nature cities

Article

https://doi.org/10.1038/s44284-026-00433-z

# Urban wetlands as hotspots of antibiotic resistomes and their potential viral transmission

Received: 19 January 2025

Accepted: 19 March 2026

Published online: 17 April 2026

Check for updates

Da Lin[1,2,3,9], Ying Liu[4,9], Xiaohui Liu[1] ✉, Shuai Du[2,5], Kenneth Mei Yee Leung[6], Yong-Guan Zhu[2,3,5,7], Fengchang Wu[8] ✉ & Dong Zhu[2,3,5] ✉

Urban wetlands serve a variety of healthful roles in cities, including as 'sponges' that absorb potential flood waters, leisure spaces for people and habitat for many species. However, urban wetlands also receive contaminated surface runoff and may become reservoirs of harmful contaminants, including related to our earlier efforts to safeguard health. Antibiotics are a public health mainstay, but overuse has promoted the spread among bacteria of genes enabling them to survive treatment. Here, by collecting and analyzing samples from 17 urban wetlands across China and comparing them with global datasets from natural lakes and urban raw sewage, we find these urban wetlands to be hotspots of antibiotic resistance genes (ARGs), with average abundances about nine times higher than in natural lakes and comparable to that in raw urban sewage. We further discover both human bacterial pathogens and indicators of the potential transfer of ARGs among bacteria ('horizontal transfer'), suggesting viruses in urban wetlands carrying ARGs might facilitate their spread within bacterial communities there. We also find higher levels of economic development associated with lower ARG abundances, suggesting socioeconomic factors could also shape the geographical distribution of ARGs in urban wetlands, perhaps through associated improvements in sewer systems. These findings emphasize the importance of collecting and treating stormwater before its release into urban wetlands to safeguard wildlife and human health.

Antibiotic resistance is a global health threat that contributes to rising mortality rates and escalating healthcare costs[1,2]. Given the interconnectedness of the environment, animals and human health under the 'One Health' framework[3,4], it is essential to investigate the fate of antibiotic resistance genes (ARGs), especially in ecosystems related to human activities[5–7]. Urban wetlands essentially elevate the livability of cities by preventing flooding during rainy season, facilitating plant growth for carbon sequestration, reducing ambient temperature, providing habitats for wildlife and creating recreation opportunities such as water sports and sport fishing[8,9]. Although the water quality in urban wetlands is generally good, urban wetlands may also represent a major yet relatively underexplored hubs of ARGs. This is due to their recipient of surface runoff and stormwater from adjacent densely populated areas with frequent anthropogenic activities[10] and several studies have demonstrated a strong correlation between human activities and ARGs[6,9,11]. Proper understanding of the presence of ARGs in urban wetlands can help us take the necessary measures to protect the health of wildlife and humans and, hence, promote sustainability.

Moreover, it remains unclear whether the accumulation of ARGs in urban wetlands can be augmented by the co-occurrence of viruses.

A full list of affiliations appears at the end of the paper. ✉e-mail: lxh7786@ouc.edu.cn; wufengchang@vip.skleg.cn; dzhu@iue.ac.cn

Viruses are widely distributed biological entities in the environment, playing crucial ecological roles within microbial communities[12]. The transfer of genetic material mediated by viruses, known as transduction, is recognized as one of the primary pathways for ARG exchange between prokaryotes[11]. Extensive ARGs and their transcriptional activity have been detected in virus from various environments, particularly those impacted by human activity[13]. This suggests that viruses may serve as an underestimated reservoir of ARGs in urban wetlands. Since viruses can survive in the environment for extended periods, they facilitate the widespread transfer of ARGs[14]. Therefore, further exploration is needed to understand the virus–host dynamics in urban wetlands.

Therefore, this study aimed to primarily test the hypothesis that urban wetlands harbor high abundances of ARGs and their potential for viral transmission. We conducted metagenomic sequencing, viral genome identification from the metagenomic data and socioeconomic data collection in 17 urban wetlands across nine representative provincial capitals/municipalities directly under the central government in China (Fig. 1a). Both intracellular and extracellular DNA were also collected in the urban wetlands within Beijing under both summer and winter seasons to comprehensively assess microbial information in these wetlands. Meanwhile, a thorough evaluation of ARG pollution in the urban wetlands was carried out by comparing them with documented ARG results in natural lakes, urban wetlands and raw sewage from various regions around the world. The main objectives of this study are to (1) reveal the level of ARG and human bacterial pathogens (HBPs) in the urban wetlands, (2) investigate the interaction between phage (viruses infecting prokaryotes) and their bacterial hosts and the relationship between the phage and ARGs and (3) examine the impact of socioeconomic factors on the occurrence of ARGs in the urban wetlands. For the latter, consider that socioeconomic inequalities may exacerbate the prevalence of ARGs in urban wetlands[15]. We hypothesize that socioeconomic factors play an important role in determining the fate of ARGs in both viruses and their hosts within these environments. Filling these knowledge gaps is essential for identifying high-risk areas and creating effective public health strategies in response to rapid urbanization and environmental changes. This study highlights the potential threat of microbial pollution in the urban wetlands, providing insightful evidences to guide informed decision on their water quality management and risk management of ARG contamination.

## Results and Discussion

### The pollution of ARGs and HBPs in urban wetlands

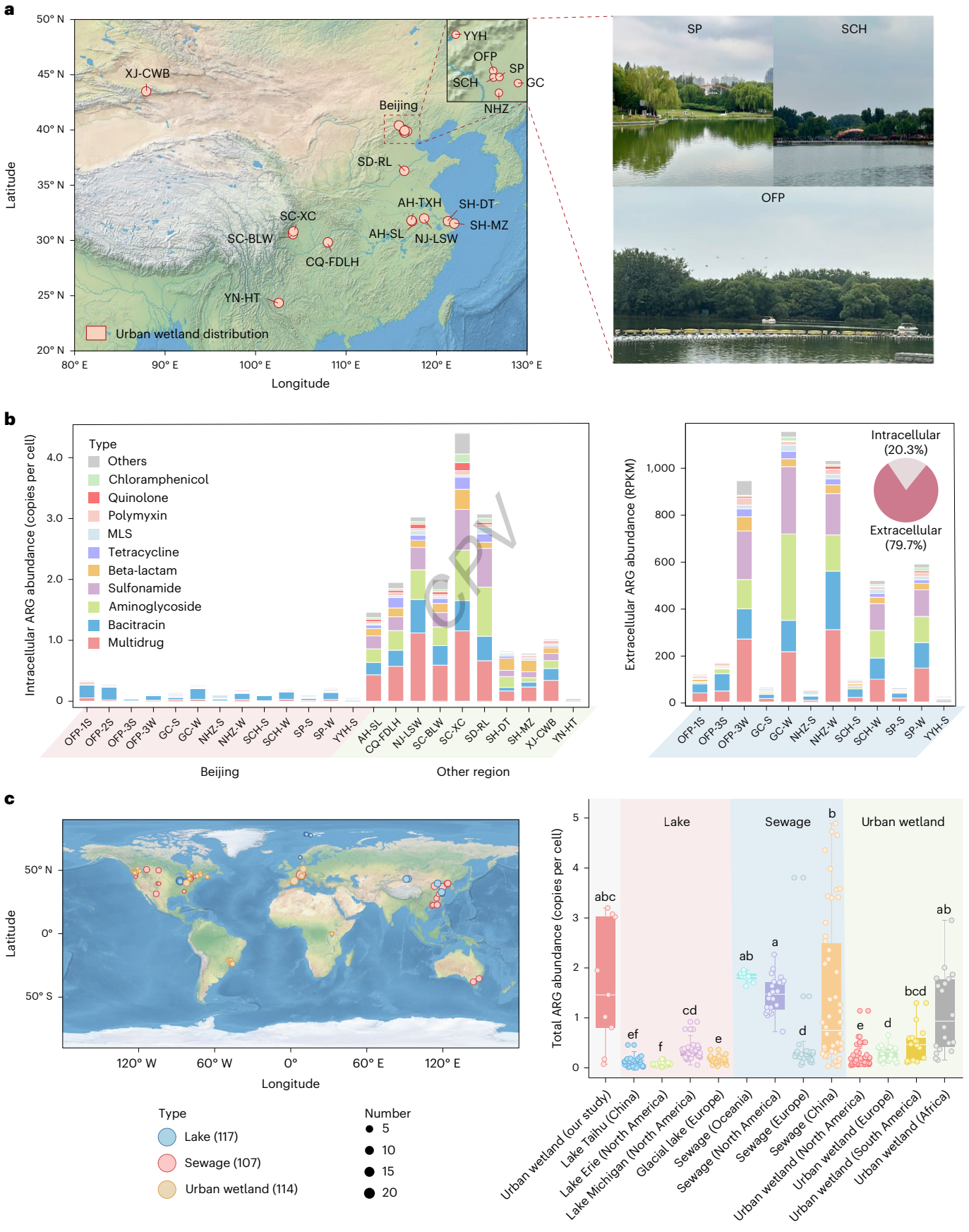
Despite the water quality of all the urban wetlands we studied was graded as IV (Chinese National Standard GB 3838-2002; moderate water quality, suitable for general uses such as recreational activities) or higher, a total of 749 different ARG subtypes were detected, with 52 distinct subtypes identified in phages (Fig. 1b and Extended Data Fig. 1). By examining the abundances of ARGs across various types of water bodies—urban wetlands, natural lakes and raw sewage—in different regions of the world, we found that the total abundances of ARGs in these urban wetlands were relatively high ($P < 0.05$; Fig. 1c and Supplementary Table 1). Specifically, the ARG abundances in these wetlands (average 1.642 copies per cell across nine provincial capitals/municipalities: Beijing, Anhui, Chongqing, Nanjing, Sichuan, Shandong, Shanghai, Xinjiang and Yunnan) are approximately 3.9× higher than the global average for other urban wetlands (average 0.421 copies per cell), about 9.0× higher than the average for natural lakes (average 0.183 copies per cell) and comparable to the levels found in most raw sewage (average 1.266 copies per cell) (Fig. 1c and Supplementary Table 1). Based on the risk indices for each ARG provided by Zhang et al.[16], which take into account human accessibility, mobility, pathogenicity and clinical availability, the ARG risk indices for all samples were quantified by multiplying the index for each ARG with the corresponding abundance in each sample and summing the results. The significant health risk posed by ARGs was identified in our urban wetlands compared to most other water bodies ($P < 0.05$; Extended Data Fig. 2). The current results confirm our hypothesis that the urban wetlands suffer from high contamination levels of ARGs.

