Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
A metagenomic-based method to study hospital air dust resistome

Xiang Li a,b, Ziqi Wu a, Chenyuan Dang a, Miao Zhang a, Bixi Zhao a, Zhanwen Cheng a, Liming Chen a, Zhenfeng Zhong c, Yuhui Ye c, Yu Xia a,b,∗

a School of Environmental Science and Engineering, College of Engineering, Southern University of Science and Technology, Shenzhen 518055, China
b State Environmental Protection Key Laboratory of Integrated Surface Water-Groundwater Pollution Control, School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China
c Shenzhen Hospital of Peking University, 1120 Lianhua Rd, Futian, Shenzhen, Guangdong 518036, China

HIGHLIGHTS
• Hospital air-conditioners can form resistome and accumulate pathogens.
• The Outpatient hall can distribute ARGs to other departments.
• Evidence-based network strategy proves that plasmid-mediated ARG transfer can occur frequently.

ARTICLE INFO
Keywords:
Airdust
Antibioticresistancegenes
Resistome
Plasmid-mediatedARGtransfer
Metagenomic-basedmethod

ABSTRACT
As a symbol of the defense mechanisms that bacteria have evolved over time, the genes that make bacteria resist antibiotics are overwhelmingly present in the environment. Currently, bacterial antibiotic resistance genes (ARGs) in the air are a serious concern. Previous studies have identified bacterial communities and summarized putative routes of transmissions for some dominant hospital-associated pathogens from hospital indoor samples. However, little is known about the possible indoor air ARG transportation. In this study, we mainly surveyed air-conditioner air dust samples under different airflow conditions and analyzed these samples using a metagenomic-based method. The results show air dust samples exhibited a complex resistome, and the average concentration is 0.00042 copies/16S rRNA gene, which is comparable to some other environments. The hospital air-conditioners can form resistome over time and accumulate pathogens. In addition, our results indicate that the Outpatient hall is one of the main ARG transmission sources, which can distribute ARGs to other departments (explains > 80% resistome). We believe that the management should focus on ARG carrier genera such as Staphylococcus, Micrococcus, Streptococcus, and Enterococcus in this hospital and our novel evidence-based network strategy proves that plasmid-mediated ARG transfer can occur frequently. Overall, these results provide insights into the characteristics of air dust resistome and possible route for how ARGs are spread in air.
1. Introduction

Antibiotics were invented for preventing bacterial infections and saving lives [1]. However, as a symbol of the defense mechanisms that bacteria have evolved over time, the genes that make bacteria resist antibiotics are overwhelmingly present in the environment after extensive antibiotic applications worldwide [2,3]. Not only are water-borne and soil-related resistant bacteria threatening our health, but bacterial ARGs in the air are a serious concern. Some research groups have examined ARGs from the outdoor air, and those resistances seemed to be closely related to air quality parameters. For instance, Li et al. [4] showed PM$_2.5$ possessed a high concentration of ARG against bacterial ARGs in the air areas is a serious concern. Some research groups have examined ARGs from the outdoor air, and those resistances seemed to be closely related to air quality parameters. For instance, Li et al. [4] showed PM$_2.5$ possessed a high concentration of ARG against bacteria.

Humans are closely tethered to the indoor environment. From the hospital we are born in to the homes and offices we live in and work in, the indoor environments have become the most intimate ecosystem for humans. Lax et al. studied many aspect related to indoor micro-ecology [6–8]. They revealed human microbiota can influence indoor microbial communities and also be shaped by a built-in environment. Micro-ecology associated with hospital settings arguably has the most profound implications because hospital-acquired infections (HAIs) have always been considered as one of the main causes of patient deaths [7]. Therefore, lots of effort has been focused on determining how microorganisms colonize, persist, and change in the hospital’s indoor environment. Many previous studies have identified dominant hospital-associated pathogens (HAPs) and summarized their putative routes of transmissions. Those starting routes mainly include physicians’ and staff’s clothing [9–11], stethoscopes [12], phones [13–16], keyboards [17], and even patient bedrails [8]. In any case, little is known about the possible indoor air ARG and/or pathogen transmission route.

In this study, a metagenomic-based method was applied to provide insights regarding possible ARG transmission in hospital indoor air. Meanwhile, bacterial cultivation, PCR, qPCR, and Nanopore sequencing approaches were also applied to provide support to our analyses. Overall, findings in this work support the notion that 1) the Outpatient hall is one of the main ARG transmission sources, and it can affect other departments under different airflow conditions; 2) management practice should focus on bacterial genera such as *Staphylococcus*, *Micrococcus*, *Streptococcus*, and *Enterococcus*; 3) plasmid-mediated ARG transfer can occur frequently among departments.

