Microbiome

RESEARCH

Open Access

House dust microbiome differentiation and phage-mediated antibiotic resistance and virulence dissemination in the presence of endocrine-disrupting chemicals and pharmaceuticals

Shicong Du[1†], Huiju Lin[2†], Qiong Luo[2], Chung Ling Man[3], Sze Han Lai[1], Kin Fai Ho[3], Kenneth M. Y. Leung[2,4] and Patrick K. H. Lee[2,5*]

Abstract

Background House dust serves as a reservoir of a diverse array of microbial life and anthropogenic chemicals, both of which can potentially influence the health of occupants, particularly those who spend significant amounts of time at home. However, the effects of anthropogenic chemicals on dust microbiomes remain poorly understood. This study investigated the presence of anthropogenic chemicals in the dust of homes occupied by elderly occupants and explored those chemicals' relationships with dust microbiomes.

Results We detected 69 out of 76 analyzed anthropogenic chemicals, including endocrine-disrupting chemicals, non-antibiotic pharmaceuticals, and antibiotics, in at least one house dust sample from 32 residential homes, with concentrations ranging from 2720 to 89,300 ng/g. Some of these detected compounds were pharmaceuticals regularly consumed by the occupants. The dust microbiomes were associated with varying levels of anthropogenic chemicals, forming two distinct clusters, each with unique diversity, taxonomy, metabolic functions, and resistome profiles. Higher concentrations and a greater variety of these chemicals were associated with an increased co-occurrence of antibiotic resistance and virulence genes, as well as an enhanced potential for their transfer through mobile genetic elements. Under these conditions, phages, especially phage-plasmids, facilitated the dissemination of antibiotic resistance and virulence among bacterial populations.

Conclusions The findings indicate that everyday anthropogenic chemicals are important factors associated with the microbes in indoor environments. This underscores the importance of improving household chemical stewardship to reduce the health risks associated with exposure to these chemicals and their effects on indoor microbiomes.

Keywords House dust, Endocrine-disrupting chemicals, Non-antibiotic pharmaceuticals, Antibiotics, Microbiomes, Phage-plasmids

†Shicong Du and Huiju Lin contributed equally to this work.

*Correspondence:
Patrick K. H. Lee
patrick.kh.lee@cityu.edu.hk
Full list of author information is available at the end of the article

© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Introduction

House dust harbors a diverse microbiome [1], including bacteria [2, 3], fungi [4, 5], and viruses [6], with both viable and non-viable microbial components [7], along with other materials, most of which originate from occupant skin [8, 9]. The composition of this microbiome is structured by the deposition of particles from indoor air [10], shedding of microbes from humans [2, 8], and household activities [11], with human-associated microbes (e.g., from skin and hair) being a major contributor [8]. While most house dust microbes are not known to be harmful, some fungal allergens [12] and bacterial pathogens [12, 13], including antibiotic-resistant strains of *Staphylococcus aureus* [13], have been detected. Furthermore, a wide repertoire of antibiotic resistance genes (ARGs)—the resistome—has been identified in house dust bacteria, including pathogenic strains [2, 14].

The co-occurrence of ARGs and virulence genes (VGs) has been reported in pathogenic bacteria [15]. In these pathogens, the shorter intergenic distance between mobile genetic elements (MGEs) from ARGs and/or VGs suggests a role for MGEs in the transfer of antibiotic resistance and virulence [16]. The inter-bacterial dissemination of ARGs often occurs through horizontal mechanisms, such as plasmid-mediated conjugation [17] and phage-mediated transduction [18]. Phage-mediated transduction can enhance the transmission of ARGs in human-impacted environments by integrating these genes into bacterial hosts as prophages [19]. Additionally, (pro)phage-plasmids (P-Ps)—genetic elements that can undergo horizontal transfer like phages and vertical transfer within cellular lineages like plasmids [20, 21]—not only carry ARGs but also facilitate the inter-bacterial spread of ARGs [21]. The distinctive biological characteristics of P-Ps make them potentially powerful drivers of bacterial evolution [21].

In addition to harboring a diverse microbiome, house dust contains numerous anthropogenic chemicals [22–24]. These include endocrine-disrupting chemicals (EDCs) that are widely used in foodstuffs, cosmetics, and personal care products [25–27], and pharmaceuticals that may enter indoor settings via excretion from the sweat of occupant skin [28], improper disposal and/or handling of unused drugs [29], and transfer from outdoor air [30]. Although the concentrations of these compounds are generally low, typically from hundreds to thousands of nanograms per gram of house dust [31–33], chronic exposure to some can have adverse health effects including respiratory diseases, neuropsychological disorders, and cancer [34]. Exposure to endocrine-disrupting phthalates is associated with the risk of allergies, particularly through ingestion of contaminated house dust [35].

While the microbial and chemical compositions of house dust have been individually characterized [2, 22, 23, 35–37], the influence of anthropogenic chemicals on the house dust microbiome remains poorly understood. There are known relationships between specific anthropogenic chemicals, such as triclocarban (TCC), triclosan (TCS), and parabens, and microbes in indoor dust across various public and private buildings [14, 38]. However, these studies have largely focused on antimicrobials and ARGs, overlooking other anthropogenic chemicals and microbial constituents like fungi and viruses and the potential mechanisms underlying the spread of ARGs among bacterial populations. To address this knowledge gap, we investigated the dust microbiome in 32 residential homes occupied by elderly individuals who regularly took medications, as well as 76 anthropogenic chemicals commonly found in pharmaceuticals and household products. Our results indicate that the presence of these chemicals was significantly associated with changes in the diversity and composition of the house dust microbiome and the co-location of ARGs and VGs near MGEs in bacteria, especially in human pathogens. Furthermore, these chemicals were associated with enhanced potential for transfer of ARGs and VGs via phages, especially P-Ps. Our findings suggest that everyday anthropogenic chemicals shape the indoor microbiome, highlighting the importance of effective household chemical stewardship.

Results

Diverse anthropogenic chemicals in house dust

Sixty-nine of the 76 analyzed anthropogenic chemicals—19 EDCs, 30 non-antibiotic pharmaceuticals, and 20 antibiotics—were detected in ≥ 1 of the 32 dust samples, with total (∑) concentrations ranging from 2720 to 89,300 ng/g dust. EDCs were the most abundant group, with a median concentration of 4890 ng/g (Fig. 1a). Median concentrations of ∑EDCs were approximately three times higher than those of ∑non-antibiotic pharmaceuticals (1700 ng/g) and approximately 10 times higher than ∑antibiotics (448 ng/g) (Fig. 1a). High detection frequencies (DFs) and elevated concentrations of bisphenols and parabens were observed within the EDC group, with bisphenol A (BPA) being the most abundant compound (DF = 100%, median concentration = 2460 ng/g) (Fig. 1b). Parabens, including methyl paraben (MP), ethyl paraben, butyl paraben, benzyl paraben (BzP), and propyl paraben, were detected in all samples, with median concentrations of 4.1 (BzP) to 308 ng/g (MP). TCC and the alkylphenol 4-nonylphenol were also ubiquitous in house dust (DF = 100%) with median concentrations of 213 and 80.8 ng/g, respectively.

Non-steroidal anti-inflammatory drugs (NSAIDs) were the most abundant subclass within non-antibiotic

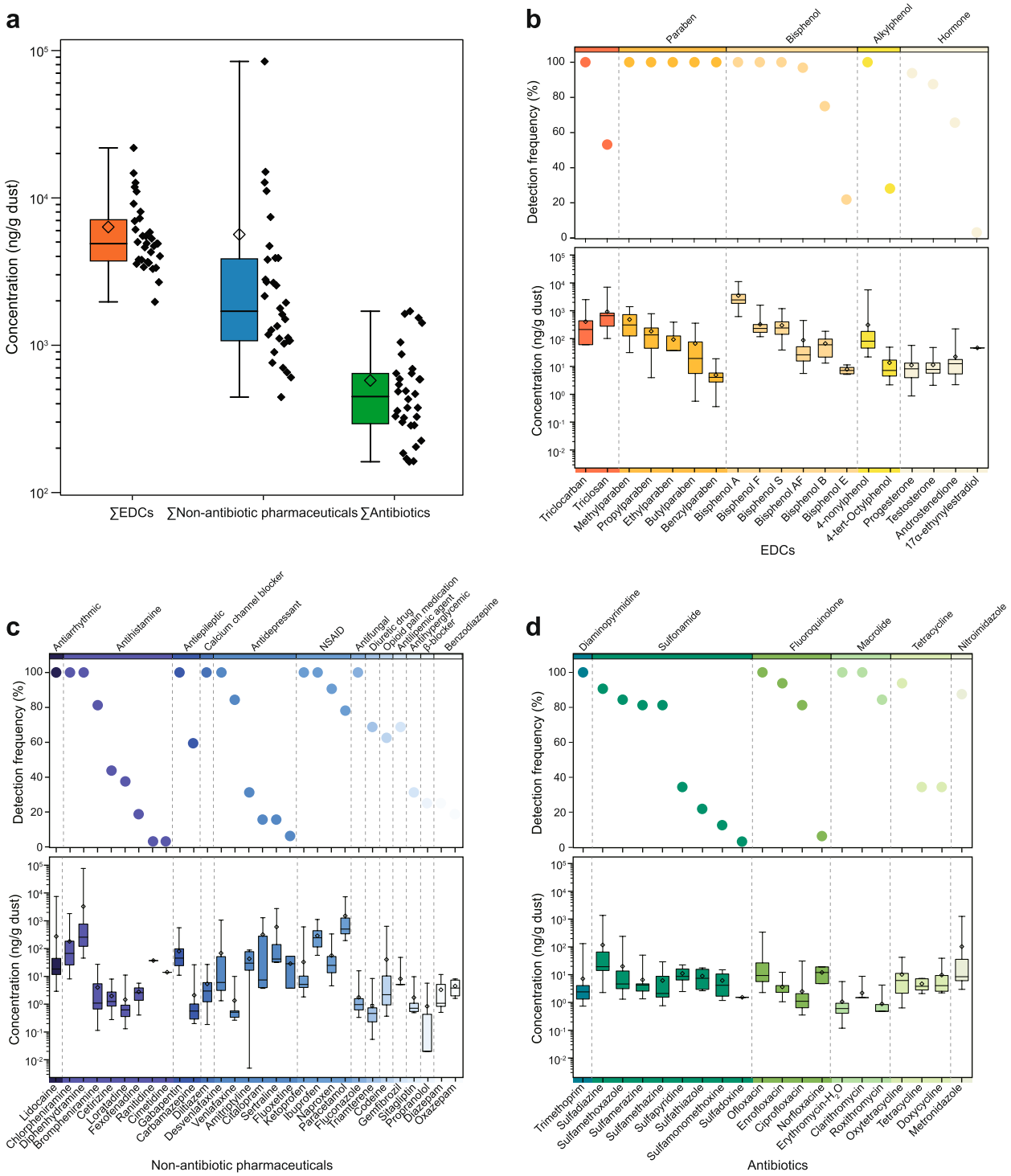

Fig. 1 Concentrations of anthropogenic chemicals in house dust. **a** Concentrations (ng/g dust) of ∑EDCs, ∑non-antibiotic pharmaceuticals, and ∑antibiotics. Each point represents an individual sample. **b**–**d** Detection frequency (%) (top panel) and concentrations (ng/g dust) (bottom panel) for individual (**b**) EDCs, (**c**) non-antibiotic pharmaceuticals, and (**d**) antibiotics. The subclass for each compound is indicated at the top, except for triclocarban and triclosan, which are not assigned to any specific subclass. In each box-and-whisker plot, the box represents the median, first quartile, and third quartile; the whiskers extend 1.5 times the interquartile range; and the diamonds indicate the mean value

pharmaceuticals, dominated by paracetamol (DF = 78%, median = 502 ng/g) (Fig. 1c). Other compounds detected in all samples included the antihistamines diphenhydramine and chlorpheniramine (median concentrations = 257 and 67 ng/g, respectively), the antiepileptic gabapentin (45.3 ng/g), the antiarrhythmic lidocaine (18.4 ng/g), and the calcium-channel blocker diltiazem (3 ng/g). The presence of four non-antibiotic pharmaceuticals consumed routinely by four elderly occupants in their respective households was investigated. Three of these medications were detected in all the corresponding house dust samples at varying concentrations. The benzodiazepine diazepam was found at 1.2 ng/g, the NSAID naproxen at 4.6 ng/g, and the antidepressant sertraline at 2744.3 ng/g. The fourth medication, the antihistamine loratadine, was detected at 11.1 ng/g in one of the two corresponding samples. These four compounds were also detected in households without known consumers, although generally with lower prevalence and concentration. Specifically, sertraline (DF = 15.6%) and loratadine (DF = 37.5%) were significantly less prevalent, with median concentrations 18 to 74 times lower than in households with known consumers. While the concentration of diazepam was similar to that in households with known consumers, it was less frequently found (DF = 25%). However, naproxen, which reduces fever and alleviates pain and inflammation [39], had a DF of 90.6% and a median concentration fivefold higher than in a household with a known consumer.

The analysis of six representative subclasses of antibiotics showed that concentrations were primarily driven by the sulfonamide sulfadiazine, which ranged from 2.3 to 1360 ng/g (median: 19.3 ng/g), followed by the nitroimidazole metronidazole, with concentrations ranging from 3 to 1250 ng/g (median: 8.5 ng/g) (Fig. 1d). Although macrolide antibiotics were observed at lower concentrations than other subclasses, they exhibited higher DFs. For example, the median concentrations of clarithromycin, erythromycin-H_2O, and roxithromycin were 1.5, 0.6, and 0.5 ng/g, with DFs of 100%, 100%, and 84%, respectively. Furthermore, the diaminopyrimidine trimethoprim and the fluoroquinolone ofloxacin were found in all samples, with median concentrations of 2.4 and 9.3 ng/g, respectively.