In parallel, 97 co-occurrence patterns were observed in the contig analysis between 168 ARGs and 119 mobile genetic elements (MGEs), including transposases, integrases and insertion elements (Supplementary Table 2). The most common co-occurrence patterns between ARGs and MGEs include: the sulfonamide resistance gene *sul1*, nonspecific multidrug resistance genes (*mdsB* and *mexD*) and the tetracycline resistance gene *tet(A)* co-located on the same contig as the MGE gene *tnpA* (transposase) (Supplementary Table 2). As shown above, the results emphasize the potential correlation between ARGs and MGEs and the possible occurrence of horizontal transfer of ARGs in urban wetlands. Moreover, 67 HBPs were identified in these urban wetlands, including *Pseudomonas aeruginosa*, *Salmonella enterica*, *Klebsiella pneumoniae* and *Escherichia coli* (Fig. 2a). These pathogens account for 40.1% of the overall HBP abundances. Although the original habitats of these four strains differ: *P. aeruginosa* from various environments, *S. enterica* from warm-blooded animals, *K. pneumoniae* from both humans and the environment, and *E. coli* as a commensal of warm-blooded animals, their widespread presence across environments, animals and humans is well documented[17–20], given that the close connection between these habitats under the 'One Health' framework. To investigate the relationship between HBPs and overall ARGs, linear regression analysis was performed. The results indicate that the abundances of these pathogens and the overall abundances of HBPs closely aligned with those of ARGs ($P < 0.05$; Fig. 2b,c). It suggests the potential for the coexistence of antibiotic resistance and pathogenicity. The coexistence of genetic and microbial contamination in urban wetlands can be transmitted via direct contact, such as tourists with open wounds or mucosal exposure while playing in the water, or through poor hygiene practices (for example, eating food using hands that have touched polluted water without washing)[21–23],

**Fig. 1 | Sampling locations and ARG distribution. a**, Distribution of the 17 studied urban wetlands. **b**, The abundances of total ARGs in these urban wetlands across nine provincial capitals/municipalities: Beijing, Anhui, Chongqing, Nanjing, Sichuan, Shandong, Shanghai, Xinjiang and Yunnan, including intracellular (left) and extracellular (right) ARGs in the urban wetlands within Beijing. The pie chart represents the proportion of intracellular and extracellular ARGs. OFP-1, OFP-2 and OFP-3 represent three locations within Olympic Forest Park, respectively. The suffixes 'S' and 'W' denote summer and winter, respectively. MLS, macrolide–lincosamide–streptogramin. **c**, Left: collection of metagenomes from natural lakes, raw sewage and urban wetlands around the world. Right: the abundances of ARGs were annotated by global dataset and compared with the data from our urban wetland samples. The global dataset includes 117 lakes, 107 sewage systems and 114 urban wetlands across 5 continents. To minimize the bias introduced by the varying amounts of urban wetland samples collected from each provincial capital/municipality, the ARG abundances in each of the nine provincial capitals/municipalities were calculated by averaging the numbers from the urban wetland samples within that specific area, for comparison with ARG abundances in different global water bodies. For the box plots, the tops of the boxes represent the 75th percentile, the bottoms indicate the 25th percentile and the center lines denote the median. The whiskers extend to the maximum and minimum nonoutlier values. As shown in Supplementary Table 10, the data did not follow a normal distribution, so *P* values were calculated using the two-sided Kruskal–Wallis *H* test as detailed in Supplementary Table 11. Different lowercase letters (a–f) indicate significant differences among the treatments at $P < 0.05$. AH-TXH, Anhui Tangxihu; AH-SL, Anhui Swan Lake; CQ-FDLH, Chongqing Fengdu Longhe; GC, Grand Canal; NHZ, Nanhaiz; NJ-LSW, Nanjing Lvshuiwan; OFP, Olympic Forest Park; SC-BLW, Sichuan Bailuwan; SCH, Shichahai; SC-XC, Sichuan Xiangcheng; SD-RL, Shandong Rose Lake; SH-DT, Shanghai Dongtan; SH-MZ, Shanghai Mingzhu; SP, Sun Park; XJ-CWB, Xinjiang Chai Wo-pu Lake; YN-HT, Yunnan Hongta; YYH, Yeyahu.

https://doi.org/10.1038/s44284-026-00433-z

posing certain health risks. This highlights the importance of monitoring microbiological indicators in urban wetlands. The relative abundances of phage ARGs and total ARGs in the urban wetlands varied by sample type, season and location, with diverse resistance mechanisms and multidrug resistance displaying the highest relative abundances (Fig. 1b and Extended Data Fig. 1). Extracellular ARGs were more abundant than intracellular ARGs (Fig. 1b and Extended Data Fig. 1). This was also observed in other studies of aquatic environments[24].