2. Materials and methods

2.1. Site description and sample collections

Two sampling rounds were performed in summer (Aug. 2018) and winter (Jan. 2019) at Shenzhen Hospital of Peking University, respectively. According to the management practice at this hospital, windows are closed in summer (enclosed space) from early April to the end of October, but windows are regularly open during winter seasons (natural ventilation). Air dust samples were collected from the Outpatient hall (sometimes called main entrance, abbr. as M), the Ophthalmology (abbr. as O), the Pediatrics (abbr. as P) and the Inpatient departments (1 denotes summer and 2 denotes winter sampling, e.g. M1 means samples were collected at the Outpatient hall in summer sampling). 2 to 3 air conditioner strainers per sampling site were firstly detached by cleaning experts and then transported back to our lab on ice within 2 h after sampling. Notably, we were asked to collect the strainers before the cleaning experts rinsed them periodically (usually thoroughly cleaned after 1 to 2 months usage). After the strainers were received at our lab, the dirt on them was shaken and washed off into sterilized autoclavable containers using sterilized ddH$_2$O. Thereafter, the non-soluble and dissolved dirt samples were all filtered immediately using 0.22 μm filters (Millipore, USA). The filtrated membranes were then transferred to purple-capped tubes in a FastDNA kit (MP Biomedicals, USA) and stored at $-20^\circ$C before downstream DNA extractions were performed (see Fig. S1 for an overall schematic overview). Besides the air dust samples, we also surveyed water, door handle and lobby chair surface swabs as internal sample comparisons in the studied hospital (refer to addendum in SI).

2.2. Bacterial isolations and DNA extractions

Portions of soluble air dust samples were used for colony counts. Six types of antibiotic LB agar plates (each has only one antibiotic of ampicillin, chloromycetin, meropenem, kanamycin, tetracycline or erythromycin) were used to screen for antibiotic resistant bacteria (see Table S1). All agar plates were prepared within one week before use and incubated at 37°C overnight with the potential for re-incubation up to 48 h. Bacterial isolates were then selected for genomic sequencing using Oxford Nanopore Technology (ONT). A whole genome DNA extraction method (Easy-DNA kit, Thermal Fisher, USA) was applied to those bacteria following manufacture instructions. We also applied a FastDNA spin kit (MP Biomedicals, USA) on filter membranes following the operational instructions. Quality controls were incorporated in both cultivations and extractions to monitor for operational errors.

2.3. PCR, qPCR, and sequencing

The universal 16S rRNA (27F and 1492R) primer pair was used to identify bacterial cells (Table S2) [18]. We used Premix Taq reagents (TaKaRa Taq V2 plus dye, Dalian, China) to perform PCR reactions with a BioRad T100 Thermal Cycler. The 25 μl reaction contained 12.5 μl premix Taq reagent, 1 μl (10 μm) forward primer, 1 μl (10 μm) reverse primer, 1 μl DNA template and 9.5 μl water. 30 cycles of 94°C 30 sec, 60°C 30 sec, and 72°C 1.5 min were used for all amplifications. Meanwhile, a total of seven qPCR assays that target antibiotic resistance and integron genes were applied. We directly used the published primers from previous studies and the seven assays included *blaNDM* [19], *intI1* [20], *qnrA* [19], *sul2* [19], *vanA* [21] and *tetW* [22]. The qPCR reactions were performed by Applied Biosystems QuantStudio 5 Real-Time PCR System (Thermo Fisher Scientific, USA). The reaction contained 5 μl 2 X PCR mixture with 1 μl forward primer (10 μm), 1 μl reverse primer (10 μm) and 1 μl DNA template in a total of 10 μl volume (2 μl water). We used 2 × plus SYBR real-time PCR mixture (Biotek Corporation, Beijing, China) to conduct the two-step qPCR cycling. 40 cycles of 94°C 15 sec and 60°C 30 sec were run with initial denaturing at 94°C for 5 min. All qPCR assays have one additional dissociation curve step. Data were summarized and analyzed by QuantStudio™ Design & Analysis Software (Thermal Fisher, USA) and Microsoft Excel. In this study, we considered results positive if 2 out of 3 qPCR replicates were amplified.

Both the Illumina and ONT sequencing technologies were used in this study. We used Illumina HiseqXten-PE150 platform to perform metagenomics sequencing (Novogene, Nanjing, China). Each department was designated to have at least two samples (one sample per season, but there were two replicates in M2 and P2, respectively). The sequencing library size was 350 bp and 10 G raw data was initially aimed for. For bacterial whole genome sequencing, the extracted genomic DNAs were purified by AMPure XP (Beckman Coulter, USA) following the instructions (gel electrophoresis was also checked), then the intact genomic DNA samples were gently and evenly mixed. Rapid Barcoding Sequencing (SQK-RBK004) was used for the library preparation following the manufacture instructions. Finally, GridION X5 platform was used to initiate Nanopore sequencing on R9.4 flowcell for 48 h.