Correlation between anthropogenic chemicals and dust microbiome clustering

The analysis of 32 house dust samples revealed that on average, 93.9% (SD ± 5.8%) of the reads were annotated as *Bacteria*, while *Fungi* accounted for 5.4% (SD ± 5.5%). Viral and archaeal reads were minor fractions, comprising only 0.14% (SD ± 0.06%) and 0.58% (SD ± 0.96%) of the reads across the samples, respectively. Across the samples, 6940 species were detected, with 1141 of these taxa considered core (present in all samples). These core taxa comprised an average of 77.6% (SD ± 11.2%) of the reads per sample (Fig. S1a).

Prediction strength analysis of the microbiomes in all dust samples indicated the presence of two distinct clusters (1 and 2). Bray–Curtis dissimilarity analysis of the microbiomes further supported this clustering (Fig. 2a), with the two clusters explaining 16.9% of the total community variation (permutational multivariate analysis of variance (PERMANOVA); Table S1). Cluster 1 had a significantly lower Shannon diversity than cluster 2 (Wilcoxon rank-sum test (WRST), $p_{adj} = 0.0305$; Fig. S1b). The microbiomes in cluster 2 were associated with various anthropogenic chemicals, particularly EDCs and antibiotics, according to distance-based redundancy analysis (Fig. 2b). Furthermore, the concentrations of specific compounds within the three major groups of anthropogenic chemicals were correlated with the first principal coordinate axis (PCoA1 in Fig. 2a) of the community composition (Fig. S2a). For example, four of the six analyzed parabens were significantly positively correlated with the community composition ($p < 0.05$) (Fig. S2a). Significant positive correlations ($p < 0.05$) were also observed among the concentrations of these four parabens (Fig. S2b), implying that these EDCs are used concurrently [40].

The median concentrations of ∑EDCs (6090 ng/g), ∑non-antibiotic pharmaceuticals (2780 ng/g), and ∑antibiotics (176 ng/g) in cluster 2 were significantly higher than in cluster 1 (∑EDCs, 4880 ng/g; ∑non-antibiotic pharmaceuticals, 1270 ng/g; ∑antibiotics, 115 ng/g)

(See figure on next page.)

Fig. 2 Differentiation of dust microbiomes into clusters and their association with anthropogenic chemicals. **a** Principal coordinate analysis (PCoA) plot based on the Bray–Curtis dissimilarity of the overall microbiome composition. Samples are colored according to the two clusters. **b** Distance-based redundancy analysis illustrating the associations between endocrine-disrupting chemicals (EDCs) (orange text), non-antibiotic pharmaceuticals (blue text), and antibiotics (green text) to cluster differentiation. Ellipses represent 95% confidence intervals. **c** Association of samples within each cluster with different host and environmental factors, including window opening frequency during winter, age group, marital status, body-mass index (BMI), and dementia. **d** The top 12 genus-level bacterial taxa based on their average relative abundance across all samples, organized by cluster. All other bacterial taxa were grouped under "Others." **e** Relative abundance of species from the genera *Streptomycetes*, *Streptococcus*, *Sphingomonas*, *Shewanella*, and *Pseudomonas* in clusters 1 and 2, as shown in the PCoA plot in panel (**a**)

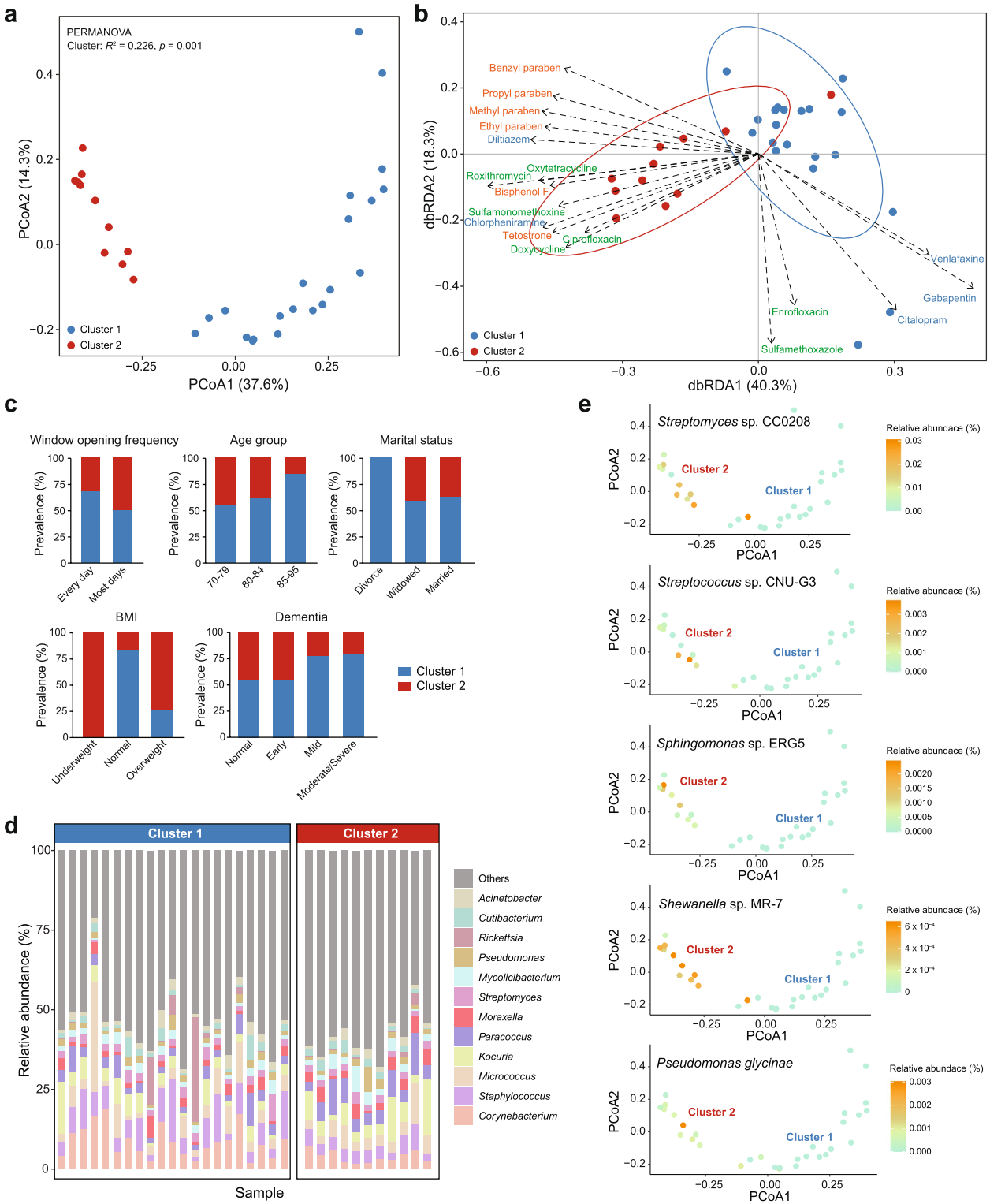

Fig. 2 (See legend on previous page.)

(WRST, $p_{adj} < 0.05$; Fig. S3a). Specifically, significantly higher concentrations were found in cluster 2 for four of the five detected NSAIDs (ketoprofen, naproxen, paracetamol, and ibuprofen), the antihistamines brompheniramine and chlorpheniramine, the paraben MeP, and the antifungal antifluconazole (WRST, $p_{adj} < 0.05$;

Fig. S3b). The elevated levels of anthropogenic chemicals in cluster 2, both in diversity and concentration, suggest that this cluster was primarily associated with greater exposure to these compounds compared with cluster 1. The clustering showed no statistically significant association with the elderly residents' body-mass index (BMI) ($\chi^2 = 4.8$, $p = 0.09$) or any other host or environmental factors, though the BMI association suggested a potential trend (Fig. 2c). Specifically, cluster 2 was more likely associated with residents whose BMI was classified as abnormal (< 18.5 or ≥ 25) [41], while cluster 1 tended to be associated with those whose BMI was classified as normal ($18.5 \leq \text{BMI} \leq 24.9$) [41]. Additionally, no significant differences in Shannon diversity were observed across any host or environmental factors (Fig. S1b). While HEPA filters effectively reduced indoor $PM_{2.5}$ (WRST, $p_{adj} = 0.043$) and BPA (WRST, $p_{adj} = 0.036$) concentrations, they did not significantly affect dust microbiome diversity and composition.

The dominant bacterial genera across all dust samples were *Corynebacterium* (7.1 ± 4.8%), *Staphylococcus* (6.6 ± 4.3%), *Micrococcus* (5.6 ± 6.0%), *Kocuria* (5.3 ± 4.1%), and *Paracoccus* (4.6 ± 2.9%) (Fig. 2d). The relative abundance of *Corynebacterium* was significantly higher in cluster 1 than cluster 2 (WRST, $p_{adj} = 0.003$). Conversely, *Paracoccus* was significantly more abundant in cluster 2 (WRST, $p_{adj} = 0.004$). Several bacterial genera known to be susceptible to antibiotics [42], including *Streptomyces*, *Streptococcus*, *Sphingomonas*, *Shewanella*, and *Pseudomonas*, were enriched in cluster 2, which was exposed to elevated levels of antibiotics and antimicrobial chemicals (Table S2), although they had relatively low abundance (Fig. 2e). *Hortaea werneckii* and a *Preussia* species were the dominant fungal species, comprising 18.4 ± 23.8% and 13.9 ± 16.9% of the fungal community, respectively. *H. werneckii* was more abundant in cluster 1 (27.4 ± 24.5%) than cluster 2 (1.3 ± 3.4%) (WRST, $p_{adj} = 9.8 \times 10^{-4}$), while *Preussia* sp. was more abundant in cluster 2 (23.9 ± 16.5%) than cluster 1 (8.7 ± 14.2%) (WRST, $p_{adj} = 0.022$) (Fig. S1c).

Elevated levels of anthropogenic chemicals enriched potential functions in dust microbiomes

To understand how the taxonomic differences between the clusters translated to variations in functional potentials, the metabolic functions of the microbiomes were examined. The dust microbiomes in clusters 1 and 2 were congruent between their taxonomy and functional pathways (Fig. S4), suggesting that samples with similar taxonomies shared similar functional characteristics. Consistent with the taxonomic analysis, microbiome clustering was the strongest determinant of the functional differences between clusters, rather than participant characteristics, household factors, or HEPA filter usage (Table S1).

The core set of MetaCyc pathways (those present in ≥ 75% of samples in each cluster) was assessed by contributional diversity analysis to compare the pathway diversity between clusters. For both clusters, the core pathways were dominated by those involved in the biosynthesis and metabolism of nucleotides and amino acids (Table S3). Despite this functional conservation, core pathways for cluster 1 predominantly exhibited high within- and between-sample diversity (i.e., "complex variable" contributional diversity) (Fig. 3a). Contrastingly, core pathways in cluster 2 exhibited a wide range of within-sample diversity, from low to high, showing particularly low between-sample diversity (i.e., "simple" and "complex" conserved contributional diversity) (Fig. 3a). The discrepancy between the clusters regarding within- and between-sample variations was most likely due to the 15 uniquely conserved pathways in cluster 2 (Table S3). Of these 15 pathways, five were related to the biosynthesis of homocysteine, cysteine, and their precursor L-methionine, and these pathways were primarily contributed by the human pathogen *Streptococcus pneumoniae* (Fig. 3b). Interestingly, these metabolic pathways have been linked to sulfur metabolism and intracellular redox homeostasis, increasing pathogens' resistance to various antibiotics subclasses, including β-lactams, tetracyclines, quinolones, and aminoglycosides [43].

Multivariate analysis revealed significant enrichment of KEGG Orthology (KO) gene families between the clusters, with 1228 KOs enriched in cluster 2, compared with 251 in cluster 1 (Table S2). Similarly, 428 MetaCyc pathways, covering various aspects of microbial physiology, were enriched exclusively in cluster 2, with none enriched in cluster 1 (Table S2). Notably, cluster 2 was characterized by more numerous pathways with high relative abundance related to nucleoside and nucleotide biosynthesis, as well as amino acid biosynthesis, especially in the synthesis of essential branched-chain amino acids (L-isoleucine, L-leucine, and L-lysine) (Fig. 3c). In addition, pathways involved in aromatic-compound degradation, such as the aerobic benzoyl-CoA degradation (PWY-1361) that is related to the degradation of ibuprofen [44], were enriched in cluster 2 (Table S2).