https://doi.org/10.1038/s44284-026-00433-z

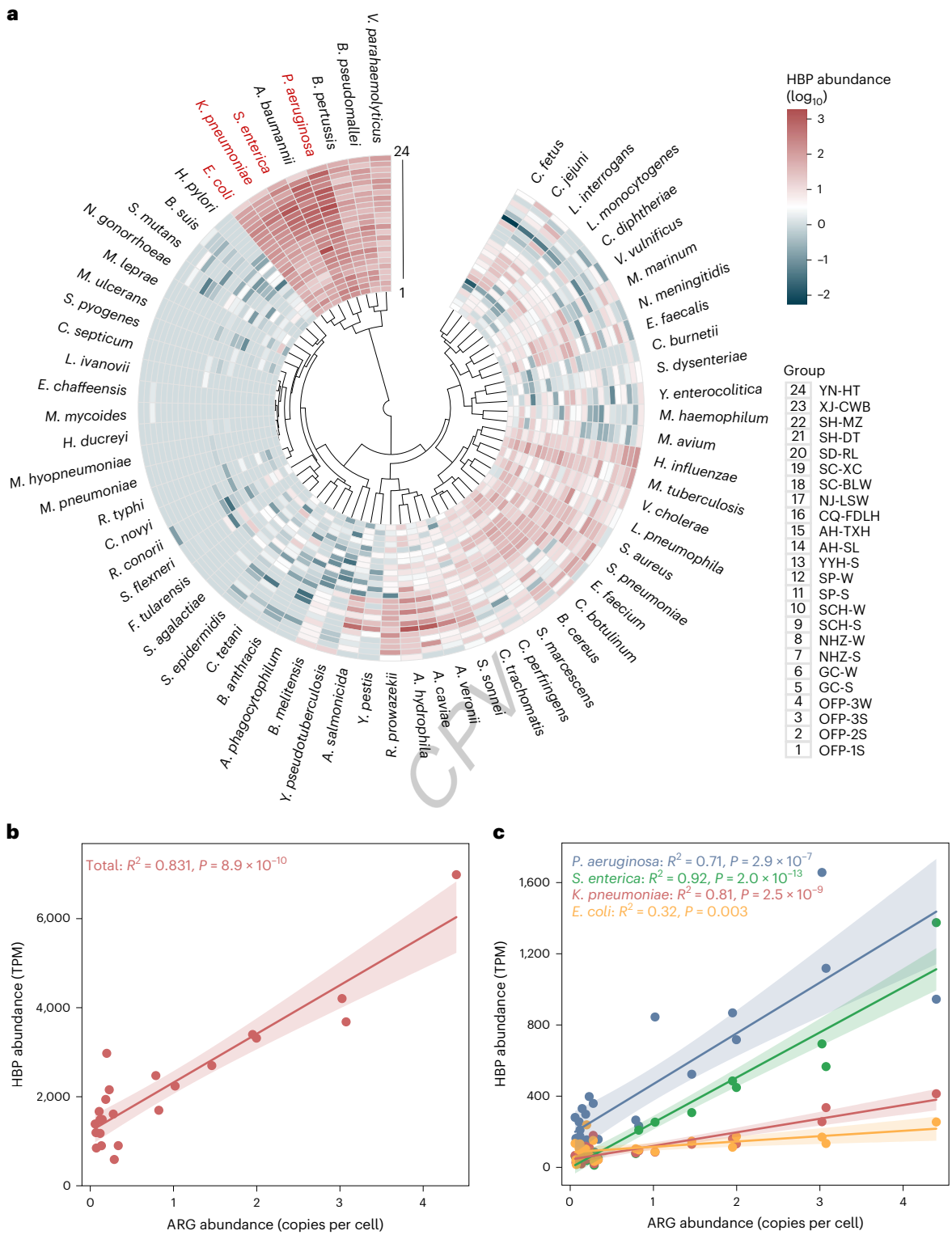



**Fig. 2 | HBPs in urban wetlands. a**, The abundances of HBPs in urban wetlands. The red markers indicate the most abundant HBPs. **b**, The relationship between the total abundances of HBPs and the total ARG abundances in these urban wetlands. **c**, The relationship between the highly abundant HBPs, including *P. aeruginosa*, *S. enterica*, *K. pneumoniae* and *E. coli*, and the total ARG abundances in these urban wetlands. Linear regression model with a two-sided test adjusted using the Benjamini–Hochberg was employed for statistical analysis. The shaded areas represent the 95% confidence interval, whereas the center lines indicate the line of best fit. The suffixes 'S' and 'W' denote summer and winter, respectively.

Given that most current studies primarily focus on intracellular ARGs, the potential issues posed by extracellular ARGs cannot be overlooked. Furthermore, the relative abundances of total ARGs were higher in winter compared to summer, both for intracellular and extracellular ARGs ($P < 0.05$; Fig. 1b and Supplementary Table 3). This may be attributed to the faster deactivation and degradation of DNA at higher temperatures[25,26] and the increase in antibiotic use during winter[27,28], which exerts selective pressure on microbial communities, leading to the

enrichment of ARGs[29]. The abundances of ARGs in developed areas such as Beijing (average 0.174 copies per cell) were generally lower compared to other regions (average 1.971 copies per cell) (Fig. 1b and Extended Data Fig. 1). This may be due to less developed cities lacking adequate sewage systems to properly manage stormwater. Stormwater runoff has been reported to carry high level of ARGs and pathogens[30,31]. Some of this untreated stormwater may end up in urban wetlands. By contrast, more developed cities such as Beijing have advanced infrastructure to collect stormwater, which may help minimize the influx of these polluted wastewaters into urban wetlands.

In addition, some representative ARGs—such as multidrug resistance genes *MexB*, *MexF* and *OprM*; polymyxin resistance gene *ugd*; and macrolide–lincosamide–streptogramin resistance gene *macB*—were selected for targeted qPCR assays to detect their absolute abundances to verify the results of metagenomic calculations of ARG abundances. These representative ARGs, characterized by high abundances and prevalence, were also detected in phages within these wetlands. Meanwhile, they exhibited significant positive correlation with total ARG abundances ($P < 0.05$; Supplementary Table 4) and have been classified in previous studies[16] as belonging to the high-risk group (Q1), with risk indices falling within the top 25% of all ARGs. The results demonstrated that the distribution of the abundances of these five ARGs, detected by targeted qPCR assays in different urban wetlands, was consistent with the metagenomic findings (Supplementary Table 5). All of them showed significant positive correlation with the abundances detected in the metagenome ($P < 0.05$; Supplementary Table 6), thereby confirming the reliability of ARG abundances derived from metagenomic data to some extent.

## The potential role of viruses in facilitating the transmission of ARGs

Our results found a strong correlation between ARGs in phages and the overall ARGs ($P < 0.05$; Extended Data Fig. 3), consistent with previous studies indicating that viruses may contribute to the evolution of antibiotic resistance in hosts as they adapt to their environment[32,33]. Based on this, a total of 274 medium and high quality metagenome-assembled genomes (MAGs) belonging to bacteria were obtained from all urban wetland samples, and their phage–host linkages were predicted. Among these, all bacterial subpopulations were hosted by phages, and 91.6% of these subpopulations were hosted phages that carried ARGs (Fig. 3). The number of ARGs carried by these subpopulations showed a significant positive correlation with the number of phages hosted by them and the number of ARGs present in those phages (Extended Data Fig. 4), emphasizing the potential effect of the viruses on ARG transmission.

The primary phage hosts at the bacterial phylum level were predicted to be Proteobacteria (56.8%), with ARGs associated with these phages representing 71.5% of the overall phage population (Fig. 3). Proteobacteria was also the key carrier of ARGs, particularly within the *Pseudomonas* genus. The *Pseudomonas* genus accounting for 52.8% of ARGs, 38.6% of virulence factor genes (VFGs) and 20.9% of MGEs within the overall bacterial communities (Fig. 3). Such bacterial subpopulations, known for their resistance to multiple antibiotics and their pathogenicity, have been frequently reported to be widely present in both the environment and the human body[34]. The phages exerted a crucial influence in this process, with 16 *Pseudomonas* genus were hosted by a total of 80 different phages (Extended Data Fig. 5). The co-occurrence network between these phages and *Pseudomonas* consisted of 96 nodes and 844 edges, with 93.8% of the nodes showing connectivity greater than 1 (Extended Data Fig. 6). This suggests that multiple *Pseudomonas* can be infected by a single phage, and a single *Pseudomonas* can be infected by multiple phages. Notably, 38.8% of these phages were connected to all *Pseudomonas* (Extended Data Fig. 5), indicating that these phages may facilitate the transfer of genetic material between different *Pseudomonas*.

In fact, we discovered that the multidrug resistance genes (for example, *MexA*, *MexB*, *MexF*, *MuxB* and *OprM*) present in the phages associated with all *Pseudomonas* were also found in every *Pseudomonas* (Extended Data Fig. 7). Our results suggest that phages may play a key role in the evolution of antibiotic resistance within these bacterial subpopulations. By comparing the genome sizes of 16 *Pseudomonas* and their associated 80 phages with those of 258 other bacterial subpopulations and their 1,610 associated phages, the results showed that *Pseudomonas* (average 5.6 Mbp) and their associated phages (average 103.8 kbp) had larger genomes than other bacteria (average 2.2 Mbp) and their associated phages (average 18.5 kbp) (Extended Data Fig. 8). This enables them to deliver larger gene cargo to their hosts and store more genetic material, potentially increasing the ARG spread and development[35].

## Potential socioeconomic factors influencing ARGs in urban wetlands

In this study, we gathered socioeconomic factors (for example, economy, population, water resource, green land area and wetland area) and climate factors (temperature and precipitation) from individual urban wetlands to investigate their contributions to the observed abundances of ARGs. These socioeconomic factors are linked to urban development and the construction of green rainwater infrastructure and urban wetlands. The abundances of total ARGs and phage ARGs were significantly negatively correlated with gross regional product (GRP) per capita ($P < 0.05$; Fig. 4a). This may further support our previous inference that the development of sponge cities[36] in China—designed to create infrastructure that absorbs runoff, controls flooding, recharges groundwater and reuses stormwater—tends to be more prevalent in areas with higher economic levels. These cities are associated with more efficient green stormwater infrastructure systems, which may help mitigate the risk of ARG pollution from stormwater runoff entering urban wetlands.

In addition, there is a trend where the abundances of total ARGs and phage ARGs firstly increase and then decrease as wetland area increases ($P < 0.05$; Fig. 4a). The relationship between wetland area and ARGs may stem from the initial increase in area, which resulted in higher foot traffic and increased human activities, ultimately leading to more severe levels of ARGs[16]. However, as the area continued to expand, improved urban wetland management practices helped reduce ARGs[37]. The random forest model further indicated that urban area and GRP per capita were the primary predictive factors for the abundances of ARGs ($P < 0.05$; Fig. 4b). The present results highlight the need for more focused classification and management of wetland areas, with particular attention to risk control in the urban wetlands of medium-sized. Our results highlight the need to integrate microbial pollution into urban wetland water quality monitoring and raising public awareness of antibiotic resistance. As global wetland sustainability requirements become increasingly multitarget specific, future efforts will focus on establishing green stormwater infrastructure systems, such as bioretention cells[38] and biofilters[39], which have been shown to effectively remove HBPs and ARGs from stormwater runoff, thereby rendering the runoff harmless before it is discharged into wetlands.

