Edgar, R. C., Haas, B. J., Clemente, J. C., Christopher, Q. & Rob, K. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**, 2194–2200 (2011).
87. Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* **7**, 335 (2010).
88. Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Methods* **10**, 996 (2013).
89. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied & Environmental Microbiology* **73**, 5261 (2007).
90. Li, B., Zhang, X., Guo, F., Wu, W. & Zhang, T. J. W. R. Characterization of tetracycline resistant bacterial community in saline activated sludge using batch stress incubation with high-throughput sequencing analysis. **47**, 4207–4216 (2013).
91. Faith, D. P. Conservation evaluation and phylogenetic diversity. *Biological Conservation* **61**, 1–10, https://doi.org/10.1016/0006-3207(92)91201-3 (1992).
92. Magali, N. R. *et al.* A microbiota signature associated with experimental food allergy promotes allergic sensitization and anaphylaxis. **131**, 201–212 (2013).

Author contributions

J.G. and C.Q. designed the study; J.G. and Y.X. performed experiments; J.G., T.K. and Z.X. analyzed the data; J.G. and Y.X. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-63543-1.

Correspondence and requests for materials should be addressed to C.Q.

Reprints and permissions information is available at www.nature.com/reprints.

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