The analysis of the ARG repertoire in dust samples revealed that the resistomes differentiated into two distinct clusters, consistent with the clustering of the microbiomes (Fig. S5a and Table S1). The multivariate analysis of the 268 detected ARGs showed the enrichment of 23 ARGs, predominantly conferring resistance to aminoglycosides ($n = 6$), tetracyclines ($n = 3$), and β-lactams ($n = 3$) in cluster 2, while no ARGs were enriched in cluster 1 (Fig. S5b and Table S2). The enrichment of ARGs conferring resistance

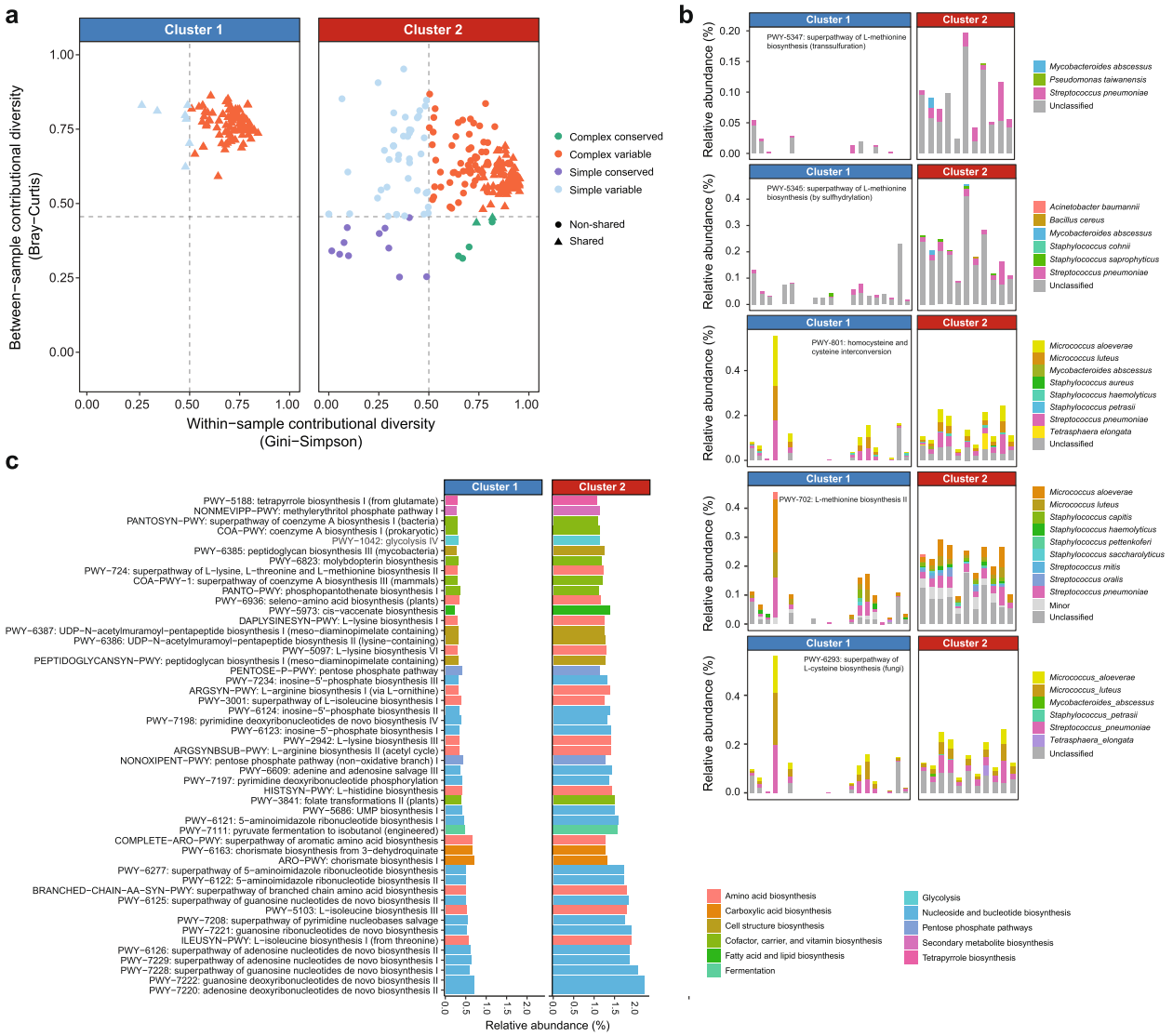

Fig. 3 Differences in contributional functional diversity and enriched functional pathways between clusters 1 and 2. **a** Differences in contributional diversity between clusters for core MetaCyc pathways (pathways present in ≥ 75% of samples within each cluster). Core pathways detected in both clusters are labeled as "shared" (triangles). **b** Examples of conserved pathways (biosynthesis of homocysteine, cysteine, and L-methionine) illustrating differences in contributional diversity between clusters. **c** Enrichment of MetaCyc pathways between clusters analyzed using MaAslin2. Only the top 50 statistically significant pathways (adjusted $p < 0.05$) are shown

to tetracyclines in cluster 2 was consistent with the analytical detection of three tetracycline-class antibiotics (tetracycline, doxycycline, and oxytetracycline) in samples associated with that cluster at average DFs of 36 to 100% and average concentrations of 2.0 to 15.1 ng/g (Fig. S3c).

Association of elevated anthropogenic chemical levels with co-occurrence among ARGs, VGs, and MGEs in dust microbiomes

The presence, distribution, and co-occurrence of ARGs, VGs, and MGEs on the contigs were investigated. In total, 2.7 million non-redundant open reading frames (ORFs) were predicted from 1.9 million contigs (≥ 1000 bp) assembled from 32 dust samples. The distribution of ARGs identified from the ORFs (Fig. S6a) was consistent with the two distinct resistome clusters observed based on the short-read analysis (Fig. S5a). Specifically, cluster 2 exhibited a greater number ($n = 289$) and a higher average relative abundance (0.036%) of ARGs than cluster 1 ($n = 29$ and average relative abundance of 0.0059%) (Fig. S6a). The spread of ARGs is often facilitated by MGEs and linked

to VGs associated with pathogenic bacteria [21]. Three hundred thirty-eight ORFs were classified into 217 VG subtypes (Fig. S6b), with VGs being more abundant and predominant in cluster 2 ($n=310$) than cluster 1 ($n=28$) (Fig. S6b and S6c). Meanwhile, 15,698 MGEs were identified, with more in cluster 2 ($n=13,685$) than cluster 1 ($n=2013$). These MGEs encompassed proteins involved in various functional processes, including integration/excision (IE; e.g., *int*), replication/recombination/repair (RRR; e.g., *repA*), transfer (e.g., *groL*), phage (e.g., *clpB*), and stability/transfer/defense (STD; e.g., *vapC*) (Fig. S6d).

The co-occurrence of ARGs, VGs, and MGEs within the house dust microbiomes in the presence of anthropogenic chemicals was investigated by co-occurrence network analysis (Fig. S7a). The network analysis, irrespective of cluster, revealed strong and significant correlations (Spearman's $|\rho|>0.8$, $p<0.001$) among 31 ARGs, 19 VGs, and different MGE categories, including IE ($n=474$), RRR ($n=161$), STD ($n=89$), phage ($n=40$), and transfer ($n=19$) (Fig. S7b). Seven of the 31 co-occurring ARGs conferred resistance to tetracyclines (*tap*, *tetO*, *tetP*, *tetW*, *tet39*, *tetA*, and *tetA(P)*), while most others were associated with multidrug efflux pumps. Further analysis of the ARG-MGE and VG-MGE connections showed that IEs were the most frequently linked ($n=231$), while phages had the fewest connections ($n=10$) (Fig. S7b). Nevertheless, phage-associated proteins played a distinct role in the network, exhibiting a higher degree of nodes and eigenvector centrality (Fig. S7c), suggesting a central and influential position within the overall co-occurrence network. Furthermore, phage-associated proteins showed the second-highest vulnerability within the network, only lower than IE proteins (Fig. S7d).

Comparing the respective co-occurrence networks of ARGs, VGs, and MGEs of clusters 1 and 2 (Fig. 4a), the network of cluster 2 exhibited significantly higher degree, betweenness centrality, and clustering density, as well as significantly lower closeness centrality (WRST, $p_{adj}=2.2\times10^{-16}$) (Fig. 4b). Consistently, the network of cluster 2 showed higher stability when subject to random attack (Fig. 4c). Differences in the co-occurring ARGs conferring resistance to tetracyclines were also observed, with four (*tetW*, *tet39*, *tetA*, and *tetA(P)*) shared and three (*tap*, *tetO*, and *tetP*) unique to cluster 2 (Fig. 4a). The variations in co-occurrence patterns between the clusters suggest a potentially stronger co-selection of ARGs, VGs, and MGEs that were associated with elevated concentrations of anthropogenic chemicals in cluster 2.

Association of elevated anthropogenic chemical levels with the co-localization of ARGs, VGs, and MGEs within genomes

The co-localization patterns of ARGs, VGs, and MGEs within the representative metagenome-assembled genomes (rMAGs) reconstructed from the samples were investigated to assess whether anthropogenic chemicals exerted selective pressure facilitating the transfer of these genetic elements among members of the dust communities. Of the 106 rMAGs identified, 75 were classified as medium-quality and 31 as high-quality genomes, with 12 identified as human pathogens (Table S4). Among these rMAGs, only three were enriched in cluster 1, two of which were *Corynebacterium* spp. In contrast, cluster 2 contained 46 enriched rMAGs, including 23 known species, six of which were human pathogens, such as *Stenotrophomonas maltophilia* [45] and *Acinetobacter johnsonii* [46] (Fig. S8a and Table S4). ARGs were detected in 38 rMAGs derived from 13 samples, including eight samples from cluster 2 (Fig. S8a and S8b). More than half of these ARGs confer multidrug resistance. Twenty-two rMAGs carried ARGs conferring resistance to antibiotics detected in their corresponding dust samples, including diaminopyrimidines, sulfonamides, fluoroquinolones, macrolides, and tetracyclines, with concentrations of 2–331.1 ng/g (Fig. S8b). To investigate potential resistance to additional antibiotics, we focused on six rMAGs carrying β-lactam resistance genes. We measured 14 antibiotics from five representative β-lactam subclasses—carbapenem ($n=2$), cephalosporin ($n=4$), cephamycin ($n=1$), monobactam ($n=1$), and penicillin ($n=6$)—in the four corresponding dust samples from which the six rMAGs were reconstructed. While no antibiotics from the cephalosporin, cephamycin, or monobactam subclasses were detectable, three from the penicillin subclass (penicillinV, amoxicillin, and piperacillin) and two from the carbapenem subclass (imipenem and metropenem) were found, with penicillinV and amoxicillin detected in all four samples at concentrations of 1.8–59 ng/g (Fig. S8c).

ARGs, VGs, and MGEs were co-localized within 26 rMAGs (Fig. S8a), 22 of which had significantly higher abundance in cluster 2 than cluster 1 (WRST, $p_{adj}<0.05$), including six human pathogens, with three of these pathogens and six additional rMAGs enriched in cluster 2 (Fig. 5a and Table S4). The numbers of ARGs and VGs detected in each rMAG were significantly correlated (Pearson's $r=0.69$, $p=5.2\times10^{-6}$), further supporting the co-selection of these genetic elements within the dust microbiomes due to elevated concentrations of anthropogenic chemicals. To evaluate the potential transfer of ARGs and VGs between genomes, the nearest distances from MGEs to these genes within the contigs

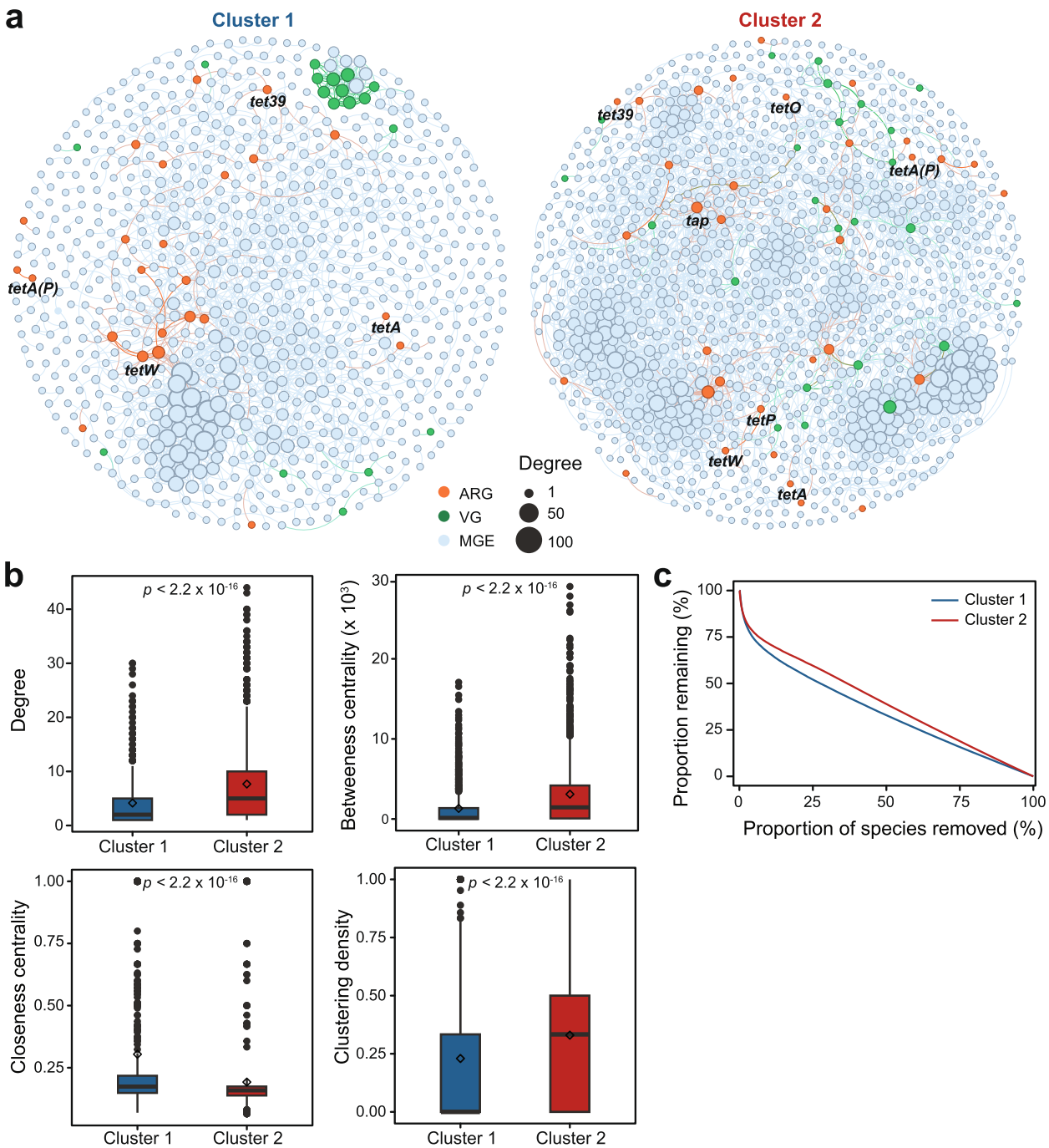

Fig. 4 Co-occurrence patterns of ARGs, VGs, and MGEs in the dust microbiomes of clusters 1 and 2. **a** Network analysis illustrating the correlations between ARGs, VGs, and MGEs in each cluster. Only connections exhibiting strong (Spearman's $|\rho| > 0.8$) and significant ($p < 0.001$) correlation are represented in the network. Node size is proportional to the number of connections (the degree). **b** Comparison of node-level topological features between subcommunities in the two clusters. Differences between clusters 1 and 2 were assessed using the Wilcoxon rank-sum test. In each box-and-whisker plot, the box represents the median, first quartile, and third quartile; the whiskers extend 1.5 times the interquartile range; and the diamonds indicate the mean value. Points beyond the whiskers are considered outliers. **c** The robustness of the two networks was assessed by evaluating the proportion of taxa remained as taxa were randomly removed from each network