2.4. Metagenomic sequencing data analyses

We applied FastQC (v0.11.8) for metagenomic data quality checks.
In brief, O1 had a “sequence duplication” error, and M2 had a “per sequence GC content” issue (floppy at the first 6 reads). Those two issues were addressed by the BBmap programs dedup.sh and bbduk2.sh, respectively before downstream analyses. We testified the reproducibility of our sequencing methodology using two replicates in M2 and P2. Notably, only one replicate from M2 and P2 samples were used for the subsequent read analysis (see details below). In this study, metagenomic analyses were based on reads and metagenome-assemblies, respectively. The following four subsections focused on read analyses, and the last subsection dealt with assembly analyses.

### 2.4.1. ARG analysis

The rapid characterization and quantification of ARGs were analyzed by ARGs-OAP v2.0, which was developed by Yin et al [23]. Briefly, both 16S and ARG type/subtype information were extracted from reads at species level. First, the ARG profiling was calculated by normalizing the extracted ARG sequences against reads that mapped to 16S rRNA gene. Both ARG type and subtype (equivalent to genotype) results were summarized in Supplementary file 1.

#### 2.4.2. Community and functionality analyses

MetaPhlAn2 was used to compute the composition of microbial communities at species-level resolution [24]. Since MetaPhlAn2 can generate species-level resolution, we searched the pathogen species in the microbial community profile to generate a pathogen list (Supplementary file 2). The community profiling results can be visualized in Fig. S2. We also applied HUMAnN2 to summarize microbial communities at species-level resolution [25]. This was achieved by measuring M2 between principle coordination analysis (PCoA) plots (Bray-Curtis distance), and compared the M2 values with a set of empirically distributed values determined from 999 permutations. Each bacterial species was determined by Sanger sequencing before Nanopore sequencing. The Nanopore sequencing fast5 files that passed QC were basecalled with albacore [39]. Next, Porechop (https://github.com/rrwick/Porechop) was used to trim adapters. The resulting fasta files were assembled by Canu 1.8 [40]. Since we performed Sanger sequencing, the estimated -genomeSize option in Canu was correctly important for correction, trim, and assemble steps. A recently published plasmid database PLSDB [41] was used to search for plasmid-like DNA fragments (99% similarity and 6000 bp in length for successfully assembled contigs or > 2000 bp for Nanopore raw reads). ResFinder 2.1 [42] was used to identify ARGs from Nanopore reads (90% similarity and 60% alignment cutoff). The detailed results were summarized in Table S4 & S5.

### 2.4.3. 16S rRNA identifications and OTU analysis

In this study, we used SortMeRNA (version 2.1b) [26] to extract 16S rRNA sequences from metagenomic reads. The silva-bac-16s-id90.fasta database was used with -fastx -paired_in -aligned output -log -a 8 reads. The detailed results were summarized in Table S3.

### 2.4.4. Bayesian source tracking

In this study, we focused on two types of source tracking analyses using SourceTracker [29]. Since patients, doctors, and visitors all start their routines at M, we assume it can serve as a “source” in our Bayesian algorithm, leaving other departments as a “sink.” Each OTU table was firstly sorted based on abundances, and then commonly shared information was extracted and reorganized into matrix format for all samples (Supplementary file 2). Besides OTU source tracking, we also took ARG genotype profiling results in source tracking analyses (Fig. 2).

### 2.4.5. ARG carrier and plasmid identifications from metagenome-assemblies

We assembled metagenomic sequencing reads separately (clustered by departments). All assembly processes were performed using CLC Genomics Workbench 12 (Qiagen Bioinformatics, USA). The program was run by default setting with minor changes: 1) the minimum contig length was set to 500 bp instead of 200 bp; 2) length fraction (mapping process) was set to 0.8. After M, O, and P metagenome-assemblies were obtained using de novo assembly method, open read frames (ORFs) in each assembly were predicted by MetaGeneMark [30]. Diamond [31] was then applied to search these ORFs against RefSeq protein (Jul. 2017) database (E = 10\(^{-5}\)). The blast format output of Diamond was parsed by MEGANS [32] for taxonomic assignment. Then assembled contigs were taxonomically classified to a Kingdom if more than 50% of the genes within a contig were attributed to the same Kingdom. The same threshold was used for Phylum and Class level taxonomies, while 40% was applied for order, followed by 34% for family, 30% for genus, and 25% for species [33–35]. Taxonomic annotations generated from this homology-based method were served to identify the phylogenetic affiliation of ARG carrying bacteria. BlastP [36] was used to search ORFs against SARG database with similarity cutoff of 80% and alignment length cutoff of 70% to identify ARGs from these metagenome-assemblies. Genus level carrier with two seasonal sequencing read partition results (normalized after mapping reads back to contigs) [37] were summarized in Fig. 4 and Supplementary file 5. In this study, the PlasFlow.py program was used to predict the existence of plasmids in all metagenome-assemblies [38]. To increase result accuracy, we changed the –threshold probability parameter to 0.9 (0.7 by default). The assemblies were then subdivided into chromosomal and plasmid contigs, and the plasmid contigs were used for further genotype sharing network analyses.