Although this research offers a detailed analysis of antibiotic resistomes and their potential viral transmission in urban wetlands, it is crucial to acknowledge certain limitations in our research approach, primarily relying on sequencing data. Given the inherent limitations of metagenomics analysis, such as the challenge of distinguishing between active and inactive organisms and the constraints of short-read assembly, metagenomics has limited capability to directly quantify the health risks associated with ARGs and the effect of viruses in ARGs transmission. Further studies are needed to fully understand and assess these issues in urban wetlands across multiple biological scales—from genomic and cellular to physiological and ecological—using multidisciplinary and multiscale approaches. These techniques should include microbial isolation and characterization, viral transplantation experiments and active microbial technology (for example, analyzing

https://doi.org/10.1038/s44284-026-00433-z

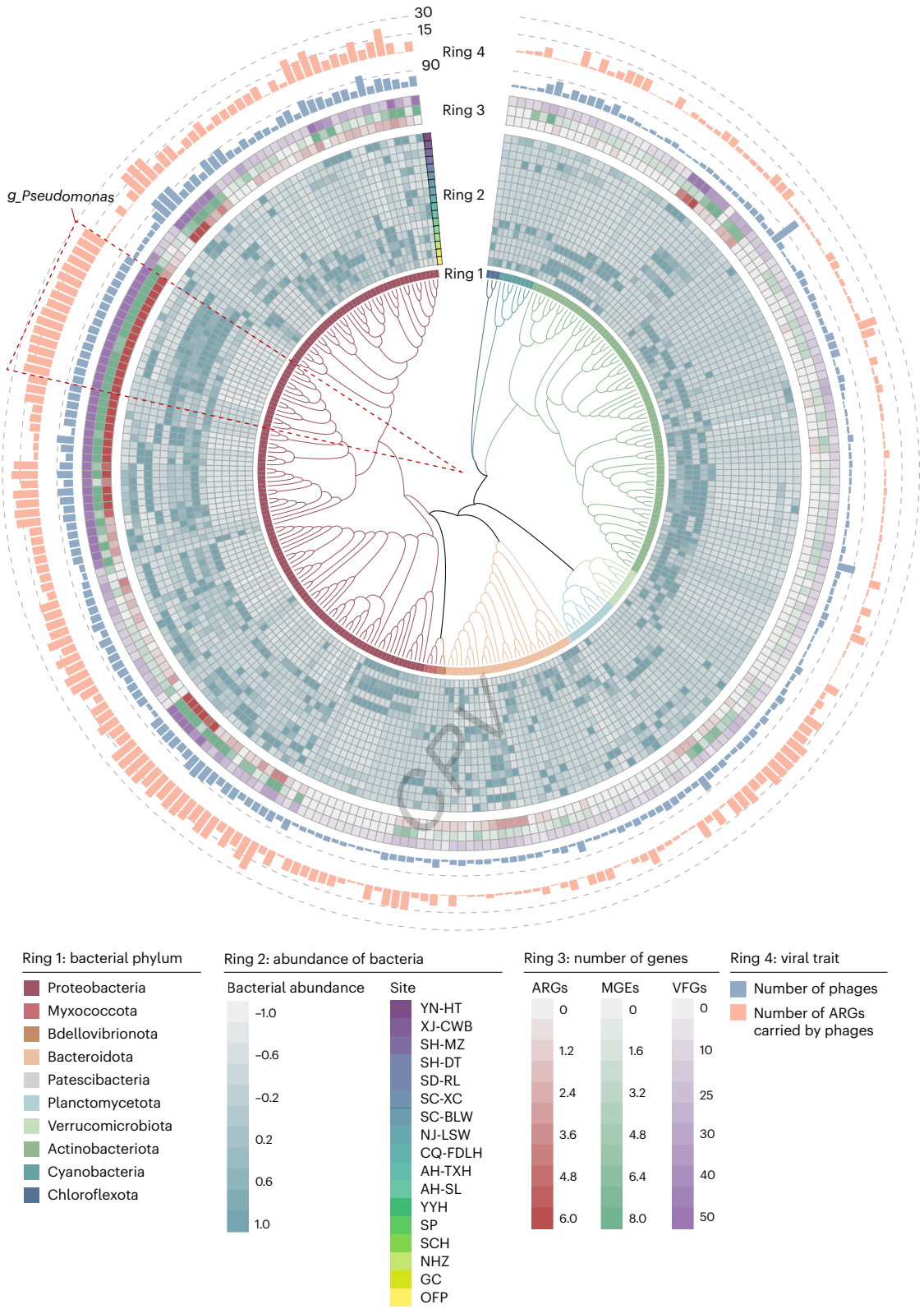



**Fig. 3 | Characterization of the interaction between MAGs and viruses.** The phylogenetic tree illustrates the taxonomy of assembled bacterial MAGs in the first ring. The heat map in the second ring represents the relative abundances of MAGs (normalized to *Z*-scores) at each site. The heat map in the third ring shows the number of ARGs, MGEs and VFGs (normalized to *Z*-scores) encoded by each MAG. The blue and orange bars in the fourth ring represent the number of host phages and the number of ARGs carried by those phages, respectively.

https://doi.org/10.1038/s44284-026-00433-z

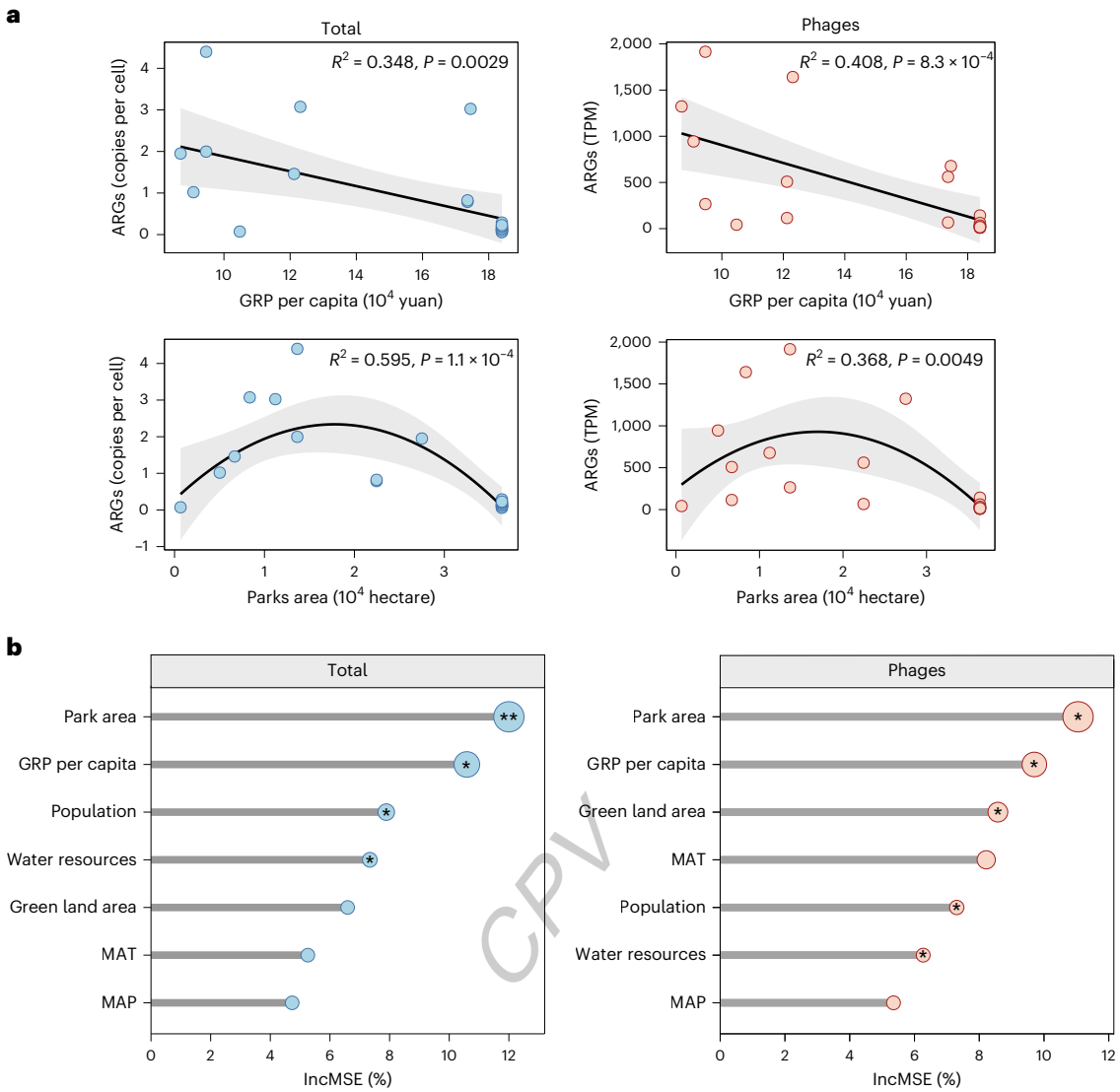



**Fig. 4 | Predictors influencing ARGs in urban wetland. a**, The relationship between park area and per capita GRP with total ARGs (blue) and phage ARGs (red) in urban wetlands. Linear and polynomial regression model with a two-sided test adjusted using the Benjamini–Hochberg was employed for statistical analysis. The shaded areas represent the 95% confidence interval, and the center lines indicate the line of best fit. **b**, The importance of predictors was assessed by calculating the percentage increase in the mean squared error (MSE) for each variable, with higher MSE% values indicating greater predictor importance. MAT, mean annual temperature; MAP, mean annual precipitation. Statistical significance was assessed using permutation tests ($n_{tree}$ = 1,000). Significance levels, as shown in Supplementary Table 12, are indicated as follows: $^*P < 0.05$, $^{**}P < 0.01$ and $^{***}P < 0.001$.

gene expression levels and transcriptional activity in active organisms), with a focus on evaluating the antibiotic resistance risks posed by active pathogens and quantifying changes in ARGs mediated by viruses. This can offer valuable insights into urban wetland ecotoxicology and its link to public health, helping to develop sustainable policies for mitigating ecological risks in urban environments and safeguarding both urban water and human health.