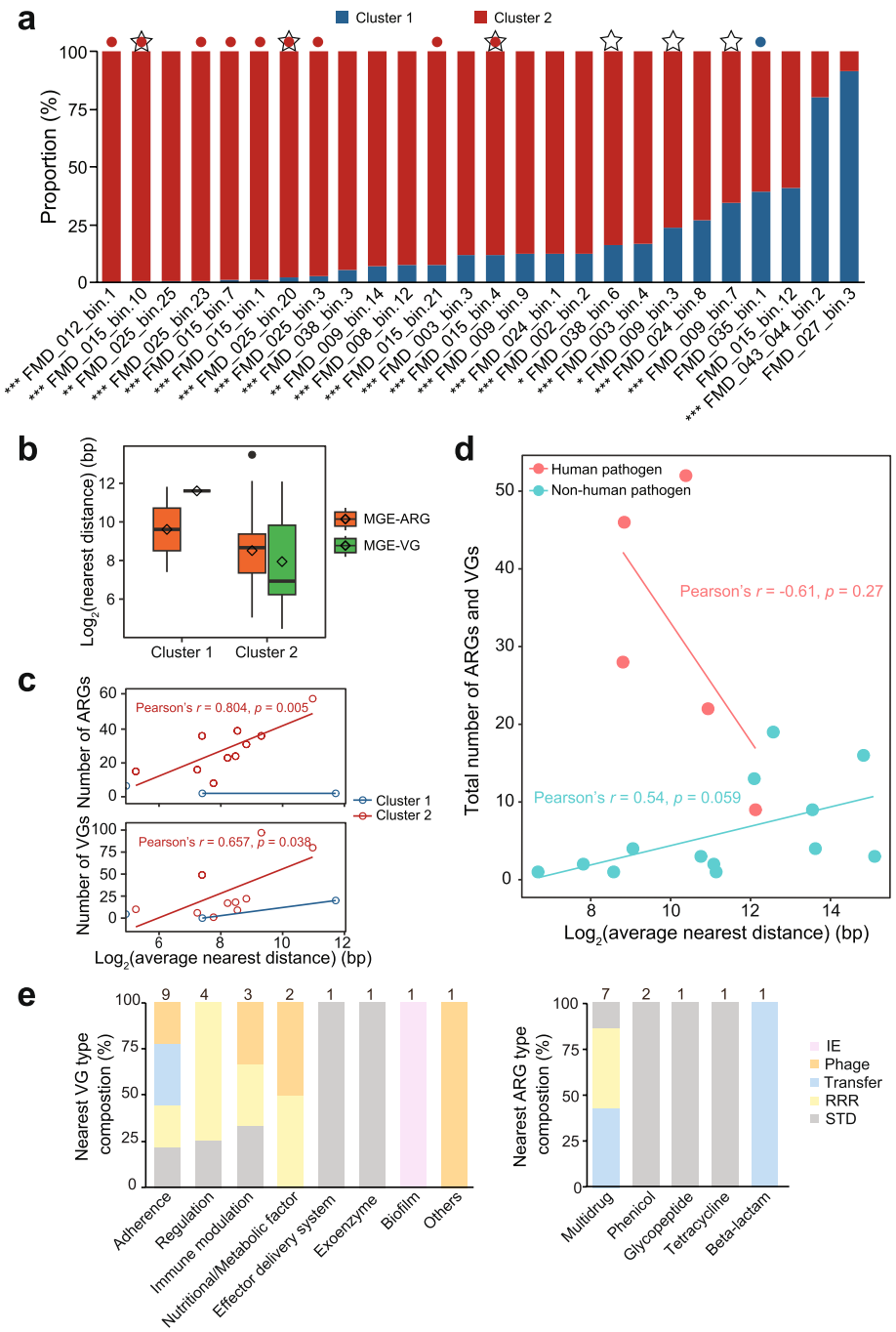

Fig. 5 Co-localization of ARGs, VGs, and MGEs in representative metagenome-assembled genomes (rMAGs). **a** Proportion of the relative abundance of rMAGs containing ARGs, VGs, and MGEs in clusters 1 and 2. The colored circle above each bar plot indicates the enrichment of rMAGs in the corresponding cluster. Statistically significant differences were assessed using the Wilcoxon rank-sum test. * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$. **b** The nearest distances of ARGs or VGs from MGEs in each cluster. In each box-and-whisker plot, the box represents the median, first quartile, and third quartile; the whiskers extend 1.5 times the interquartile range; and the diamonds indicate the mean value. **c** Correlations between the total number of ARGs (top) and VGs (bottom) and the average nearest distance of ARGs or VGs from MGEs. **d** Correlations between the total number of ARGs and VGs and the average nearest distance of ARGs or VGs from MGEs across human and non-human pathogens. **e** Composition of the nearest VGs (left) and ARGs (right) across five major categories of MGEs in the rMAGs. Numbers above the bars indicate the total count of VGs (left) and ARGs (right)

were assessed. The average nearest distances from MGEs to ARGs or VGs were shorter in cluster 2 than cluster 1 (Fig. 5b). Furthermore, these distances showed significant positive correlations with the number of ARGs (Pearson's $r=0.804$, $p=0.005$) and VGs (Pearson's $r=0.657$, $p=0.038$) in cluster 2 (Fig. 5c). Regardless of cluster, for the rMAGs with co-localization of MGEs with ARGs and VGs, the average nearest distances from MGEs to ARGs or VGs were marginally significantly positively correlated with the total number of ARGs and VGs in non-pathogenic taxa (Pearson's $r=0.52$, $p=0.059$). Conversely, a negative trend was observed in human pathogenic taxa, although not statistically significant (Pearson's $r=-0.61$, $p=0.27$), likely due to the small number of genomes (Fig. 5d). Investigation of the transfer vehicles revealed that ARGs, primarily encoding multidrug efflux pumps, were predominantly transferred via MGEs with a transfer life cycle, with VGs mainly transferred by phage-related MGEs (Fig. 5e).

Association of elevated anthropogenic chemical levels with increased prevalence of phages with ARGs and/or VGs and their insertion into hosts

In total, 1108 viral operational taxonomic units (vOTUs) were identified (Fig. S9a), with phages classified based on putative links between these vOTUs and their in situ and ex situ bacterial hosts [47] (Fig. S9b, S9c, and Table S5). This analysis identified 530 phages, including 64 prophages and 94 putative P-Ps, with two of the P-Ps also classified as prophages. The putative P-Ps were clustered with 24 plasmid types, the dominant types being P1, pSLy3, and AB (Fig. S10a). Twenty-nine phages (including one putative P-P and one prophage) were enriched in cluster 1, while 70 phages (including 26 putative P-Ps and three prophages) were enriched in cluster 2 (Table S2). Among the seven high-confidence P-Ps, six showed a significantly higher relative abundance in cluster 2 than cluster 1 (WRST, $p_{adj}<0.05$; Fig. S10b). Significantly more ARGs and VGs were identified in the phages enriched in cluster 2, especially in P-Ps, where all five ARGs and 9 of the 11 VGs were detected. In contrast, no ARGs or VGs were detected in the phages enriched in cluster 1.

The role of phage-mediated horizontal gene transfer (HGT) of ARGs and VGs among bacteria in the two clusters was investigated. All three rMAGs enriched in cluster 1, along with 41 of the 46 rMAGs enriched in cluster 2, were linked to phages, including both P-Ps and prophages. Notably, three pathogenic rMAGs enriched in cluster 2 and one in cluster 1 were predominantly infected by P-Ps (Fig. S11a). Of the 44 identified links between the enriched phages and hosts in cluster 2, P-Ps accounted for 80%, other phages 16%, and prophages 4% (Fig. 6a). Many of these P-Ps carried ARGs and/or VGs and were linked to non-pathogenic rMAGs, although some were also linked to pathogenic ex situ hosts (e.g., *P. aeruginosa* [48]). Interestingly, P-P S7C825, linked to the rMAG FMD_006_bin.6 (classified as *Paracoccus marinus* and known to exhibit antibiotic susceptibility [49]), carries the same ARG *ugd*, which confers resistance to antimicrobial peptides. However, based on nucleotide matches, there is no evidence that this P-P specifically inserted *ugd* into the host. In contrast, only a single link was identified between the enriched phages and hosts in cluster 1.

Although the P-Ps enriched in cluster 2 were not found to have inserted ARGs into their linked hosts, we identified one prophage (S11C70841_1) and two putative P-Ps (S11C610 and S16C4026) in three samples from cluster 2 that had inserted ARGs, such as *erm*(46), *arr*−1, and *golS*, into their respective linked hosts (Fig. 6b). Furthermore, a putative P-P (S11C2677) and a high-confidence P-P (S11C3896), present in a single sample from cluster 2, were identified as having integrated into a host chromosome as prophages (Fig. S11a). Specifically, P-P S11C3896 was identified within the chromosome of the pathogenic rMAG FMD_015_bin.10, where several ARGs, VGs, and MGEs co-localized within the genome (Fig. S11b). This P-P is phylogenetically closely related to *Pseudomonas* phage Y1 (OQ572403.1), which infects the opportunistic pathogen *P. aeruginosa* [48] (Fig. 6c). Although the host bacterium FMD_015_bin.10 was classified as *P. mendocina*, this species is also phylogenetically related to *P. aeruginosa* (Fig. 6d).

Discussion

Dust is ubiquitous indoors and a significant reservoir of diverse anthropogenic chemicals and microbes [3]. Studies of the relationships between anthropogenic chemicals and dust microbes in public and private buildings have primarily focused on antimicrobial chemicals and ARGs, overlooking other anthropogenic chemicals, fungi, viruses, and the mechanisms of ARG dissemination [14, 38]. In this study, we analyzed 76 anthropogenic chemicals, including EDCs, non-antibiotic pharmaceuticals, and antibiotics, in 32 dust samples, while characterizing the corresponding metagenomes. We found that both the prevalence and concentration of these chemicals in dust were key factors influencing microbiome differentiation into two clusters, each with unique diversity, taxonomic composition, functional profiles, and resistomes. In the presence of varying levels of anthropogenic chemicals, we observed a distinct co-occurrence of ARGs, VGs, and MGEs, characterized by higher potential for co-localization within genomes. Moreover, evidence suggests that phages, particularly P-Ps, enhance the potential for transfer of ARGs and VGs between bacterial populations.

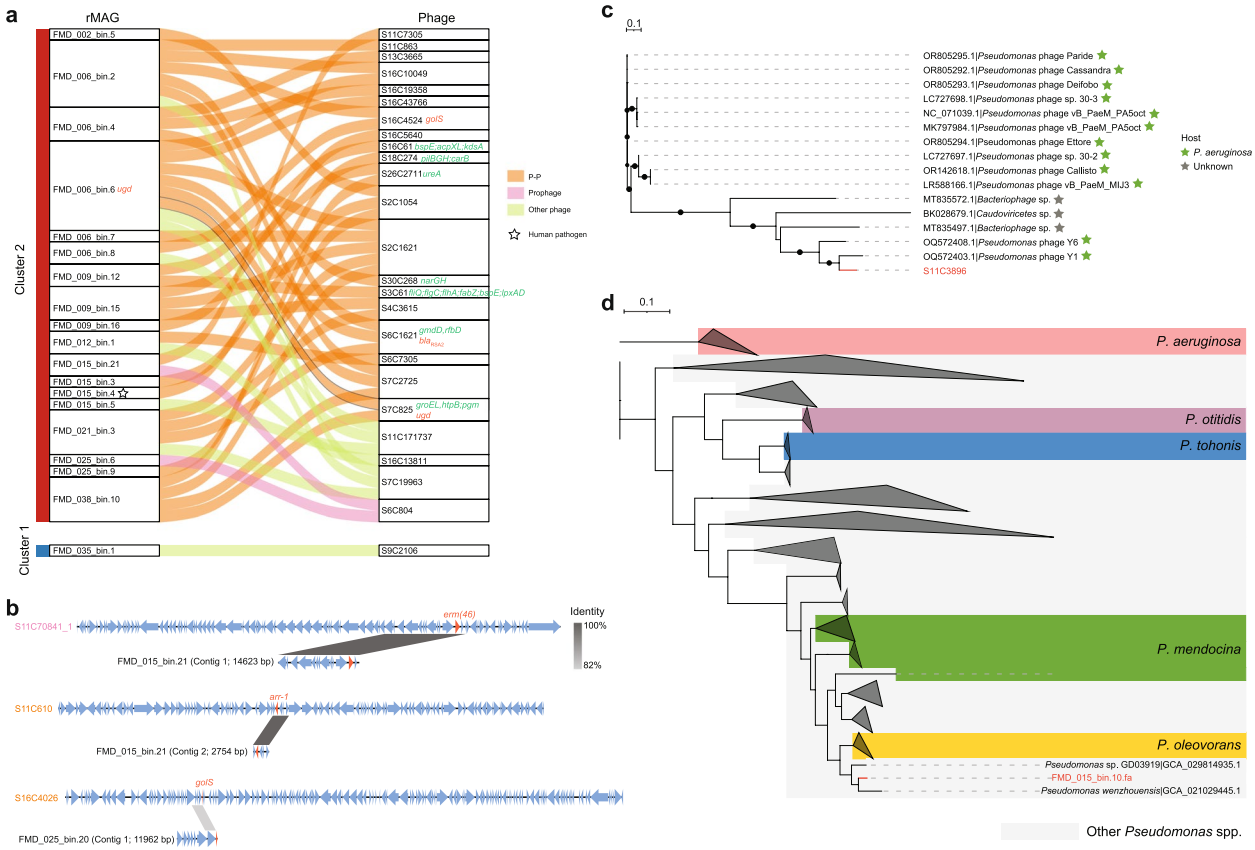

Fig. 6 Phage-mediated dissemination of ARGs and VGs. **a** The Sankey plot illustrates the enriched links between in situ hosts and phages, including prophages and phage-plasmids (P-Ps). ARGs and/or VGs carried by phages are highlighted in orange and green text, respectively (right column). The only shared ARG between the in situ host and its linked phage is highlighted (left column). **b** A schematic depicting the insertion of ARGs by one prophage (top panel) and two putative P-Ps (bottom two panels) into their respective host genomes. **c** A maximum-likelihood phylogenetic tree of P-P S11C3896 from this study, alongside 15 phages from the NCBI Virus database. The branch representing this P-P is shown in red, with the colored stars indicating the hosts of these 15 phages. Black circles on the trees represent bootstrap values > 50%. **d** A maximum-likelihood phylogenetic tree of the rMAG FMD_015_bin.10 from this study and 1765 non-redundant complete strains of the genus *Pseudomonas* from the NCBI GenBank database. The branch representing this rMAG is shown in red. Black circles on the tree indicate bootstrap values > 50%