### 2.5. ARG and plasmid identifications from bacterial isolates

We selected 11 cultural isolates for Nanopore sequencing but only 9 isolates were found to have plasmids (see Table S4). Briefly, the selected bacterial colony was firstly enriched in LB broth, and then after cell precipitation by centrifuge, whole genome DNA extraction was performed. Each bacterial species was determined by Sanger sequencing before Nanopore sequencing. The Nanopore sequencing fast5 files that passed QC were basecalled with albacore [39]. Next, Porechop (https://github.com/rrwick/Porechop) was used to trim adapters. The resulted fasta files were assembled by Canu 1.8 [40]. Since we performed Sanger sequencing, the estimated -genomeSize option in Canu was correctly important for correction, trim, and assemble steps. A recently published plasmid database PLSDB [41] was used to search for plasmid-like DNA fragments (99% similarity and 6000 bp in length for successfully assembled contigs or > 2000 bp for Nanopore raw reads). ResFinder 2.1 [42] was used to identify ARGs from Nanopore reads (90% similarity and 60% alignment cutoff). The detailed results were summarized in Table S4 & S5.

### 2.6. Statistics and network analyses

Spearman’s correlation tests were performed using the cor function in base-R (3.5) to confirm the reproducibility between replicates. We checked the community and ARG correlations, respectively on all replicates (Fig. S3). Meanwhile, to compare air dust samples, we performed an analysis of similarities (ANOSIM, R package “vegan” [43]) using the shared bacterial OTU matrix. Procrustes analysis was performed to illustrate the correlation between resistome and community profiling [44]. In brief, the Procrustes transformation significance was calculated by measuring M2 between principle coordination analysis (PCoA) plots (Bray-Curtis distance), and compared the M2 values with a set of empirically distributed values determined from 999 computational permutations. In each permutation, the sum of squared distances between matched samples (M2) was used to compute a p-value. PCoA analysis was also applied to distinguish ARG subtypes of various environmental samples based on ARGs-OAP v2.0 results.

An ARG genotype sharing network was constructed to search for plasmid-mediated ARG transfer. First, two-round metagenomic sequencing results were clustered (normalized results that were generated from MetaPhlAn2 and ARGs-OAP v2.0). Second, only ARG hits with similarity greater than 0.99 were considered. Third, identified plasmids were used to clarify the possible ARG transfer among hospital departments. Forth, the plasmid-like fragments identified from Nanopore sequencing were also integrated to confirm plasmid movements (Fig. 3). The ARG carriers’ networks were shown in Fig. 4 and Fig. S4, and we also networked the ARGs that were identified by Nanopore isolate sequencing regarding shared ARG genotypes and plasmid-like fragments information (Figs. S5 and S6). All network graphs were created using Gephi 0.9.2 and draw.io software.
3. Results and discussion

3.1. Metagenomic data summary and α diversity of air dust samples

The quality of all post-QC clean reads was high (Q30 over 93% for all samples). The metagenomics datasets had 162 Gbp post-QC clean reads, meaning each air dust sample had over 20 Gbp high-quality data on average (see Table S6). In this study, two replicates from winter sampling, namely P2 and M2, were used for reproducibility analyses, and the non-parametric Spearman’s tests showed the correlations ranged from 0.77 to 0.93 (r-value) regarding OTU and ARGs, respectively (Fig. S3). The reproducibility results indicate that our metagenomic sequencing results are reliable. After CLC de novo assembly, P had 1347130 assembled contigs (at least 500 bp in length), followed by M (607008 contigs) and O (240311 contigs). N50 contig measurements (including scaffolded region) from O, P and M were 791 (max 162632), 1118 (max 1287280) and 1733 (max 120352) bp in length, respectively. From rRNA α diversity results, we can see summer sampling had higher Good’s coverage but lower Shannon index’s values, suggesting the diversity of summer air dust samples is lower than winter air dust samples.