Although our research on the separation of intracellular and extracellular DNA provided valuable insights, more effective and rigorous separation methods should be considered in the future, as 0.22-μm filter membranes may not fully remove bacteria. In addition, this study highlights the possible link between basal socioeconomic factors and ARG abundances in urban wetlands. However, future research should further explore the impact of specific infrastructure changes, particularly those related to green rainwater systems, driven by urban development on ecological health of wetlands. This could enhance our understanding of the evolving relationship between urban society and water health and inform future urban policy development.

## Conclusion

Our study reveals the potential risk of ARGs in urban wetlands, reflected primarily by unexpectedly high abundances and the diversity of ARGs they harbor, and the potential for phage-mediated transmission of ARGs among bacterial communities. This provides critical warning information regarding the ecological health and public health risks associated with urban wetland and new insights into virus–host dynamics in these environments. We also discovered that this risk is influenced by socioeconomic factors, and positive economic development and effective urban wetland management can help reduce ARGs in urban wetlands. From a 'One Health' perspective, these observations are vital as they underscore the need for developing effective human intervention strategies to curb the spread of antibiotic resistance in urban wetlands.

## Methods

### Typical urban wetlands in China

We collected samples from 17 urban wetlands in China (Fig. 1). Detailed sample type, season and location information are provided

in Supplementary Table 7. To ensure sufficient and uniform DNA extraction from urban wetland samples, 10 l of water from each wetland in every region was first filtered through 8-µm filter membranes to remove impurities, and intracellular DNA was then collected using multiple 0.22-µm filter membranes, with each filter processing more than 1 l of water. Although extracellular DNA was obtained by precipitating the filtered water with ferric chloride, followed by filtration through 0.8-µm membranes as previously described[40]. Specifically, the filtered water was treated with 10 g $l^{-1}$ $FeCl_3$, incubated for 1 h at room temperature, and then filtered through 0.8-µm polycarbonate membranes, which retained the extracellular DNA and flocculated viruses. DNA was extracted using the FastDNA SPIN Kit for Soil (MP Biomedicals) according to the manufacturer's instructions. Extracted DNA concentration and purity were determined via micro-spectrophotometry (NanoDrop LITE, Thermo Scientific).

### Metagenomic sequencing and data analysis

The DNA sequencing was conducted using the Illumina Hiseq X platform operated by Majorbio Biotechnology. The extracellular DNA from sample 'OFP-2S', representing one location within Olympic Forest Park collected during the summer, failed to produce a metagenomic library. To minimize systematic errors and batch effects and to ensure that this would not impact our research (as two other locations in Olympic Park already contain both extracellular and intracellular DNA), sample 'OFP-2S' was not subjected to secondary sequencing. The raw metagenomic data were trimmed using Fastp v0.20.0 (https://github.com/OpenGene/fastp)[41]. Potential ARGs were annotated using local ARG-OAP (version 3.2) based on the structured ARG reference database (SARG) (identity ≥80%, hit length ≥75% and E-value ≤$10^{-7}$)[42]. The abundances of intracellular ARGs were normalized by adjusting the read abundance based on the length of each gene and by calculating gene copies relative to the total cell numbers in each sample (that is, copies per cell).

To evaluate the ARG risk of each sample, ARG risk indicators for each were calculated. The risk indices for each ARG were provided by Zhang et al.[16], with the risk index calculated as human accessibility × mobility × human pathogenicity × clinical availability. Furthermore, the ARG risk index for all samples was calculated based on previously established standards[43]. The specific formula used for the calculation is as follows:

$$\mathrm{RI}_{\mathrm{sample}} = \sum_{i=1}^{n} \mathrm{abundance}_i \times \mathrm{RI}_i.$$

Among these variables, 'abundance' and 'RI', respectively, denote the abundance and risk indices of each ARGs ($i$) within the sample, respectively, where $n$ represents the total number of ARGs detected in the sample.

Due to the potential inaccuracies in standardizing extracellular ARGs based on cell counting, the abundances of both intracellular and extracellular ARGs were also normalized according to read counts and gene length (that is, reads per kilobase of exon model per million mapped reads, RPKM) for comparison.

The clean metagenomic data from all samples were independently assembled into contigs using MEGAHIT v1.2.9[44]. Open reading frames were predicted using the Prodigal v2.6.3[45], and the CD-HIT algorithm was applied to generate a nonredundant gene catalog[46]. To obtain annotation and taxonomic information for the bacteria, sequences from the nonredundant gene set were compared to those in the NCBI Non-Redundant Database (NR) database, with an E-value threshold of ≤$10^{-5}$. Bacteria were classified as HBPs by referencing a compilation of emerging/re-emerging pathogens, Virulence Factor Database (VFDB) and the World Health Organization (WHO) priority ARB list, as reported in previous studies[47]. MAGs were recovered from each metagenome via MetaBAT2[48]. The quality of the obtained MAGs was assessed using CheckM v2.1.1[49], and only medium and high quality MAGs with a completeness of >50% and a contamination of <10% were kept for further analysis. The relative abundances of MAGs were determined by calculating the proportion of the average coverage for each MAG across the metagenomic sequencing samples. Taxonomic and phylogenetic analyses were performed using GTDB-Tk v1.7.0[50]. Open reading frames predicted from contigs and MAGs by Prodigal v2.6.3[45] were queried against the SARG[42], VFDB[51] and MGE[52] databases using BLASTX of DIAMOND v2.0.14.152[53], with thresholds set at identity >60%, hit length ≥80%, and E-value ≤$10^{-5}$. ARGs and MGEs were considered co-contigs if they were located on the same contig[54].

### Targeted qPCR assay

Based on the results of our metagenomic analysis, five representative ARGs (*MexB*, *MexF*, *OprM*, *ugd* and *macB*) were selected for targeted qPCR assays using the ABI QuantStudio 12K Real-Time PCR System (Applied Biosystems) to measure their absolute abundances and verify the reliability of ARG abundances obtained from metagenomic data. The primer sequences for these ARGs are shown in Supplementary Table 8. PCR amplification of these ARGs was performed using the PrimerSTAR Max DNA Polymerase Kit (Takara) according to the manufacturer's instructions. The resulting PCR products were then ligated into the pClone007 vector. After that, the recombinant products were transformed into competent *E. coli* cells using the Rapid Competent Cell Preps Kit (Takara). The plasmids were then extracted according to the protocol provided by the SanPrep Column Endotoxin-Free Plasmid Mini-Preps Kit (Sangon). The approximate concentration and purity of plasmids were assessed using micro-spectrophotometry (NanoDrop LITE, Thermo Scientific) to confirm successful extraction. The exact plasmid concentration was subsequently measured using the Qubit 3.0 fluorometer (Invitrogen). The plasmids were diluted in a tenfold gradient to construct a standard curve for subsequent qPCR assays. SYBR Green Premix Ex Taq (Takara) was used for qPCR assays. The amplification conditions were as follows: one cycle of initial denaturation at 95 °C for 5 min, followed by 45 cycles of 95 °C for 15 s, annealing at 60 °C for 15 s and final extension at 72 °C for 35 s. Each sample was analyzed in triplicate to minimize potential bias and assess the reproducibility of the qPCR results. The absolute copy number of genes was determined by constructing standard curves based on threshold cycles versus the log-copy number of genes. Gene abundances were then normalized to copies per milliliter based on the amount of water used for DNA extraction.

### Socioeconomic factors of urban wetlands

We obtained socioeconomic factors related to urban development and the construction of green rainwater infrastructure and urban wetlands, from the investigated urban wetlands. These factors include population size, GRP per capita, water resources, green land area and park area. These factors, along with climate variables such as mean annual temperature and mean annual precipitation, were obtained for each city from the China City Statistical Yearbook. These factors were identified as predictors of ARG abundances through random forest analyses using the rfPermute package. Regression and Spearman correlation analyses were conducted to explore the relationship between ARG abundances and multiple factors.

### Global dataset collection

In November 2024, we conducted a literature search on Web of Science, Google Scholar and PubMed using the keywords 'lake', 'sewage', 'urban water body' and 'metagenome' to gather metagenomic data on natural lakes, raw sewage and urban wetlands globally. To reduce potential confusion during the analysis, we carefully filtered the metagenomic samples from these studies according to the following criteria: (1) only water body samples were included, excluding sediment samples; (2) samples without geographic coordinates were excluded, as these would not enable the identification of biogeographic patterns; and

(3) samples treated with chemical compounds in the laboratory were excluded. After careful screening, we compiled metagenomic samples from 117 lakes, 107 sewage systems and 112 urban wetlands across five continents. Detailed information on the metagenomes is available in Supplementary Table 9. To ensure a more accurate comparison between the collected metagenomic data and our experiments, we gathered metagenomic samples from various locations in lakes and sewage systems in China, specifically those closer to our experimental site. Moreover, the annotation methods and conditions for ARGs were consistent with those used in our experiment. Detailed data on the overview of all ARGs across the samples are provided in Supplementary Table 1.