Among the detected anthropogenic chemicals, EDCs showed a significantly higher median concentration than non-antibiotic pharmaceuticals and antibiotics. The identified EDCs comprise a diverse group, including hormones, some of which are used for oral supplementation, as well as bisphenols and antimicrobial chemicals such as parabens, TCC, and TCS, which are commonly found in plastics, pharmaceuticals, and personal care products (PCPs) [50]. Although the use of TCC and TCS in PCPs (e.g., detergents, toothpastes, and deodorants) has been restricted in recent years due to their acute and chronic toxicity and environmental persistence [50], considerably high concentrations (from a few dozen to thousands ng/g) can still be detected in indoor dust across North America [14, 38], Asia [17, 51], and Europe [24]. The present study also detected high concentrations and prevalence of TCC (100% DF) and TCS (53% DF) in the sampled houses, with median concentrations of 10^2–10^4 ng/g dust, which was approximately tenfold higher than previously detected in public facilities (e.g., offices, gyms, and classrooms) [14, 38] and private homes [24]. Bisphenols (20–100% DF) used in plastic products and parabens (100% DF) used as preservatives in PCPs and pharmaceuticals were other frequently detected EDCs. Multiple bisphenols and parabens were often co-detected, suggesting their ubiquitous presence and accumulation in indoor environments. Specifically, MP was the most abundant paraben determined in this study with a median concentration (308 ng/g) comparable to those in indoor dust from China ($n = 52$; median = 320 ng/g), but approximately two- to fivefold lower than those collected from the US ($n = 40$; median = 760 ng/g), Japan ($n = 22$; median = 1470 ng/g), and South Korea ($n = 41$; median = 1310 ng/g) [32]. Additionally, the median concentration of BPA in this study

(2460 ng/g) was comparable to the levels found in indoor dust from Japan ($n = 22$; median = 2700 ng/g), but lower than the concentrations in dust samples from South Korea ($n = 41$; median = 3260 ng/g) [52]. The variations in the concentrations of these anthropogenic chemicals may reflect differences in the use of specific PCPs or pharmaceuticals across populations. A meta-analysis showed a positive association between urinary and serum BPA levels and the risk of type 2 diabetes mellitus [53]. However, the present study did not find significant differences in BPA concentrations in house dust between participants with or without a diagnosis of diabetes mellitus. NSAIDs, including paracetamol and ibuprofen, were among the most abundant and prevalent pharmaceuticals detected in a comprehensive study of European house dust [24], with concentrations comparable to those found in our study. Four non-antibiotic pharmaceuticals consumed by elderly occupants were detected in their house dust, with two found at higher concentrations than in dust from households without known consumers, suggesting that these compounds originate from medication use and accumulate in the surrounding environment.

In addition to EDCs, the widespread use of antibiotics has undesirable consequences, primarily by promoting antibiotic resistance and facilitating the transfer of resistance among bacterial populations [54]. Research also indicates that microbial secondary metabolites contribute antibiotics to house dust [55, 56]. Our study detected many common antibiotics and identified a diverse array of bacterial taxa with potential antibiotic resistance enriched in cluster 2. Previous studies have also reported that certain antimicrobials (TCC, TCS, and parabens) are associated with the increased presence of ARGs in bacteria [14, 38]. Moreover, growing evidence suggests that exposure to non-antibiotic pharmaceuticals may also contribute to antibiotic resistance and potentially increase HGT rates among bacteria [57]. Consistent with these findings, we observed a higher proportion of ARGs encoding multidrug efflux pumps, which exhibited an increased potential for transfer in samples exposed to elevated levels of anthropogenic chemicals, particularly non-antibiotic pharmaceuticals such as antihistamines (e.g., fexofenadine and lidocaine) and NSAIDs (e.g., ibuprofen). Although not directly tested in our study, exposure to these non-antibiotic pharmaceuticals may increase the expression of efflux systems, potentially increasing resistance to multiple antibiotics and antimicrobial chemicals.

The presence of ARGs in bacteria is often accompanied by VGs due to co-selection under selective pressure, especially in pathogenic strains [58]. A recent study reported the co-occurrence of ARGs and VGs in indoor dust [36]; however, the specific role of anthropogenic chemicals in the co-selection of these genetic elements remains unclear. Our study shows a significant positive correlation between ARGs and VGs, identifying a co-occurrence pattern of these two types of genes in cluster 2. This finding supports the notion that selective pressures from high levels of anthropogenic chemicals promote the co-selection of both antibiotic resistance and virulence in bacteria. These resistant, virulent bacteria may be opportunistic pathogens, posing an infection risk for immunocompromised individuals [58]. Consistently, in cluster 2, we observed a higher prevalence of human pathogens that carry both ARGs and VGs, suggesting that house dust is both a reservoir and a vehicle for the dissemination of these resistant, virulent bacteria.

ARGs and VGs can disseminate widely through HGT, facilitated by MGEs, potentially accelerating bacterial evolution under environmental stress [59]. In cluster 2, the distances between MGEs and ARGs or VGs within rMAGs were shorter than in cluster 1. Further investigation into human and non-pathogens showed contrasting patterns. In human pathogens, as the number of ARGs and VGs increased, the distance between these genes and MGEs decreased. Conversely, in non-pathogens, shorter distances between MGEs and ARGs or VGs were significantly correlated with lower numbers of these genes. Notably, most co-localized ARGs conferred multidrug resistance, suggesting that pathogenic and non-pathogenic bacteria selectively acquire functions for adaptation to environments with diverse antibiotics. ARGs and VGs located near MGEs in pathogens or rMAGs exposed to high chemical stress in the dust samples are likely to have increased functional mobility. This gene arrangement is similar to those found in pathogenic bacteria from hospital wastewater treatment systems, where pharmaceutical concentrations are presumably high [60, 61]. These findings suggest that bacteria under environmental stress are more prone to transferring ARGs and VGs within microbial communities, contributing to the spread of antibiotic resistance and virulence in bacterial populations.

Phage-mediated lateral transduction can be more efficient than the transfer of classical MGEs via conjugation or generalized transduction [62], particularly in P-Ps [20]. Some phages, particularly P-Ps, preferentially transfer ARGs and VGs between bacterial populations, including pathogens, within dust samples exposed to elevated levels of anthropogenic chemicals. The spread of antibiotic resistance by P-Ps, which has been experimentally confirmed [63], is particularly concerning, as it does not require direct cell-to-cell contact and P-P integrons serve as genetic platforms facilitating the acquisition of novel ARGs [20, 21]. Despite their potential role in antibiotic resistance, P-Ps have been documented in only a few studies, including those on glaciers [64],

marine environments [65], and the mammalian gut [66]. Although not all P-Ps carry ARGs or VGs, their strategies to infect hosts increase their potential for transfer. For example, the conversion of P-Ps into prophages that integrate into bacterial chromosomes was observed in a sample with elevated levels of anthropogenic chemicals. One possible reason for the conversion from P-Ps to prophages is that integration into the bacterial chromosome enables P-Ps to evade host defenses that specifically target circular MGEs [67]. Additionally, integrated P-Ps may replicate with the bacterial chromosome and no longer require plasmid replicases and segregation systems [21]. Thus, the effective transfer capabilities of P-Ps may facilitate the spread of antibiotic resistance and virulence among bacteria, including pathogens.

This study highlights the interplay between anthropogenic chemicals and microbiomes in residential dust. However, it has limitations. First, while a number of anthropogenic chemicals were measured, the study did not include certain common EDCs (e.g., phthalates and flame retardants), medications commonly used by the participants (e.g., metformin and aspirin), or other dust-related components (e.g., mites and skin-derived material). Additionally, because household dust gathers material from multiple sources, such as surfaces, the origins of chemical-microbiological interactions remain uncertain and may begin elsewhere before accumulating and persisting in dust. Second, although we observed that distinct fungal taxonomic compositions were associated with these anthropogenic chemicals, the specific roles of fungi in the interconnected systems between chemicals and human health remain underexplored. Third, metagenomic sequencing did not elucidate bacterial viability or the absolute abundance and functionality of key genes (e.g., ARGs and VGs) and elements (e.g., prophages and P-Ps). Future research should include laboratory experiments on the functionality of ARGs and VGs, identify phages that facilitate the transfer of functional genes between bacteria, and use quantitative PCR for absolute gene quantification and metatranscriptomic sequencing for gene expression analysis. Last, recent studies suggest that indoor chemicals and their metabolites in dust have a greater impact on occupant health—particularly in relation to asthma and allergic rhinitis—than the dust microbiome itself [68, 69]. Our findings indicate that exposure to these chemicals can alter the dust microbiome composition, increasing pathogen presence and facilitating phage-mediated transmission of antibiotic resistance and virulence, which may have significant health implications. To establish a causal link between anthropogenic chemicals, dust microbiome composition, and health outcomes, future research should conduct longitudinal cohort studies across diverse populations and household settings, integrating environmental monitoring, microbiome analysis, and health assessments.

Conclusions

Our study highlights the role of house dust as a reservoir for numerous EDCs, non-antibiotic pharmaceuticals, and antibiotics. These anthropogenic chemicals are associated with the diversity and composition of microbiomes and the spread of antibiotic resistance and virulence among bacterial populations. Additionally, our findings highlight the often-overlooked roles of phages in the dissemination of antibiotic resistance and virulence, posing increased health risks to vulnerable occupants, such as children and the elderly. As most people spend significant periods indoors, long-term exposure to ARGs and VGs accumulated in dust raises concerns about their potential impact on occupants' microbiomes, particularly those of the airway [70] and gut [71]. The diversity of anthropogenic chemicals in house dust likely far exceeds our measurements, suggesting that numerous substances from everyday products, including household items and pharmaceuticals, accumulate in dust and influence its microbiology. Therefore, there is an urgent need for increased awareness and improved chemical stewardship to reduce exposure to these chemicals and associated microbes in indoor environments.

Materials and methods

Subject characteristics and sample collection

The study was conducted across 32 residential homes in urban Hong Kong, each with an average of 2 residents and 1–2 elderly individuals aged ≥ 72. On average, the elderly participants had lived in their homes for 23 years and spent ≥ 20 h a day there. The elderly residents underwent comprehensive clinical assessments at a local hospital, which included evaluations of common age-related health conditions such as diabetes mellitus, hypertension, and dementia. All but one of the elderly individuals regularly consumed medication during the study period. No non-elderly residents required long-term medication. The elderly residents' BMI was calculated based on the provided height and weight data using the formula BMI = weight (kg)/height^2 (m^2). Detailed personal information and daily living habitats of the elderly participants are presented in Table S6.

Household characteristics, including house size, frequency of opening windows, and frequency of air conditioner use, were also collected (Table S6). Each household had an air-purifier (#LA352, LIFAair, Helsinki, Finland) containing four activated-carbon plates (#LA31, LIFAair) installed in the living room. The device had sensors to monitor the airborne levels of carbon dioxide, formaldehyde, and particulate matter ($PM_{2.5}$), and could

automatically adjust the airflow rate to maintain optimal air quality according to its default settings. To investigate the impact of additional HEPA filters (#LA21, LIFAair) on air quality and microbes, they were randomly installed in the air-purifiers of 20 households, while the remaining 12 households served as a control group. The air-purifiers had been in use for ~ 1 year prior to dust sampling. A dust sample was collected from each of the 32 households by brushing the activated-carbon plates with a sterile soft-bristle brush. The samples were then placed in sterile glass containers and stored at − 80 °C until further processing.

Chemical extraction and analytical procedures

The dust samples were homogenized using a vortex mixer and filtered through a 100-mesh stainless steel sieve to obtain particles smaller than 150 μm. Approximately 50 mg of the sieved dust was weighed and spiked with 4 ng of mass-labeled internal standards. The sample was allowed to equilibrate in the dark for 1 h and then extracted as described in previous studies [72, 73] with minor modifications. Briefly, each dust sample was extracted three times using 4 mL of 0.1 M citric buffer (citric acid and sodium citrate, pH 4), followed by 4 mL of methanol (HPLC grade; Merck KGaA, Darmstadt, Germany), with ultrasonication used to facilitate the extraction. The sample was centrifuged at 8000 rpm for 10 min between each round of extraction. The supernatants from the three extractions were then combined and reduced to approximately 5 mL under a gentle stream of ultra-high-purity nitrogen (≥ 99.995%). Finally, the concentrated extracts were diluted to 125 mL with a solution of 0.02% formic acid (≥ 98%; Supelco, PA, USA) in Milli-Q water, and 0.2 g disodium ethylenediaminetetraacetate (Sigma-Aldrich, MO, USA) was added to the solution.