3.2. Prevalence of resistome in air dust

All bacterial isolation negative controls exhibited no bacteria growth and the positive controls showed antibiotics functioned properly in this study. From Table S1, we can see that the resistome were present in P1, P2 and O1, and 5 out of 6 antibiotic types were detected (no ampicillin). Meanwhile, PCR showed all air dust samples were bacterial positives but all the extraction controls were negative, indicating our DNA extraction step is contamination free (see Table S2). Both cultivation and PCR showed that the resistome were prevalent in air dust at this hospital. Our findings suggest antibiotic resistant airborne bacteria could freely move in the hospital air and accumulate in

Fig. 1. Hospital air dust comparative metagenomic results. Fig. 1A is Procrustes superimposition plot that depicts the confident correlation between air dust resistome (Bray-Curtis) and community compositions (Bray-Curtis); Fig. 1B is principal coordinate analysis based on ARG subtypes of various environmental samples: CFA represents chicken farm air, HWA is hospital winter air dust, HSA is hospital summer air dust; STP represents sewage treatment plant.

Fig. 2. Bayesian source tracking results. M stands for the Outpatient hall, O represents the Ophthalmology and the Pediatrics is abbreviated as P. Two season’s M air dust samples as source and other departments as sink for SourceTracker.
Fig. 3. ARG genotype sharing network. The different colored oval represents different departments studied. Vertical ellipse within each oval represents ARG genera that were identified with their relative abundances in summer (yellow) and winter (blue, shown as a bar chart). Circle-pairs represent ARG genotypes. Genotypes are linked to the carriers as network edges. The size of the circle is proportional to its abundance (normalized against 16S rRNA gene) within air dust samples. Circle border represents the mobile mechanisms: plasmid carrying (dotted line) and chromosome encoding (full line). ARG genotypes shared among departments are shown in the middle area outside oval and each sharing instance was further linked with dash lines; those plasmids (red) and chromosomal (green) are shown as different colors. Those sharing instances are confirmed by whole-genome Nanopore sequencing (\( > 0 \rightarrow < 0 \)).

Fig. 4. Network of ARG genus carriers among different departments within the studied hospital. The grey node size is proportional to the total relative abundance (in term of ppm) of ARG carriers. The double-layers ring overlaying each grey node shows the proportion of ARG type carried by the corresponding genus. The inner and outer rings represent the ARG type proportions in the enclosed space (summer) and natural ventilation (winter) airflow conditions, respectively. Edges are color coded according to the sampling departments (color nodes) and the width is proportional to the relative abundance of corresponding ARG carriers.
the air-conditioner, which was also consistent with others’ results [45–49]. For qPCR assays, we only obtained positive results from intI1 and tetW. One explanation might be we have only applied correct candidate primers for those two assays since both intI1 and tetW primers used here were previously applied in hospital settings. Therefore, our results indicated that the selection of hospital ARG primers was necessary in our case since ARG genotypes varied greatly [23]. The standard curves (intI1 and tetW) have higher R² values (> 0.983), and the amplification efficiencies varied from 1.18 to 1.23 (see Supplementary file 4). The lowest limit of detection (LOD) value is 10⁴ gene copies/reaction. The qPCR enumerations showed intI1 exhibited 5.95 log₁₀ copies/reaction (on average), and the average tetW concentration was 6.14 log₁₀ copies/reaction. Metagenomic sequencing only generated short reads and required computational resources to reconstruct the “broken” genomic fragments. Therefore, inconsistencies between cultural and metagenomic sequencing are unavoidable [50]. In this study, we noticed there are inconsistencies among culture, PCR technologies, and metagenomic sequencing, indicating different methods have different sensitivity and specificity abilities to some degree. But the throughput advantage is outstanding for the metagenomic sequencing approach. Notably, both qPCR and metagenomic results detected tetW, but ARG carrier analysis missed its presence, suggesting there are also disadvantages when using metagenomic sequencing (e.g., sequencing coverage, depth and PCR bias issues, etc.).

3.3. Air dust resistome and gene function profiles identified by metagenomic-based method

From ARGs-OAP v2.0 data analyses of 1208 total ARG subtypes, P1 had 167 hits followed by 144 in M1 and only 42 in O1 (see Supplementary file 1). However, P2 had 384 hits followed by O2 and M2 with 357 and 344, respectively. Meanwhile, similar findings were noticed using a diversity and network analyses (Fig. S4). Moreover, after we grouped ARG subtype detections, 86 occurred in winter but only 11 occurred in summer (only β-lactam and multidrug subtypes). The average normalized ARG concentration in summer was 0.0000257, 0.0000761 and 0.0000132 copies per 16S rRNA gene for M1, P1 and O1, respectively; and the corresponding values changed to 0.000522, 0.0000536 and 0.000524 in winter. Overall, air dust had an average of 0.00042 copies per 16S rRNA gene, which was less than the wastewater influent samples (0.000601). Besides, drinking water, wastewater effluent, and ocean samples had concentrations of 0.00024, 0.000178, and 0.000341 copies per 16S rRNA gene, respectively.