### Identification of viral sequences and host prediction

Viral sequences were identified from the assembled contigs ≥5,000 bp using a combination of DeepVirfinder (v1.1)[55], Virsorter2 (v2.0)[56] and VIBRANT (v1.2.1)[57]. Contigs that met the criteria of 95% average nucleotide identity and 85% alignment fraction were then were then clustered into viral operational taxonomic units (vOTUs) using CD-HIT v4.6[46]. The completeness and contamination of these vOTUs were then estimated using CheckV v1.0.1 (database v1.2)[58]. The protein-coding sequences of vOTUs were predicted using Prodigal v2.6.3[45]. BLASTX of DIAMOND v2.0.14.152[53] was used to align the sequences, and a sequence was classified as an ARG fragment if it met the following criteria: identity ≥40%, hit length ≥75% and E-value ≤ $10^{-7}$. The abundances of ARGs were normalized based on gene length and read counts (that is, transcripts per kilobase of exon model per million mapped reads (TPM)) using Bowtie2 v2.5.0[59].

In this study, three computational strategies based upon nucleotide sequence similarity were used for host prediction, namely: (1) CRISPR spacer bar matching: CRISPR spacer sequences from the bacterial genome were predicted using a CRISPR recognition tool (CRT, v2.1)[60]. Detected spacer sequences were queried against the viral genome using the BLASTn-short[61] function with the parameter of at least 95% identity over the entire spacer length, allowing only one to two SNPs at the end of the sequence, and E-value ≤ $10^{-10}$. (2) Transfer RNA (tRNA) matching: tRNA genes were annotated using tRNAscan-SE (v2.0.9)[62] and then compared to prokaryotic genomes using BLASTn[61], with a threshold of >95% sequence identity. Moreover, viral genomes were aligned with prokaryotic genomes based on shared genomic regions using BLASTn[61]. Matches that met the following criteria were retained: bit score ≥50, nucleotide identity ≥70%, match length ≥2,500 bp and E-value ≤ $10^{-3}$. (3) Nucleotide sequence homology between the bacterial genome and viral genome: vOTU sequences were compared to a microbial genome dataset using BLASTn[61]. The match criteria identity ≥70% nucleotide, coverage ≥75%, bit score ≥50 and E-value ≤ $10^{-3}$. The co-occurrence network between phages and their hosts was constructed using the Gephi interactive platform.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

All sequence data are available in the NCBI database under SRA accession number PRJNA1172644. Source data are provided with this paper.

## References


1. The Lancet The antimicrobial crisis: enough advocacy, more action. *Lancet* **395**, 247 (2020).
2. Murray, C. J. et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet* **399**, 629–655 (2022).
3. Singh, B. K., Yan, Z. Z., Whittaker, M., Vargas, R. & Abdelfattah, A. Soil microbiomes must be explicitly included in One Health policy. *Nat. Microbiol.* **8**, 1367–1372 (2023).
4. Banerjee, S. & van der Heijden, M. G. A. Soil microbiomes and One Health. *Nat. Rev. Microbiol.* **21**, 6–20 (2023).
5. Yin, Y. et al. Nanoplastics released from textile washing enrich antibiotic resistance and virulence genes in sewage sludge microbiomes. *Environ. Int.* **202**, 109611 (2025).
6. Ni, N. et al. Fibrous and fragmented microplastics discharged from sewage amplify health risks associated with antibiotic resistance genes in aquatic environments. *Environ. Sci. Technol.* **59**, 15919–15930 (2025).
7. Lin, D., Xu, J. Y., Wang, L., Du, S. & Zhu, D. Long-term application of organic fertilizer prompting the dispersal of antibiotic resistance genes and their health risks in the soil plastisphere. *Environ. Int.* **183**, 108431 (2024).
8. Alikhani, S., Nummi, P. & Ojala, A. Urban wetlands: a review on ecological and cultural values. *Water* **13**, 3301 (2021).
9. Mao, D., Wang, Z., Song, K. & Yang, H. Rescue urban wetlands for flood resilience. *Nature* **624**, 42 (2023).
10. Li, R. et al. Impact of urbanization on antibiotic resistome in different microplastics: evidence from a large-scale whole river analysis. *Environ. Sci. Technol.* **55**, 8760–8770 (2021).
11. Li, R. et al. Viral metagenome reveals microbial hosts and the associated antibiotic resistome on microplastics. *Nat. Water* **2**, 553–565 (2024).
12. Paez-Espino, D. et al. Uncovering Earth's virome. *Nature* **536**, 425–430 (2016).
13. Liao, H. et al. Prophage-encoded antibiotic resistance genes are enriched in human-impacted environments. *Nat. Commun.* **15**, 8315 (2024).
14. Debroas, D. & Siguret, C. Viruses as key reservoirs of antibiotic resistance genes in the environment. *ISME J.* **13**, 2856–2867 (2019).
15. Gupta, S., Wu, X., Pruden, A., Zhang, L. & Vikesland, P. Global scale exploration of human faecal and sewage resistomes as a function of socio-economic status. *Nat. Water* **2**, 975–987 (2024).
16. Zhang, Z. et al. Assessment of global health risk of antibiotic resistance genes. *Nat. Commun.* **13**, 1553 (2022).
17. Pulford, C. V. et al. Global diversity and evolution of *Salmonella enterica* serovar Panama: a genomic epidemiology study. *Lancet Microbe* **6**, 101150 (2025).
18. Kim, D. D. et al. Contaminated drinking water facilitates *Escherichia coli* strain-sharing within households in urban informal settlements. *Nat. Microbiol.* **10**, 1198–1209 (2025).
19. Letizia, M., Diggle, S. P. & Whiteley, M. *Pseudomonas aeruginosa*: ecology, evolution, pathogenesis and antimicrobial susceptibility. *Nat. Rev. Microbiol.* **23**, 701–717 (2025).
20. Rolbiecki, D. et al. Global dissemination of *Klebsiella pneumoniae* in surface waters: genomic insights into drug resistance, virulence, and clinical relevance. *Drug Resist. Updat.* **79**, 101204 (2025).
21. Uberoi, A., McCready-Vangi, A. & Grice, E. A. The wound microbiota: microbial mechanisms of impaired wound healing and infection. *Nat. Rev. Microbiol.* **22**, 507–521 (2024).
22. Geng, Y. et al. Anthropogenic activity and climate change exacerbate the spread of pathogenic bacteria in the environment. *Sci. Adv.* **11**, eads4355 (2025).
23. Leoni Swart, A. et al. *Pseudomonas aeruginosa* breaches respiratory epithelia through goblet cell invasion in a microtissue model. *Nat. Microbiol.* **9**, 1725–1737 (2024).
24. Li, Y. et al. Engineered DNA scavenger for mitigating antibiotic resistance proliferation in wastewater treatment. *Nat. Water* **2**, 758–769 (2024).
25. McCartin, L. J. et al. Temperature controls eDNA persistence across physicochemical conditions in seawater. *Environ. Sci. Technol.* **56**, 8629–8639 (2022).