Next, the sample was loaded onto a Bond Elut SAX solid-phase extraction (SPE) cartridge (500 mg, 6 mL; Agilent Technologies, CA, USA) connected in tandem with an HLB SPE cartridge (200 mg, 6 mL; Waters Corporation, MA, USA). The SAX cartridge was used for cleanup of impurities, while the HLB cartridge was used for extraction. Both cartridges were preconditioned with 8 mL MeOH, 4 mL Milli-Q water, and 4 mL Milli-Q water with 0.02% formic acid. After sample loading, the HLB cartridge was washed with Milli-Q water containing 0.02% formic acid and vacuum-dried for 30 min. The target chemicals were subsequently eluted from the HLB cartridge first with 6 mL MeOH followed by 6 mL of a 1:1 (*v:v*) mixture of MeOH and acetone (Merck KGaA). The two eluate fractions were combined and concentrated to near-dryness under a gentle stream of ultra-high-purity nitrogen (≥ 99.995%). Finally, the extract was reconstituted to 200 μL with a 4:1 (*v:v*) mixture of Milli-Q water and MeOH.

Ninety chemicals, as listed in Table S7, were analyzed as described previously. These included 24 EDCs (including TCC, TCS, parabens, hormones, bisphenols, and alkylphenols), 32 non-antibiotic pharmaceuticals (including antihistamines, antidepressants, NSAIDs, and 10 other subclasses), and 20 antibiotics (including nitroimidazole, diaminopyrimidines, sulfonamides, fluoroquinolones, tetracyclines, macrolides) for all dust samples, together with 14 β-lactams (including monobactams, cephamycins, cephalosporins, carbapenems, and penicillins) for four selected dust samples. The target chemicals were quantified using an Agilent 1290 Infinity II ultra-high-performance liquid chromatography instrument interfaced with an AB Sciex 6500 triple-quadrupole tandem mass spectrometer (Applied Biosystems, CA, USA), utilizing positive and negative electrospray ionization (ESI + and ESI-) modes. Chromatographic separation was achieved using a Zorbax Eclipse plus C18 column (2.1 × 100 mm id, 1.8 μm; Agilent Technologies) connected in series with a Zorbax Eclipse plus C18 guard column (2.1 × 5 mm id, 1.8 μm). The column temperature was 40 ℃. The sample injection volume was 5 μL, and the mobile phase flowrate was 0.2 mL/min.

For the ESI - mode, the mobile phases were ultrapure water with 5 mM ammonia acetate and MeOH with 5 mM ammonia acetate. For the ESI+ mode, the mobile phases were ultrapure water with 0.02% formic acid and MeOH with 0.02% formic acid. The gradient profiles for the mobile phases are provided in Table S8. The source parameters were curtain gas at 30 psi, spray voltage at 4500 V, both ion source gases 1 and 2 at 50 psi, and the source temperature at 300 ℃. The optimized MS/MS parameters for the target analytes are provided in Table S9.

Quality assurance and quality control for chemical analysis

Duplicate analyses were conducted for each dust sample, excluding three with insufficient amounts, along with procedural blanks ($n = 10$) and procedural recoveries ($n = 5$). Averaged values of the replicate samples were used for statistical analysis. Procedural recoveries were determined by fortifying a known amount of the target chemicals on 50 mg of anhydrous sodium sulfate. The procedural blanks and recovery samples were subjected to the same extraction and cleanup procedures as the dust samples. The recoveries for all target chemicals ranged from 28 to 226% (average = 104%), with standard deviations varying between 0.1 and 35% (average = 9%). Method quantification limits (MQLs) were set at three times the standard deviation of the procedural

blank levels. Values below the MQLs were replaced with MQL/2 for statistical analysis.

Metagenomic sequencing and quality control

Approximately 100 mg of the collected dust samples were used for genomic DNA extraction. The samples were first cut into smaller, manageable pieces and then homogenized using sterile scissors. The genomic DNA was then extracted using the DNeasy PowerSoil Pro DNA extraction kit (Qiagen, CA, USA) following the manufacturer's protocol. Four extraction-reagent-only samples were processed in parallel with the dust samples to serve as negative controls. Library preparation and sequencing were performed by Novogene (Beijing, China), generating 150-bp paired-end reads according to the manufacturer's protocol. Briefly, extracted DNA was fragmented to an average size of 350 bp using the Covaris LE220R-plus system (Woburn, MA, USA). The fragments then underwent end-polishing, A-tailing, and ligation with full-length Illumina sequencing adapters, followed by PCR amplification using the Nextera DNA Flex Library Preparation Kit (Illumina, San Diego, CA, USA). The resulting PCR libraries were quantified, pooled, and sequenced on a NovaSeq 6000 platform (Illumina, San Diego, CA, USA).

The raw sequence data was first processed to remove sequence adapters using AdapterRemoval (v.2.3.3) [74]. Next, quality filtering and removal of human sequences were performed using KneadData (v.0.10.0), utilizing the human genome hg37 as the reference and the default parameters. To identify and remove potential contaminant sequences, a co-assembly of reads from the four quality-filtered negative control samples was carried out using MetaWRAP (v.1.2.1) [75] with the parameter "-m 1000." Reads from the dust samples that could be mapped to the contigs assembled from the negative controls were removed using KneadData. Additionally, all the contaminating species identified by the R package "decontam" (v.1.12; https://github.com/benjjneb/decontam) using the default threshold were removed from the dataset using KrakenTools (v.1.2). Finally, unpaired reads were removed from the paired-end FastQ files using fastq-pair (v.1.0) [76], yielding an average of ~ 37.9 (SD ± 15.8) million paired-end reads retained per dust sample for downstream analysis.

Diversity and community composition

Taxonomic classification of the paired-end reads was performed using Kraken2 (v.2.1.2) [77] with the standard Kraken2 database (v.2024.1.12), and species-level abundance was estimated using Bracken (v.2.9). Species-level taxonomy was used to identify factors associated with community diversity and compositional changes. Based on the taxonomic data, prediction strength analysis was performed to estimate the optimal number of clusters in the community using the K-means algorithm and the "prediction. strength" command (with 100 random splits) in the R package "fpc" (v.2.2–12). The analysis indicated that the highest mean prediction strength was 0.84 when $k=2$, where k is the number of clusters. This indicated the presence of two clusters among all samples. FindFungi (v.0.23.3) [78] was used to assign taxonomy to members of fungal communities.

HUMAnN3 (v.3.7) [79] was used to profile the potential metabolic functions of the metagenomes, including the analysis of KO assignments and metabolic pathways. Based on the MetaCyc v24.0 pathway data, contributional diversity (Gini–Simpson for within-sample and Bray–Curtis for between-sample diversity) [80] was calculated for each cluster and compared by selecting pathways detected in > 75% of samples within each cluster. ARG family markers were detected using the "shortbred_quantify" function in ShortBRED (v0.9.4) [81]. Principal coordinate analysis based on Bray–Curtis dissimilarity was performed using the "vegdist" function in the R package "vegan" (v.2.6–4), and the multivariate homogeneity of variances was analyzed to test for differences in multivariate dispersions between clusters using the "betadisper" function in the same R package. PERMANOVA with Benjamini–Hochberg adjustment was applied using the "adonis2" function in "vegan" with 999 permutations to test the influence of cluster, HEPA filter usage, household characteristics ($PM_{2.5}$, house size, number of occupants, window opening frequency, and air conditioner usage frequency), personal information (age, BMI, and marital status), and medical diagnoses (diabetes mellitus, hypertension, and dementia) on microbiome composition. The effects of anthropogenic chemicals on community differentiation, and antibiotics and antimicrobial chemicals on resistome differentiation, were assessed using distance-based redundancy analysis (db-RDA) with the "dbrda" function in "vegan." The Procrustes test was performed to determine the congruency between taxonomic and functional composition data using the "protest" function in "vegan." To assess α-diversity, paired-end reads were subsampled to a read depth of 3 million reads per sample using seqtk (v.1.3), which corresponded to the sequencing depth of the sample with the lowest number of reads. The Shannon index was calculated for the subsampled dataset in "vegan." The associations of taxonomic, functional, and antimicrobial resistance data with cluster were determined using MaAsLin2 [82] with the generalized linear model. An adjusted p-value (q-value) ≤ 0.05 was considered statistically significant. Statistical significance between two groups was assessed using the WRST with false discovery rate adjustment via

the Benjamini–Hochberg method. Comparisons among more than two groups were analyzed with the Kruskal–Wallis rank-sum test. Both tests were performed in the R package "stats" (v.4.2.2). A p-value of < 0.05 was considered statistically significant.

Functional annotation of contigs

The paired-end reads were first assembled into contigs using the "assembly" function of MetaWRAP (v.1.2.1) [75] with the parameter "-m 1000." Prodigal (v.2.6.3) [83] was then used to identify ORFs from the assembled contigs, including their amino acid and nucleic acid sequences, using the parameter "-p meta." CD-HIT (v.4.8.1) [84] was used to construct a non-redundant gene catalog, with the suggested parameters (sequence identity $\geq 95\%$ and sequence coverage $\geq 90\%$), yielding 2,730,259 ORF clusters. ARGs were annotated using the Comprehensive Antibiotic Resistance Database (CARD; v.3.2.8) [85] and BLASTp, with an E-value $\leq 10^{-5}$, identity $\geq 80\%$, and coverage $\geq 80\%$. VGs and MGEs were annotated based on the VFDB database [86] and mobileOG-db (Beatrix-1.6) [87], respectively, using the same annotation parameters as for ARGs.

The paired-end reads from all 32 metagenomes were mapped to the nucleic acid sequences of the ORFs using Bowtie2 (v.2.4.4) with the parameter "–fast." The resulting BAM files were processed using Samtool (v.1.2.0)—the "view" function was used to transfer the files, and the "sort" and "index" functions were used to sort and index the files, respectively. The coverage of the ORFs was then calculated for each sorted BAM file using the "coverage" function in CheckM (v.1.2.2). Finally, the relative abundance of each ORF was estimated using the calculated coverage and the total number of sequencing reads in each sample.

Network analysis

Co-occurrence networks were constructed based on strong ($|\rho| > 0.8$) and statistically significant ($p < 0.001$) correlations between ARGs, VGs, and MGEs that were present in at least 50% of the samples. The "corAndPvalue" function in the R package "WGCNA" was used to calculate the correlations and associated p-values between the variables, and the "graph_from_data_frame" function in the R package "igraph" was used to create the graph objects. The co-occurrence networks were visualized using Gephi (v.0.10.1). Edges connecting nodes within the same MGE category were excluded from the overall co-occurrence network construction (regardless of cluster), eliminating redundant features and minimizing collinearity. The topological properties of the nodes in the networks were analyzed using Gephi, and the statistical differences in these properties between clusters were examined using WRST. The vulnerability and robustness of the constructed networks were analyzed as previously described [88].

Reconstruction and annotation of MAGs

Contigs with lengths > 1000 bp were binned into MAGs using the "binning" function of MetaWRAP. The resulting MAGs were further refined using the "bin_refinement" function of MetaWRAP with the parameter "-c 50 -× 10" and dereplicated using the "dRep dereplicate" function of dRep (v.3.2.2). In total, 106 representative MAGs (rMAGs) with contamination $\leq 10\%$ and completeness $\geq 50\%$ were generated. The taxonomy of the 106 rMAGs was annotated using GTDB-TK (v.2.1.0). The phylogeny of the MAGs was performed using PhyloPhlAn3 [89] and visualized using the Interactive Tree of Life (iTOL) tool (https://itol.embl.de). rMAGs were classified as human pathogens if their corresponding species were found in both the gcPathogen database [90] (accessed June 15, 2024) and the curated human pathogen database [91]. The ORFs of the rMAGs were predicted using Prodigal with the parameter "-p meta," and the annotation of ARGs, VGs, and MGEs followed the same methods used for the contigs, with E-value $\leq 10^{-5}$, identity $\geq 60\%$, and coverage $\geq 80\%$.

Determination of viruses

VAMB (v.4.1.3) [92] was used to cluster contigs > 1000 bp into putative microbial taxa, generating k-mer-compositions of the sequences and the abundance of the contigs, which were then parsed with PHAMB (v.1.0.1) [93] to identify putative viral bins. The quality of these bins was assessed using CheckV (v.0.8.1; database v.1.0) [94] with the "end_to_end" pipeline. For the proviruses identified by CheckV, host regions were removed, retaining only the proviral regions of the viral bins for further analysis. All viral bins with completeness $\geq 50\%$ were clustered into species-level vOTUs based on a 95% ANI with a minimum coverage of 85%, using centroid-based clustering. Genus- and family-level vOTUs were generated using a combination of shared genes and amino acid identity (AAI) based on Markov clustering as described previously [95]. Briefly, viral bins with $< 20\%$ AAI or $< 10\%$ shared genes (using an inflation factor of 1.2) were clustered into family-level vOTUs, while those with $< 50\%$ AAI or $< 20\%$ shared genes (using an inflation factor of 2.0) were clustered into genus-level vOTUs. The ORFs within the vOTUs were predicted using Prodigal with the "-p meta" parameter. The annotation of ARGs and VGs in these ORFs was performed using the same method and threshold as for the original rMAGs. The lifestyle of vOTUs was predicted using BACPHLIP [96] and VIBRANT [97].