ARG carrier results showed Escherichia, Bacteroides, Staphylococcus, and Acinetobacter were the main ARG carriers at M, but Xanthobacter, Pseudodermonas, Pseudomonas, Spirochaeta, Neisseria, and Legionella, etc. were commonly found at both O and P (Fig. 4). The most abundant aminoglycoside resistome was Pseudomonas (21.24) at O2, and M1 had 23.86 tetracycline abundance in Acinetobacter (Supplementary file 5). The previously published indoor airborne bacterial studies reported similar findings. For instance, Pseudomonas was found to be the most frequently detected airborne bacteria in two intensive care units (ICU) according to a study that was performed in central Taiwan [51]. Another consistency proof was that a group of researchers from Iran found the hospital indoor air was a potential route of transmission of beta-lactam resistant Acinetobacter and Staphylococcus [47]. Even though ANOSIM showed no significant differences among all air dust samples (p = 0.67), the Procrustes analysis could help us separate the air dust resistome samples. In Fig. 1A, we can see a dispersed scattering in summer (enclosed-space) but a more tightly clustered shape in winter (natural ventilation), indicating there might be distinct bacterial community and resistome characteristics under different airflow conditions. Additionally, from the PCoA analysis (Fig. 1B), we noticed hospital air dust ARG compositions were closely related to human activities.

The MetaPhlAn2 microbial community results can be seen in Fig. S2. We can see that most of the species were bacteria, but we also detected polymavirus, and porcine type-C oncovirus. Some of the bacteria were closely associated with humans. For instance, Streptococcus sanguinis and Rothia dentocariosa are known from the human mouth [52,53]. Some studies related Acinetobacter johnsonii, Pseudomonas stutzeri and Acinetobacter junii to clinical settings [54,55], and Aerococcus viridans, Enhydrbacter aerosec, and Micrococcus luteus were thought to be related to air dust [56]. Therefore, our results showed the hospital air-conditioner accumulated a variety of human-related bacteria that can be “detached” from humans in some way. For instance, a microbiologist recently demonstrates the importance of wearing masks to protect health under COVID-19 pandemic [57]. The results showed that speaking, coughing and sneezing could produce tremendous bacteria associated droplets in short time, indicating the likelihood of indoor human exposure to bacterial pollution is high. Another evidence was reported by Kennedy et al. [46], which they directly sampled exhaled aerosols from chronic obstructive pulmonary disease patients and healthy volunteers. The results showed blaTEM was detected on face mask surfaces from both patients and healthy persons; therefore, the authors concluded that patient-generated aerosols may cause antimicrobial resistance dissemination. This “detachment” might be endangering human health in hospital indoor environment since many pathogens were detected in our study (see Supplementary file 2). Notably, another longitudinal analysis of microbial interaction between humans and home indoor environment showed that a new house could be rapidly converged by occupants’ microbial communities, indicating that the hospital indoor environment exhibits high health risk if occupied by contagious patients [6]. From HUMAnN2 results, a clear aerobic lifestyle was observed in M air dust community for both summer and winter. It is interesting to observe some prevalent anaerobic pathways (such as sulfate reduction, homolactic fermentation, and pyrimidine deoxyribonucleotide) in P1, implying that a non-negligible anaerobic bacterial growth happens in air dust. But this anaerobic metabolic pathway disappeared in wintertime, which might explain α diversity results. Our metagenomic data analyses demonstrated that 1) the hospital air dust samples have a variety of ARG carriers and 2) airflow conditions seem to play roles in distributing and accumulating resistome in air dust samples.

3.4. Plasmid-mediated ARG transfer

As can be seen from Bayesian source tracking results (Fig. 2), P2 still harbored summer bacterial species, and the overall impact of M increased greatly from summer (around 20%) to winter (around 50%). A similar trend was also observed regarding ARG source tracking. For instance, more than 80% of winter ARGs from both P and O could be explained by M. Such a huge impact on the overall air dust resistome from M2 indicated heavy resistome exchanges could occur in this hospital frequently. This observation was also following the Procrustes analysis. Notably, Lax et al. claimed that under natural ventilation conditions (e.g., wintertime in our case), a significant reduction of the pathogenic percentage of airborne bacteria could be found [8].