26. Schnecker, J. et al. Improving measurements of microbial growth, death, and turnover by accounting for extracellular DNA in soils. *Soil* **10**, 521–531 (2024).
27. Sun, D. S. et al. Analysis of multiple bacterial species and antibiotic classes reveals large variation in the association between seasonal antibiotic use and resistance. *PLoS Biol.* **20**, e3001579 (2022).
28. Serletti, L. M. et al. Analysis of seasonal variation of antibiotic prescribing for respiratory tract diagnoses in primary care practices. *Open Forum Infect. Dis.* **9**, ofac492.777 (2022).
29. Brown, C. L. et al. Selection and horizontal gene transfer underlie microdiversity-level heterogeneity in resistance gene fate during wastewater treatment. *Nat. Commun.* **15**, 5412 (2024).
30. O'Malley, K. et al. Environmental drivers impact the accumulation and diversity of antibiotic resistance in green stormwater infrastructure. *J. Hazard. Mater.* **469**, 133923 (2024).
31. Lee, S. et al. Residential urban stormwater runoff: a comprehensive profile of microbiome and antibiotic resistance. *Sci. Total Environ.* **723**, 138033 (2020).
32. Kauffman, K. M. et al. Resolving the structure of phage–bacteria interactions in the context of natural diversity. *Nat. Commun.* **13**, 372 (2022).
33. Ruan, C. et al. Phage predation accelerates the spread of plasmid-encoded antibiotic resistance. *Nat. Commun.* **15**, 5397 (2024).
34. Qin, S. et al. *Pseudomonas aeruginosa*: pathogenesis, virulence factors, antibiotic resistance, interaction with host, technology advances and emerging therapeutics. *Signal Transduc. Target. Ther.* **7**, 199 (2022).
35. Yi, X. et al. Giant viruses as reservoirs of antibiotic resistance genes. *Nat. Commun.* **15**, 7536 (2024).
36. Liu, D. Water supply: China's sponge cities to soak up rainwater. *Nature* **537**, 307 (2016).
37. Wang, M. et al. Impacts of net cages on pollutant accumulation and its consequence on antibiotic resistance genes (ARGs) dissemination in freshwater ecosystems: insights for sustainable urban water management. *Environ. Int.* **183**, 108357 (2024).
38. Lancaster, E., Winston, R., Martin, J. & Lee, J. Urban stormwater green infrastructure: evaluating the public health service role of bioretention using microbial source tracking and bacterial community analyses. *Water Res.* **259**, 121818 (2024).
39. Rugh, M. B. et al. Highly variable removal of pathogens, antibiotic resistance genes, conventional fecal indicators and human-associated fecal source markers in a pilot-scale stormwater biofilter operated under realistic stormflow conditions. *Water Res.* **219**, 118525 (2022).
40. Buscaglia, M., Iriarte, J. L., Schulz, F. & Díez, B. Adaptation strategies of giant viruses to low-temperature marine ecosystems. *ISME J.* **18**, wrae162 (2024).
41. Chen, S., Zhou, Y., Chen, Y. & Gu, J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).
42. Yin, X. et al. ARGs-OAP v3.0: Antibiotic-resistance gene database curation and analysis pipeline optimization. *Engineering* **27**, 234–241 (2023).
43. Lin, D. et al. Climate warming fuels the global antibiotic resistome by altering soil bacterial traits. *Nat. Ecol. Evol.* **9**, 1512–1526 (2025).
44. Li, D., Liu, C. M., Luo, R., Sadakane, K. & Lam, T. W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**, 1674–1676 (2015).
45. Hyatt, D. et al. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinf.* **11**, 119 (2010).
46. Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150–3152 (2012).
47. Zhu, L. et al. Insights into microbial contamination in multi-type manure-amended soils: the profile of human bacterial pathogens, virulence factor genes and antibiotic resistance genes. *J. Hazard. Mater.* **437**, 129356 (2022).
48. Kang, D. D., Froula, J., Egan, R. & Wang, Z. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* **3**, e1165 (2015).
49. Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**, 1043–1055 (2015).
50. Parks, D. H. et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* **36**, 996 (2018).
51. Liu, B., Zheng, D., Zhou, S., Chen, L. & Yang, J. VFDB 2022: a general classification scheme for bacterial virulence factors. *Nucleic Acids Res.* **50**, D912–D917 (2022).
52. Pärnänen, K. et al. Maternal gut and breast milk microbiota affect infant gut antibiotic resistome and mobile genetic elements. *Nat. Commun.* **9**, 3891 (2018).
53. Buchfink, B., Reuter, K. & Drost, H. G. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods* **18**, 366–368 (2021).
54. Song, J. et al. Carbendazim shapes microbiome and enhances resistome in the earthworm gut. *Microbiome* **10**, 63 (2022).
55. Ren, J. et al. Identifying viruses from metagenomic data using deep learning. *Quant. Biol.* **8**, 64–77 (2020).
56. Guo, J. et al. VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. *Microbiome* **9**, 37 (2021).
57. 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* **8**, 90 (2020).
58. Nayfach, S. et al. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. *Nat. Biotechnol.* **39**, 578–585 (2021).
59. Peng, Y., Leung, H. C. M., Yiu, S. M. & Chin, F. Y. L. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* **28**, 1420–1428 (2012).
60. Bland, C. et al. CRISPR Recognition Tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. *BMC Bioinf.* **8**, 209 (2007).
61. Ye, J., McGinnis, S. & Madden, T. L. BLAST: improvements for better sequence analysis. *Nucleic Acids Res.* **34**, 6 (2006).
62. Lowe, T. M. & Chan, P. P. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res.* **44**, W54–W57 (2016).


## Acknowledgements

This work was financially supported by the National Key Research and Development Program of China (grant no. 2024YFE0106300 to Y.-G.Z.), National Natural Science Foundation of China (grant nos. 22193062 to D.Z. and 42207260 to X.L.), Fujian Provincial Natural Science Foundation of China (grant no. 2023J02031 to D.Z.), Taishan Scholars Program of Shandong Province (grant no. tsqn202312094 to X.L.) and Shandong Provincial Higher Education Institution Youth Innovation Teams (grant no. 2023KJ034 to X.L.).

## Author contributions

D.L., X.L., D.Z. and F.W. conceived of and designed the research. D.L., Y.L. and X.L. performed the experiments. D.L., S.D. and D.Z. analyzed the data and prepared the figures. D.L., X.L., D.Z., K.M.Y.L, Y.-G.Z. and F.W. wrote and revised the paper. All authors read and approved the paper.

## Competing interests
The authors declare no competing interests.

## Additional information
**Extended data** is available for this paper at https://doi.org/10.1038/s44284-026-00433-z.

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s44284-026-00433-z.

**Correspondence and requests for materials** should be addressed to Xiaohui Liu, Fengchang Wu or Dong Zhu.

**Peer review information** *Nature Cities* thanks Andrea di Cesare, Kazuaki Matsui and Lisa Paruch for their contribution to the peer review of this work.

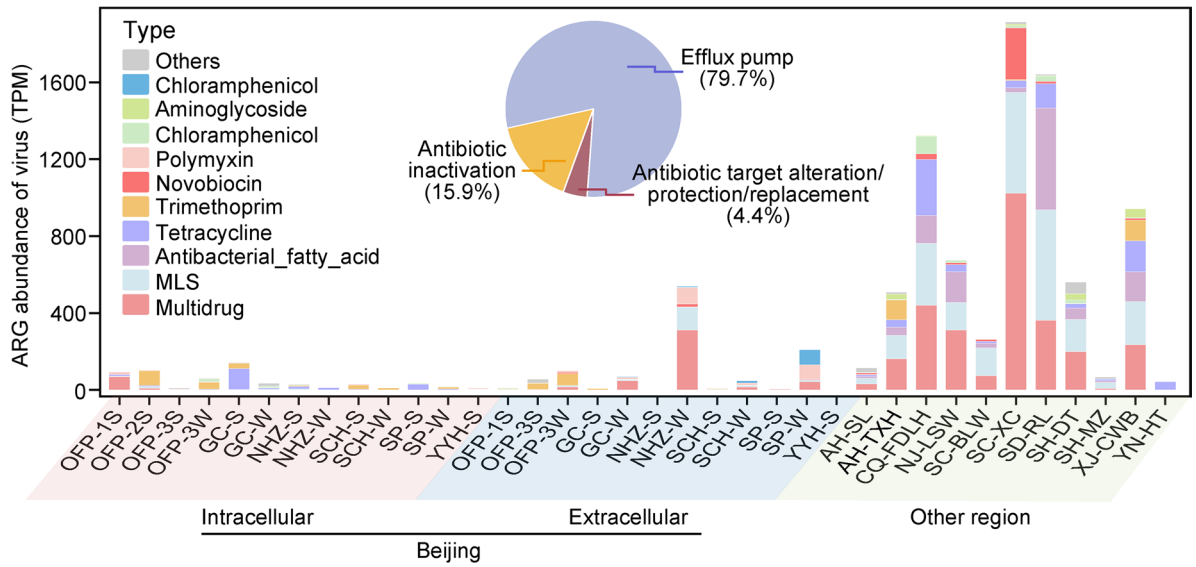
**Reprints and permissions information** is available at www.nature.com/reprints.

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



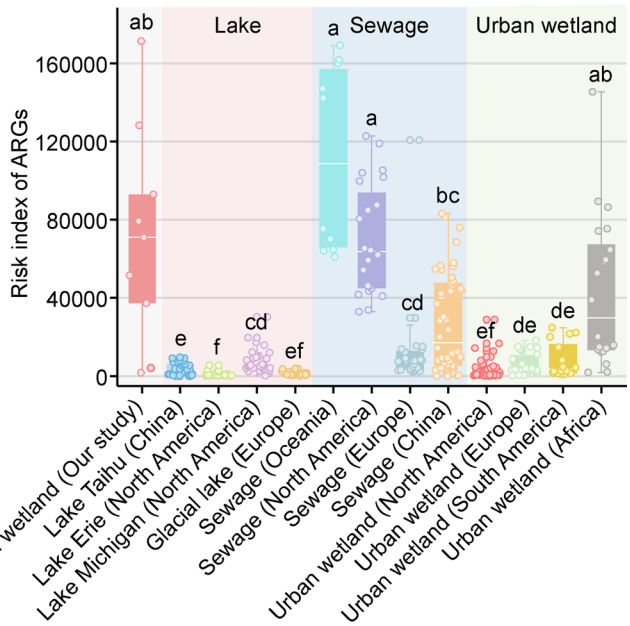
[1]Key Laboratory of Marine Environment and Ecology, Ministry of Education and College of Environmental Science and Engineering, Ocean University of China, Qingdao, China. [2]State Key Laboratory of Regional and Urban Ecology, Ningbo Observation and Research Station, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen, China. [3]University of Chinese Academy of Sciences, Beijing, China. [4]School of Environmental Science and Engineering, Qingdao University, Qingdao, China. [5]Zhejiang Key Laboratory of Pollution Control for Port-Petrochemical Industry, CAS Haixi Industrial Technology Innovation Center in Beilun, Ningbo, China. [6]State Key Laboratory of Marine Pollution, Department of Chemistry, and School of Energy and Environment, City University of Hong Kong, Hong Kong, China. [7]State Key Laboratory of Regional and Urban Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing, China. [8]State Key Laboratory of Environmental Criteria and Risk Assessment, Chinese Research Academy of Environmental Sciences, Beijing, China. [9]These authors contributed equally: Da Lin, Ying Liu. ✉e-mail: lxh7786@ouc.edu.cn; wufengchang@vip.skleg.cn; dzhu@iue.ac.cn