Determination of phage–host interactions

Both the in situ and ex situ hosts of the phages were identified as previously described [98]. Briefly, two host databases were used to establish the phage–host link, in which 552,481 complete bacterial genomes from RefSeq (downloaded from the NCBI database in November 2023) were used for identifying ex situ hosts, while 106 rMAGs were used for identifying in situ hosts. CRISPR spacers were extracted from the two host databases using a custom Python script. Hosts for the vOTUs were predicted using a combination of bioinformatic methods. These methods included (i) identifying viral sequences with exact matches (or close similarity) to host CRISPR spacers, (ii) detecting integrated viral fragments within host genomes, (iii) locating matches between viral sequences and host tRNA genes, and (iv) analyzing host k-mer signatures to identify associations with viral sequences. For ex situ host prediction, only methods (i) and (ii) were used. For method (i), BLASTn was used to compare CRISPR spacer sequences with the viral genomes, and matches with ≤ 1 mismatch and an E-value ≤ $1e^{-5}$ were retained. For any CRISPR spacer with a match in a viral genome, the repeat sequence from the same assembled CRISPR region was compared with all bacterial and archaeal genomes via BLASTn (E-value ≤ $1e^{-5}$, 100% nucleotide identity, and 95% coverage) to link that CRISPR region (and any viruses harboring spacers in that CRISPR region) to a host. For method (ii), a bit-score threshold of 50 with an E-value ≤ $1e^{-5}$ and a ≥ 96% ANI were used for identifying shared genomic regions via BLASTn, and only hits ≥ 1000 bp were considered. For method (iii), viral and host tRNA genes were predicted by tRNA-scan SE-2.0 using the general and bacterial/archaeal models, respectively, and BLASTn comparison was then performed between the predicted viral and bacterial tRNA genes. For method (iv), WIsH (v.1.1) [99] was used for host prediction after masking tRNA sequences on the viral genomes to improve performance. Additionally, 3024 viral genomes (downloaded from the NCBI Virus portal in November 2022) whose hosts are invertebrates were used as a decoy database after conservatively excluding viruses known to infect a host genus under prediction. For each viral genome, the WIsH-predicted host with the lowest *p*-value (≤ $1e^{-5}$) was retained for conservativeness with host assignments.

Identification of prophages and phage-plasmids

Prophages were identified using CheckV, PHASTER [100], and VIBRANT. To identify P-Ps, all vOTUs (including phages determined from phage–host links and prophages) were screened for plasmid functions by searching for proteins specific to plasmid replication and partition systems [101, 102]. Specifically, phages smaller than 300 kb (cutoff based on chromids at 250 kb and domesticated megaplasmids at 300 kb) and larger than 10 kb were searched for plasmid-associated genes using hidden Markov models (HMMs) specific to plasmid replication and partition systems [102]. A positive hit was assigned if the alignment covered ≥ 50% of the protein profile with a domain E-value < 10-3. The putative P-Ps were classified by clustering with 740 known P-P groups [102].

Construction of phylogenetic trees

The phylogenetic relationships between rMAGs affiliated with *Pseudomonas* and their reference genomes from NCBI GenBank were inferred using the UBCG2 pipeline (v3.0) [103]. The analysis included 1765 non-redundant complete strains within the genus *Pseudomonas*. A maximum-likelihood phylogenetic tree was constructed using FastTree (v.2.1.11) [104] and visualized using iToL.

To infer the phylogeny of the identified P-P S11C3896, BLASTn searches were performed against the NCBI Virus database (https://www.ncbi.nlm.nih.gov/labs/virus/; accessed May 13, 2024) using default parameters, yielding 112 reference sequences. The ORFs of the reference sequences and P-P S11C3896 were predicted using Prodigal. For marker selection, HMM searches of the ORFs were conducted against the profile HMMs of virus orthologous groups from the Pfam (v.37.0) [105], VOGDB (release 224; https://vogdb.org/), and TIGRFAMs (v.15.0) [106] databases using an E-value threshold of 1×10^{-3} and alignment coverage ≥ 50%, yielding 63 markers. The top HMM hits with the highest bit score were then individually aligned to the profile HMMs of these 63 markers using FAMSA (v.1.5.12) [107]. The individual marker alignments were then trimmed with trimAl (v.1.4) [108] to retain positions with less than 50% gaps. The trimmed alignments were concatenated, with gaps filled in for missing markers where necessary. Finally, a concatenated protein phylogeny was inferred from multiple-sequence alignment using FastTree and visualized using iToL.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40168-025-02081-2.

Supplementary Material 1.
Supplementary Material 2.
Supplementary Material 3.
Supplementary Material 4.
Supplementary Material 5.
Supplementary Material 6.
Supplementary Material 7.

Supplementary Material 8.

Supplementary Material 9.

Supplementary Material 10.

Acknowledgements
We thank all the participants of this study.

Authors' contributions
S.D. and H.L. prepared the samples, analyzed and interpreted the data, and wrote the manuscript. Q.L. developed instrumental methods for analyzing anthropogenic chemicals. C.L.M. and S.H.L. handled sample collection and processing. K.M.Y.L provided guidance on data analysis. K.F.H. conceived the study and supervised the fieldwork. P.K.H.L. conceived the study and supervised the data analysis and manuscript preparation. All authors have read and approved the final manuscript.

Funding
This research was supported by the Hong Kong Research Grants Council Research Impact Fund (R1016-20F) and the General Research Fund (11214721 and 11206224). Open Access made possible with partial support from the Open Access Publishing Fund of the City University of Hong Kong.

Data availability
The raw paired-end metagenomics sequences have been deposited in NCBI's BioProject under the accession number PRJNA1152926. Custom scripts and input and output files generated for analysis and figures, as well as the reference genomes used for decontamination, are available at the GitHub page https://github.com/Dorothydu12/dust_microbiome.

Declarations

Ethics approval and consent to participate
This research was approved by the Joint Chinese University of Hong Kong-New Territories East Cluster Clinical Research Ethics Committee (2017.304-T) and conducted in accordance with the standard operating procedures and principles of the Declaration of Helsinki and ICH Good Clinical Practice.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]School of Energy and Environment, City University of Hong Kong, Kowloon, Hong Kong SAR, China. [2]School of Energy and Environment and State Key Laboratory of Marine Pollution, City University of Hong Kong, Kowloon, Hong Kong SAR, China. [3]The Jockey Club School of Public Health and Primary Care, The Chinese University of Hong Kong, New Territories, Hong Kong SAR, China. [4]Department of Chemistry, City University of Hong Kong, Kowloon, Hong Kong SAR, China. [5]Low-Carbon and Climate Impact Research Centre, City University of Hong Kong, Kowloon, Hong Kong SAR, China.

Received: 25 November 2024 Accepted: 7 March 2025
Published online: 09 April 2025

References

1. Rintala H, Pitkäranta M, Täubel M. Microbial communities associated with house dust. Adv Appl Microbiol. 2012;78:75-120.
2. Ding LJ, Zhou XY, Zhu YG. Microbiome and antibiotic resistome in household dust from Beijing. China Environ Int. 2020;139:105702.
3. Lee M, Kaul A, Ward JM, Zhu Q, Richards M, Wang Z, et al. House dust metagenome and pulmonary function in a US farming population. Microbiome. 2024;12:129.
4. Estensmo ELF, Morgado L, Maurice S, Martin-Sanchez PM, Engh IB, Mattsson J, et al. Spatiotemporal variation of the indoor mycobiome in daycare centers. Microbiome. 2021;9:220.
5. Balasubrahmaniam N, King JC, Hegarty B, Dannemiller KC. Moving beyond species: fungal function in house dust provides novel targets for potential indicators of mold growth in homes. Microbiome. 2024;12:231.
6. Nastasi N, Renninger N, Bope A, Cochran SJ, Greaves J, Haines SR, et al. Persistence of viable MS2 and Phi6 bacteriophages on carpet and dust. Indoor Air. 2022;32:e12969.
7. Wang Y, Thompson KN, Yan Y, Short MI, Zhang Y, Franzosa EA, et al. RNA-based amplicon sequencing is ineffective in measuring metabolic activity in environmental microbial communities. Microbiome. 2023;11:131.
8. Hospodsky D, Qian J, Nazaroff WW, Yamamoto N, Bibby K, Rismani-Yazdi H, et al. Human occupancy as a source of indoor airborne bacteria. PLoS One. 2012;7:e34867.
9. Weschler CJ, Langer S, Fischer A, Bekö G, Toftum J, Clausen G. Squalene and cholesterol in dust from danish homes and daycare centers. Environ Sci Technol. 2011;45:3872–9.
10. Leppänen HK, Täubel M, Jayaprakash B, Vepsäläinen A, Pasanen P, Hyvärinen A. Quantitative assessment of microbes from samples of indoor air and dust. J Expo Sci Env Epid. 2018;28:231–41.
11. Leung MHY, Lee PKH. The roles of the outdoors and occupants in contributing to a potential pan-microbiome of the built environment: a review. Microbiome. 2016;4:21.
12. Zhou JC, Wang YF, Zhu D, Zhu YG. Deciphering the distribution of microbial communities and potential pathogens in the household dust. Sci Total Environ. 2023;872:162250.
13. Ludwig S, Jimenez-Bush I, Brigham E, Bose S, Diette G, McCormack MC, et al. Analysis of home dust for *Staphylococcus aureus* and staphylococcal enterotoxin genes using quantitative PCR. Sci Total Environ. 2017;581–582:750–5.
14. Hartmann EM, Hickey R, Hsu T, Betancourt Román CM, Chen J, Schwager R, et al. Antimicrobial chemicals are associated with elevated antibiotic resistance genes in the indoor dust microbiome. Environ Sci Technol. 2016;50:9807–15.
15. Darmancier H, Domingues CPF, Rebelo JS, Amaro A, Dionísio F, Pothier J, et al. Are virulence and antibiotic resistance genes linked? A comprehensive analysis of bacterial chromosomes and plasmids. Antibiotics. 2022;11:706.
16. Pan Y, Zeng J, Li L, Yang J, Tang Z, Xiong W, et al. Coexistence of antibiotic resistance genes and virulence factors deciphered by large-scale complete genome analysis. mSystems. 2020;5. https://doi.org/10.1128/msystems.00821-00819.
17. Wang X, Zhang H, Yu S, Li D, Gillings MR, Ren H, et al. Inter-plasmid transfer of antibiotic resistance genes accelerates antibiotic resistance in bacterial pathogens. ISME J. 2024;18:wrad032.
18. Xiao X, Ma XL, Han X, Wu LJ, Liu C, Yu HQ. TiO2 photoexcitation promoted horizontal transfer of resistance genes mediated by phage transduction. Sci Total Environ. 2021;760:144040.
19. Liao H, Liu C, Zhou S, Liu C, Eldridge DJ, Ai C, et al. Prophage-encoded antibiotic resistance genes are enriched in human-impacted environments. Nat Commun. 2024;15:8315.
20. Pfeifer E, Bonnin RA, Rocha EPC. Phage-plasmids spread antibiotic resistance genes through infection and lysogenic conversion. mBio. 2022;13:e01851-01822.
21. Pfeifer E, Rocha EPC. Phage-plasmids promote recombination and emergence of phages and plasmids. Nat Commun. 2024;15:1545.
22. Mercier F, Glorennec P, Thomas O, Bot BL. Organic contamination of settled house dust, a review for exposure assessment purposes. Environ Sci Technol. 2011;45:6716–27.
23. Moschet C, Anumol T, Lew BM, Bennett DH, Young TM. Household dust as a repository of chemical accumulation: new insights from a comprehensive high-resolution mass spectrometric study. Environ Sci Technol. 2018;52:2878–87.
24. Haglund P, Alygizakis NA, Covaci A, Melymuk L, Nizzetto PB, Rostkowski P, et al. Comprehensive characterization of European house dust contaminants: concentrations and profiles, geographical variability, and implications for chemical regulation and health risk. Sci Total Environ. 2024;957:177639.