Since we noticed the frequent exchanges, it is necessary to search for the possible media that can initiate ARG transfer. Thus, we constructed an evidence-based ARG genotype sharing network (Fig. 3). Unlike statistical association networks [58], we focused on an evidence-based strategy to establish network edges. Meanwhile, strict cutoffs were applied (e.g., only genotypes with over 99% similarity to a known genotype can be used). As can be seen, the Outpatient hall had the largest number of shared ARG genotypes, and Staphylococcus, Micrococcus, Streptococcus, and Enterococcus were identified. Notably, Staphylococcus carried the largest number of ARGs (14 ARG genotypes), and fewer genera were identified in the Ophthalmology and the Pediatrics departments. Aminoglycoside, bacitracin, β-lactam, chloramphenicol, macrolide, lincomamide and streptogramin (MLS), multi-drug, tetracycline, bleomycin, and sulfonamide resistant genes were commonly detected (Fig. 4 and Supplementary file 5). The most
commonly shared ARG genotypes among different departments mainly included aadD, CE, tetK, tetA, tetZ, and norA. Nanopore sequencing results confirmed the sharing of *Staphylococcus* tetK between different departments. Additionally, plasmid-mediated *CE* transfer was observed in the Outpatient hall between *Staphylococcus* and *Corinebacterium*. The tetK carried by *Staphylococcus* plasmid in M was also observed in the chromosome of *Staphylococcus* and *Corinebacterium* in P and *Methylobacter* in O, indicating plasmid-mediated ARG transfer may be highly possible. Nanopore sequencing helped us successfully assemble 5 bacterial genomes, and some of the long reads were classified as plasmid category (BLASTn against PLSDB, in Fig. S6), suggesting there are active plasmid-related activities. As shown in Table S4, tetracycline, aminoglycoside, macrolide, and β-lactam were all detected from these live cells, and the movement of *Staphylococcus saprophyticus* plasmid among departments might be important for ARG and HAP transfer patterns in this hospital.

### 3.5. Limitations and future directions

Prior studies either applied amplicon sequencing to identify aerosolized microbial communities [8] or used PCR based methods to profile resistome [59]. Some studies applied metagenomic approaches on aerosolized microbial communities in hospitals, but they did not connect ARGs with specific carriers [60,61]. In this study, we integrated a state-of-the-art metagenomic method to link ARGs to specific bacteria and we also confirmed the ARG transfer using cultural isolate whole genome Nanopore sequencing. Even though whole study plan worked (from wet lab to data interpretation), we speculated that our survey (two-timepoints) only demonstrated limited ARG profiling and transfer information, meaning multiple survey years and more sampling frequencies (e.g. weekly or monthly) are needed. Additionally, Illumina sequencing still suffers from fragmented assembly and needs more turnaround time. The short-read sequencing methods were not able to address repetitive insertion sequences, which is often not avoidable in antibiotic resistance gene regions [62,63]. On the other hand, the Nanopore technology has a short turnaround time and much longer sequence read, which can favor studying the structure and genomic context of resistance genes. Therefore, even if Nanopore sequencing has higher sequencing error problems, we still propose to develop Nanopore metagenomic sequencing methods that are suitable in the context of hospital settings in the future.

### 4. Conclusion

We show that the studied hospital air dust samples exhibit complex resistome, and the average concentration is higher than in some other environment (e.g., drinking water, wastewater effluent, and ocean). As a “hotspot”, hospital air-conditioners can accumulate resistome and pathogens overtime, indicating hospital indoor air could serve as one of the transmission routes for bacteria. The Outpatient hall is one of the main ARG transmission sources, and it can affect other departments under different airflow conditions. Inspired by the fact that the frequent cleaning practice reduces the ARGs in surface swab samples, we believe that more frequent cleaning actions towards the Outpatient hall air conditioner should be made to stop ARG spreading in the air in this hospital. Meanwhile, aminoglycoside, bacitracin, β-lactam, chloramphenicol, MLS, multidrug, tetracycline, bleomycin, and sulfonamide resistant genes were commonly detected and the management practice should lean towards ARG carrier genera such as *Staphylococcus, Micrococcus, Streptococcus, and Enterococcus* in this hospital. Our novel evidence-based network strategy proves plasmid-mediated ARG transfer can occur frequently among departments.

### 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.

### Acknowledgment

Our research was funded by Guangdong Province Foundation for Young Scholars (2019A1515110959). Also, we want to thank Center for Computational Science and Engineering at Southern University of Science and Technology (SUSTech) and core research facilities at SUSTech to provide quality resources and services. We thank Dr. Chuanlun Zhang from Department of Ocean Science and Engineering at SUSTech for generously sharing the usage of QuantiStudio qPCR.

### Availability of data

The Illumina metagenomics sequencing raw reads were deposited into EMBL-EBI under the following accession number ER5420782-ER5420789. Cultural isolate Nanopore sequencing reads were deposited under accessions from ER5420826 to ER5420836.

### Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cej.2020.126854.

### References

[1] W.A. Adebeye, The treasure called antibiotics, Ann. Ibadan Postgrad. Med. 14 (2016) 56–57.
[2] I.M. Gould, A.M. Bal, New antibiotic agents in the pipeline and how they can help overcome microbial resistance, Virulence 4 (2013) 185–191, https://doi.org/10.4161/viru.22507.
[3] C.I. Ventola, The antibiotic resistance crisis part 1: causes and threats, P T. 40 (2015) 277–283.
[4] J.L. Ji, J. Cao, Y.-G. Zha, Q.-L. Chen, F. Shen, Y. Wu, S. Xu, H. Fan, G. Du, R. J. Huang, J. Wang, A.L. de Jesus, L. Morawska, C.K. Chan, J. Peccia, M. Yao, Global survey of antibiotic resistance genes in air, Environ. Sci. Technol. 52 (2018) 10975–10984, https://doi.org/10.1021/acs.est.8b02204.
[5] J. Xie, L. Jin, X. Luo, Z. Xiao, X. Li, Seasonal disparities in airborne bacteria and associated antibiotic resistance genes in PM2.5 between urban and rural sites, Environ. Sci. Technol. Lett. 5 (2018) 74–79.
[6] S. Lax, D.P. Smith, J. Hampton-Marcell, S.M. Owens, K.M. Handley, N.M. Scott, S.M. Gibbons, P. Larsen, B.D. Shogan, S. Weiss, J.L. Metcalf, L.K. Ureli, Y. Vazquez-Baeza, W. Van Treuren, N.A. Hasan, M.K. Gibson, R. Colwell, G. Dantas, R. Knight, J.A. Gilbert, Longitudinal analysis of microbial interaction between humans and the indoor environment, Science 345 (2014) 1048–1052, https://doi.org/10.1126/science.1254529.
[7] S. Lax, J.A. Gilbert, Hospital-associated microbiota and implications for nosocomial infections, Trends Mol. Med. 21 (2015) 427–432, https://doi.org/10.1016/j.trends.2015.03.005.
[8] S. Lax, D. Smith, N. Sangwan, K. Handley, P. Larsen, M. Richardson, S. Taylor, E. Landon, J. Alverdy, J. Siegel, R. Stephens, R. Knight, J.A. Gilbert, Colonization and succession of hospital-associated microbiota, Sci. Transl. Med. 9 (2017).
[9] Y. Wiener-Well, M. Galaty, B. Rudensky, Y. Schlesinger, D. Attia, A.M. Yinnon, Nursing and physician attire as possible source of nosocomial infections, Am. J. Infect. Control. 39 (2011) 555–559, https://doi.org/10.1016/j.ajic.2010.12.016.
[10] P.J. Lopez, O. Ron, P. Pachanaarathy, J. Soothill, I. Spitz, Bacterial counts from hospital doctors’ ties are higher than those from shirts, Am. J. Infect. Control. 37 (2009) 79–80, https://doi.org/10.1016/j.ajic.2008.09.018.
[11] A.M. Trevalle, K.A. Thom, J.P. Furuno, S.M. Strauss, A.D. Harris, E.N. Perencevich, Bacterial contamination of health care workers’ white coats, Am. J. Infect. Control. 37 (2009) 101–105, https://doi.org/10.1016/j.ajic.2008.03.009.
[12] M.A. Mariella, C. Pecor, C. Chenoweth, The stethoscope. A potential source of nosocomial infection? Arch. Intern. Med. 157 (1997) 786–790.
[13] R.R.W. Brady, J. Verran, N.N. Damani, A.P. Gibb, Review of mobile communication devices as potential reservoirs of nosocomial pathogens, J. Hosp. Infect. 71 (2009) 290–300, https://doi.org/10.1016/j.jhin.2008.12.009.
[14] F. Ulger, S. Eser, A. Dilek, K. Yanik, M. Gunaydin, H. Leblebicicigil, Are we aware how contaminated our mobile phones with nosocomial pathogens? Ann. Clin. Microbiol. Antimicrob. 8 (2009) 7, https://doi.org/10.1186/1476-0711-8-7.
[15] R.R.W. Brady, A. Wasson, I. Stirling, C. McAllister, N.N. Damani, Is your phone bugged? The incidence of bacteria known to cause nosocomial infection on healthcare workers’ mobile phones, J. Hosp. Infect. 62 (2006) 123–125, https://doi.org/10.1016/j.jhin.2005.05.005.
[16] K.O. Akinyemi, A.D. Atapu, O.O. Adekon, A.O. Coker, The potential role of mobile phones in the spread of bacterial infections, J. Infect. Dev. Ctries. 3 (2009) 628–632.