25. Guo Y, Kannan K. A survey of phthalates and parabens in personal care products from the United States and its implications for human exposure. Environ Sci Technol. 2013;47:14442–9.
26. Halden RU, Paull DH. Co-occurrence of triclocarban and triclosan in US water resources. Environ Sci Technol. 2005;39:1420–6.
27. Halden RU, Lindeman AE, Aiello AE, Andrews D, Arnold WA, Fair P, et al. The Florence statement on triclosan and triclocarban. Environ Health Perspect. 2017;125:064501.
28. Björnsdotter MK, Romera-García E, Borrull J, de Boer J, Rubio S, Ballesteros-Gómez A. Presence of diphenyl phosphate and aryl-phosphate flame retardants in indoor dust from different microenvironments in Spain and the Netherlands and estimation of human exposure. Environ Int. 2018;112:59–67.
29. Anwar M, Iqbal Q, Saleem F. Improper disposal of unused antibiotics: an often overlooked driver of antimicrobial resistance. Expert Rev Anti Infect Ther. 2020;18:697–9.
30. Li X, Zhang W, Lv J, Liu W, Sun S, Guo C, et al. Distribution, source apportionment, and health risk assessment of phthalate esters in indoor dust samples across China. Environ Sci Eur. 2021;33:19.
31. Zhu Q, Wang M, Jia J, Hu Y, Wang X, Liao C, et al. Occurrence, distribution, and human exposure of several endocrine-disrupting chemicals in indoor dust: a nationwide study. Environ Sci Technol. 2020;54:11333–43.
32. Wang L, Liao C, Liu F, Wu Q, Guo Y, Moon H-B, et al. Occurrence and human exposure of p-hydroxybenzoic acid esters (parabens), bisphenol A diglycidyl ether (BADGE), and their hydrolysis products in indoor dust from the United States and three East Asian countries. Environ Sci Technol. 2012;46:11584–93.
33. Yang J, Ching YC, Kadokami K. Occurrence and exposure risk assessment of organic micropollutants in indoor dust from Malaysia. Chemosphere. 2022;287:132340.
34. Maertens RM, Bailey J, White PA. The mutagenic hazards of settled house dust: a review. Mutat Res. 2004;567:401–25.
35. Ait Bamai Y, Araki A, Kawai T, Tsuboi T, Saito I, Yoshioka E, et al. Exposure to phthalates in house dust and associated allergies in children aged 6–12years. Environ Int. 2016;96:16–23.
36. Li L-J, Xu F, Xu J-X, Yan Y, Su J-Q, Zhu Y-G, et al. Spatiotemporal changes of antibiotic resistance, potential pathogens, and health risk in kindergarten dust. Environ Sci Technol. 2024;58:3919–30.
37. Jarma D, Maestre JP, Sanchez J, Brodfuehrer S, Katz LE, Horner S, et al. Participant-collected household dust for assessing microorganisms and semi-volatile organic compounds in urban homes. Sci Total Environ. 2024;908:168230.
38. Fahimipour AK, Ben Maamar S, McFarland AG, Blaustein RA, Chen J, Glawe AJ, et al. Antimicrobial chemicals associate with microbial function and antibiotic resistance indoors. mSystems. 2018;3:e00200-00218.
39. Angiolillo DJ, Weisman SM. Clinical pharmacology and cardiovascular safety of naproxen. Am J Cardiovasc Drug. 2017;17:97–107.
40. Li C, Chen X. Parabens in indoor dust from houses, university dormitories, and cosmetics stores in Nanjing, China: occurrence and human exposure. Environ Sci Pollut R. 2023;30:26929–37.
41. Weir CB, Jan A. BMI classification percentile and cut off points. In: StatPearls. 2020.
42. Urban-Chmiel R, Marek A, Stępień-Pyśniak D, Wieczorek K, Dec M, Nowaczek A, et al. Antibiotic resistance in bacteria—a review. Antibiotics. 2022;11:1079.
43. Tao Y, Zheng D, Zou W, Guo T, Liao G, Zhou W. Targeting the cysteine biosynthesis pathway in microorganisms: mechanism, structure, and drug discovery. Eur J Med Chem. 2024;271:116461.
44. de Melo PL, Camargo FP, Dornelles HS, Granatto CF, Sakamoto IK, Grosseli GM, et al. Biodegradation of diclofenac and ibuprofen in Fluidized Bed Reactor applied to sanitary sewage treatment in acidogenic and denitrifying conditions. J Water Process Eng. 2022;49:102964.
45. Brooke JS. Stenotrophomonas maltophilia: an emerging global opportunistic pathogen. Clin Microbiol Rev. 2012;25:2–41.
46. Montaña S, Schramm STJ, Traglia GM, Chiem K, Parmeciano Di Noto G, Almuzara M, et al. The genetic analysis of an Acinetobacter johnsonii clinical strain evidenced the presence of horizontal genetic transfer. PLoS One. 2016;11:e0161528.
47. Du S, Tong X, Lai ACK, Chan CK, Mason CE, Lee PKH. Highly host-linked viromes in the built environment possess habitat-dependent diversity and functions for potential virus-host coevolution. Nat Commun. 2023;14:2676.
48. de Bentzmann S, Plésiat P. The *Pseudomonas aeruginosa* opportunistic pathogen and human infections. Environ Microbiol. 2011;13:1655–65.
49. Dominguez-Moñino I, Jurado V, Hermosin B, Saiz-Jimenez C. *Paracoccus cavernae* sp. nov., isolated from a show cave. Int J Syst Evol Micr. 2016;66:2265–70.
50. Zhu L, Hajeb P, Fauser P, Vorkamp K. Endocrine disrupting chemicals in indoor dust: a review of temporal and spatial trends, and human exposure. Sci Total Environ. 2023;874:162374.
51. Tran TM, Trinh HT, Anh HQ, Van Le T, Le SN, Minh TB. Characterization of triclosan and triclocarban in indoor dust from home micro-environments in Vietnam and relevance of non-dietary exposure. Sci Total Environ. 2020;732:139326.
52. Liao C, Liu F, Guo Y, Moon H-B, Nakata H, Wu Q, et al. Occurrence of eight bisphenol analogues in indoor dust from the United States and several Asian countries: implications for human exposure. Environ Sci Technol. 2012;46:9138–45.
53. Hwang S, Lim J-e, Choi Y, Jee SH. Bisphenol A exposure and type 2 diabetes mellitus risk: a meta-analysis. BMC Endocr Disord. 2018;18:1–10.
54. Larsson DGJ, Flach C-F. Antibiotic resistance in the environment. Nat Rev Microbiol. 2022;20:257–69.
55. Kirjavainen PV, Täubel M, Karvonen AM, Sulyok M, Tiittanen P, Krska R, et al. Microbial secondary metabolites in homes in association with moisture damage and asthma. Indoor Air. 2016;26:448–56.
56. Vishwanath V, Sulyok M, Weingart G, Kluger B, Täubel M, Mayer S, et al. Evaluation of settled floor dust for the presence of microbial metabolites and volatile anthropogenic chemicals in indoor environments by LC–MS/MS and GC–MS methods. Talanta. 2011;85:2027–38.
57. Murray LM, Hayes A, Snape J, Kasprzyk-Hordern B, Gaze WH, Murray AK. Co-selection for antibiotic resistance by environmental contaminants. npj Antimicrob Resist. 2024;2:9.
58. Geisinger E, Isberg RR. Interplay between antibiotic resistance and virulence during disease promoted by multidrug-resistant bacteria. J Infect Dis. 2017;215:S9–17.
59. Penadés JR, Chen J, Quiles-Puchalt N, Carpena N, Novick RP. Bacteriophage-mediated spread of bacterial virulence genes. Curr Opin Microbiol. 2015;23:171–8.
60. Kang Y, Wang J, Li Z. Meta-analysis addressing the characterization of antibiotic resistome in global hospital wastewater. J Hazard Mater. 2024;466:133577.
61. Xu C, Hu C, Li F, Liu W, Xu Y, Shi D. Antibiotic resistance genes risks in relation to host pathogenicity and mobility in a typical hospital wastewater treatment process. Environ Res. 2024;259:119554.
62. Humphrey S, Fillol-Salom A, Quiles-Puchalt N, Ibarra-Chávez R, Haag AF, Chen J, et al. Bacterial chromosomal mobility via lateral transduction exceeds that of classical mobile genetic elements. Nat Commun. 2021;12:6509.
63. Wang M, Jiang L, Wei J, Zhu H, Zhang J, Liu Z, et al. Similarities of P1-Like phage plasmids and their role in the dissemination of $bla_{CTX\text{-}M\text{-}55}$. Microbiol Spectr. 2022;10:e01410-01422.
64. Bellas CM, Anesio AM, Barker G. Analysis of virus genomes from glacial environments reveals novel virus groups with unusual host interactions. Front Microbiol. 2015;6:6.
65. Shan X, Szabo RE, Cordero OX. Mutation-induced infections of phage-plasmids. Nat Commun. 2023;14:2049.
66. Schmidtke DT, Hickey AS, Liachko I, Sherlock G, Bhatt AS. Analysis and culturing of the prototypic crAssphage reveals a phage-plasmid lifestyle. bioRxiv. 2024;585998. https://doi.org/10.1101/2024.03.20.585998.
67. Sternberg N, Hamilton D. Bacteriophage P1 site-specific recombination: I. Recombination between loxP sites. J Mol Biol. 1981;150:467–86.
68. Sun Y, Tang H, Du S, Chen Y, Ou Z, Zhang M, et al. Indoor metabolites and chemicals outperform microbiome in classifying childhood asthma and allergic rhinitis. Eco Environ Health. 2023;2:208–18.
69. Sun Y, Zhang M, Ou Z, Meng Y, Chen Y, Lin R, et al. Indoor microbiome, microbial and plant metabolites, chemical compounds, and asthma symptoms in junior high school students: a multicentre association study in Malaysia. Eur Respir J. 2022;60:2200260.
70. Zhou ZC, Liu Y, Lin ZJ, Shuai XY, Zhu L, Xu L, et al. Spread of antibiotic resistance genes and microbiota in airborne particulate matter, dust, and human airways in the urban hospital. Environ Int. 2021;153:106501.

71. Li X, Stokholm J, Brejnrod A, Vestergaard GA, Russel J, Trivedi U, et al. The infant gut resistome associates with *E. coli*, environmental exposures, gut microbiome maturity, and asthma-associated bacterial composition. Cell Host Microbe. 2021;29:975-987.e974.
72. Wu R, Ruan Y, Huang G, Li J, Lao J-Y, Lin H, et al. Source apportionment, hydrodynamic influence, and environmental stress of pharmaceuticals in a microtidal estuary with multiple outlets in South China. Environ Sci Technol. 2022;56:11374–86.
73. Zhou L-J, Ying G-G, Liu S, Zhao J-L, Chen F, Zhang R-Q, et al. Simultaneous determination of human and veterinary antibiotics in various environmental matrices by rapid resolution liquid chromatography–electrospray ionization tandem mass spectrometry. J Chromatogr A. 2012;1244:123–38.
74. Schubert M, Lindgreen S, Orlando L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC Res Notes. 2016;9:88.
75. Uritskiy GV, DiRuggiero J, Taylor J. MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. Microbiome. 2018;6:158.
76. Edwards JA, Edwards RA. Fastq-pair: efficient synchronization of paired-end fastq files. bioRxiv. 2019;552885. https://doi.org/10.1101/552885.
77. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. Genome Biol. 2019;20:257.
78. Donovan PD, Gonzalez G, Higgins DG, Butler G, Ito K. Identification of fungi in shotgun metagenomics datasets. PLoS One. 2018;13:e0192898.
79. Beghini F, McIver LJ, Blanco-Míguez A, Dubois L, Asnicar F, Maharjan S, et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. eLife. 2021;10:e65088.
80. Franzosa EA, McIver LJ, Rahnavard G, Thompson LR, Schirmer M, Weingart G, et al. Species-level functional profiling of metagenomes and metatranscriptomes. Nat Methods. 2018;15:962–8.
81. Kaminski J, Gibson MK, Franzosa EA, Segata N, Dantas G, Huttenhower C. High-specificity targeted functional profiling in microbial communities with ShortBRED. PLoS Comput Biol. 2015;11:e1004557.
82. Mallick H, Rahnavard A, McIver LJ, Ma S, Zhang Y, Nguyen LH, et al. Multivariable association discovery in population-scale meta-omics studies. PLoS Comput Biol. 2021;17:e1009442.
83. Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics. 2010;11:119.
84. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics. 2012;28:3150–2.
85. Alcock BP, Huynh W, Chalil R, Smith KW, Raphenya Amogelang R, Wlodarski MA, et al. CARD 2023: expanded curation, support for machine learning, and resistome prediction at the comprehensive antibiotic resistance database. Nucleic Acids Res. 2023;51:D690–9.
86. Liu B, Zheng D, Zhou S, Chen L, Yang J. VFDB 2022: a general classification scheme for bacterial virulence factors. Nucleic Acids Res. 2022;50:D912–7.
87. Brown Connor L, Mullet J, Hindi F, Stoll James E, Gupta S, Choi M, et al. mobileOG-db: a manually curated database of protein families mediating the life cycle of bacterial mobile genetic elements. Appl Environ Microb. 2022;88:e00991-00922.
88. Yuan MM, Guo X, Wu L, Zhang Y, Xiao N, Ning D, et al. Climate warming enhances microbial network complexity and stability. Nat Clim Chang. 2021;11:343–8.
89. Asnicar F, Thomas AM, Beghini F, Mengoni C, Manara S, Manghi P, et al. Precise phylogenetic analysis of microbial isolates and genomes from metagenomes using PhyloPhlAn 3.0. Nat Commun. 2020;11:2500.
90. Guo C, Chen Q, Fan G, Sun Y, Nie J, Shen Z, et al. gcPathogen: a comprehensive genomic resource of human pathogens for public health. Nucleic Acids Res. 2024;52:D714–23.
91. Bartlett A, Padfield D, Lear L, Bendall R, Vos M. A comprehensive list of bacterial pathogens infecting humans. Microbiology. 2022;168:001269.
92. Nissen JN, Johansen J, Allesøe RL, Sønderby CK, Armenteros JJA, Grønbech CH, et al. Improved metagenome binning and assembly using deep variational autoencoders. Nat Biotechnol. 2021;39:555–60.
93. Johansen J, Plichta DR, Nissen JN, Jespersen ML, Shah SA, Deng L, et al. Genome binning of viral entities from bulk metagenomics data. Nat Commun. 2022;13:965.
94. Nayfach S, Camargo AP, Schulz F, Eloe-Fadrosh E, Roux S, Kyrpides NC. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. Nat Biotechnol. 2021;39:578–85.
95. Nayfach S, Páez-Espino D, Call L, Low SJ, Sberro H, Ivanova NN, et al. Metagenomic compendium of 189,680 DNA viruses from the human gut microbiome. Nat Microbiol. 2021;6:960–70.
96. Hockenberry AJ, Wilke CO. BACPHLIP: predicting bacteriophage lifestyle from conserved protein domains. PeerJ. 2021;9:e11396.
97. Kieft K, Zhou Z, Anantharaman K. VIBRANT: automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences. Microbiome. 2020;8:90.
98. Jian H, Yi Y, Wang J, Hao Y, Zhang M, Wang S, et al. Diversity and distribution of viruses inhabiting the deepest ocean on Earth. ISME J. 2021;15:3094–110.
99. Galiez C, Siebert M, Enault F, Vincent J, Söding J. WIsH: who is the host? Predicting prokaryotic hosts from metagenomic phage contigs. Bioinformatics. 2017;33:3113–4.
100. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, et al. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res. 2016;44:W16–21.
101. Cury J, Touchon M, Rocha Eduardo PC. Integrative and conjugative elements and their hosts: composition, distribution and organization. Nucleic Acids Res. 2017;45:8943–56.
102. Pfeifer E, Moura de Sousa JA, Touchon M, Rocha EPC. Bacteria have numerous distinctive groups of phage–plasmids with conserved phage and variable plasmid gene repertoires. Nucleic Acids Res. 2021;49:2655–73.
103. Kim J, Na S-I, Kim D, Chun J. UBCG2: up-to-date bacterial core genes and pipeline for phylogenomic analysis. J Microbiol. 2021;59:609–15.
104. Price MN, Dehal PS, Arkin AP. FastTree 2 – approximately maximum-likelihood trees for large alignments. PLoS One. 2010;5:e9490.
105. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, et al. The Pfam protein families database in 2019. Nucleic Acids Res. 2018;47:D427–32.
106. Haft DH, Selengut JD, White O. The TIGRFAMs database of protein families. Nucleic Acids Res. 2003;31:371–3.
107. Deorowicz S, Debudaj-Grabysz A, Gudyś A. FAMSA: fast and accurate multiple sequence alignment of huge protein families. Sci Rep. 2016;6:1–13.
108. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25:1972–3.

Publisher's Note

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