https://doi.org/10.1038/s44284-026-00433-z



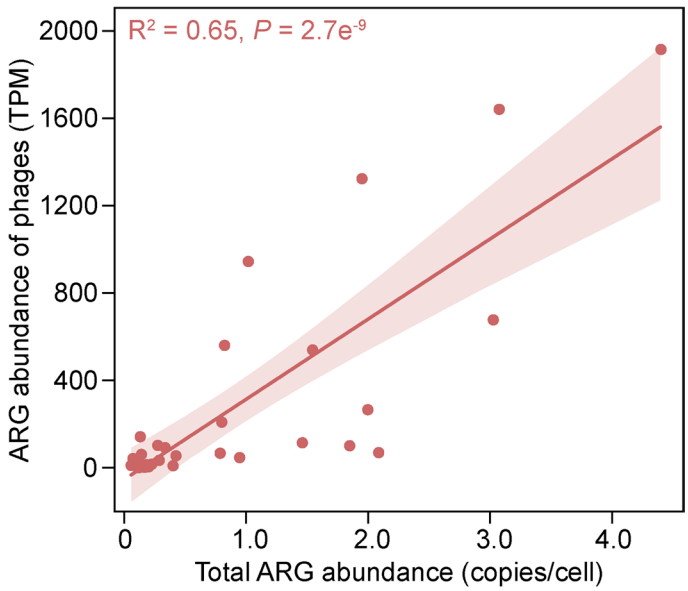

**Extended Data Fig. 1 | The abundances of phage ARGs in urban wetlands, including intracellular and extracellular ARGs in Beijing.** Pie chart represents the proportion of different antibiotic resistance mechanisms.

https://doi.org/10.1038/s44284-026-00433-z



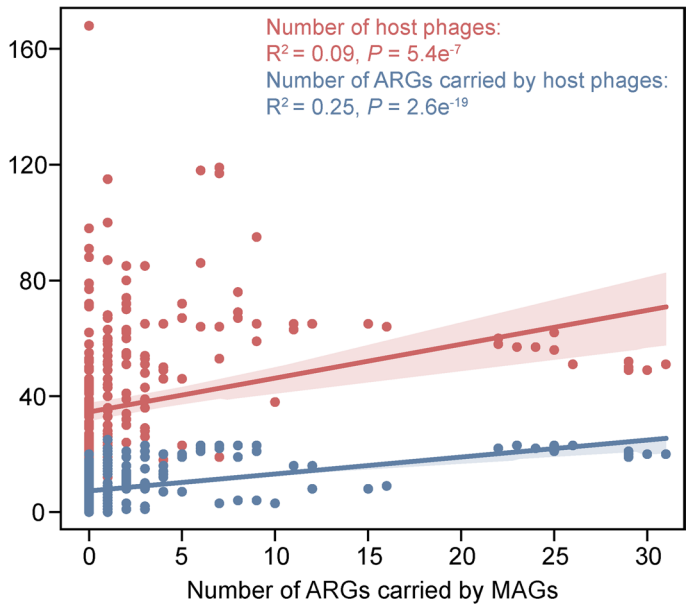

**Extended Data Fig. 2 | Comparison of the abundances of risk indices for ARGs in this study with those from natural lakes, sewage effluents, and urban wetlands worldwide.** For box plots, the tops of the boxes represent the 75th percentile, the bottoms indicate the 25th percentile and the centre lines denote the median. The whiskers extend to the maximum and minimum non-outlier values. As shown in Supplementary Table 13, the data (n = 347) did not follow a normal distribution, so *P* values were calculated using the two-sided Kruskal-Wallis H test as detailed in Supplementary Table 14. Different lowercase letters indicate significant differences among the treatments at $P < 0.05$.

https://doi.org/10.1038/s44284-026-00433-z



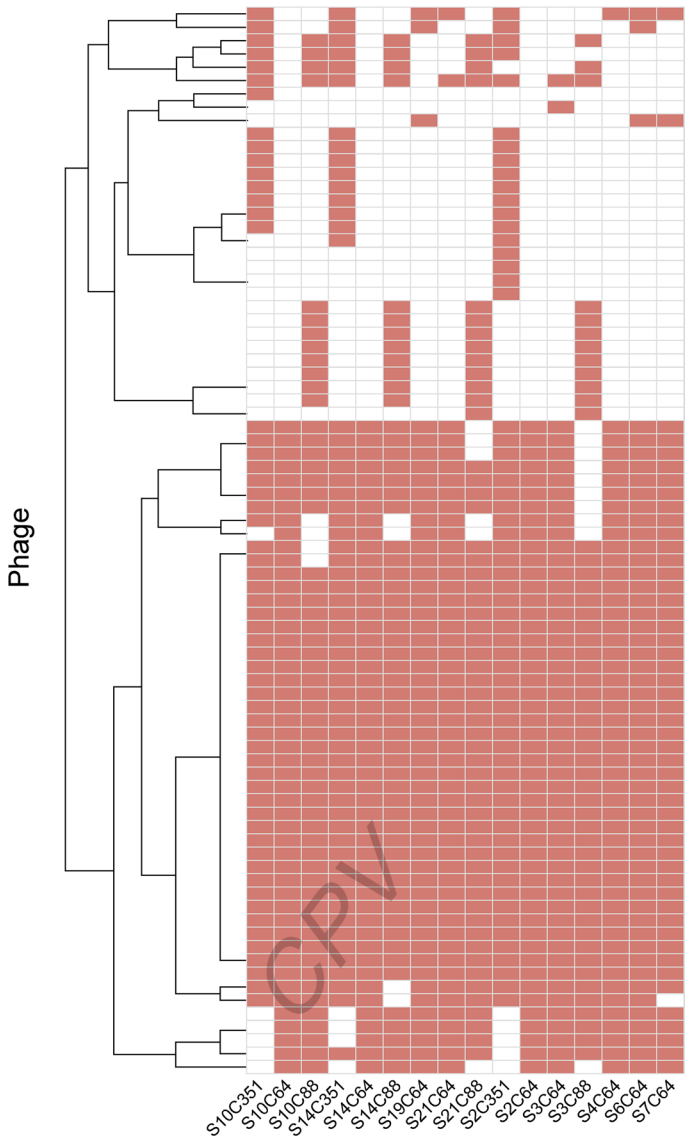

**Extended Data Fig. 3 | The relationship between phage ARG abundances and total ARG abundances.** Linear regression model with a two-sided test adjusted using the Benjamini-Hochberg was employed for statistical analysis. The shaded areas represent the 95% confidence interval, while centre lines indicate the line of best fit.

https://doi.org/10.1038/s44284-026-00433-z



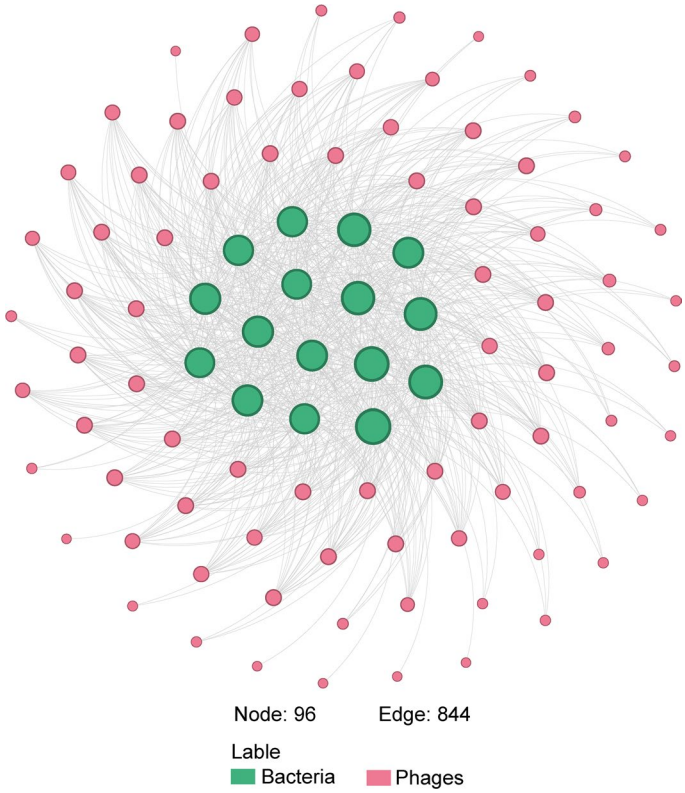

**Extended Data Fig. 4 | The relationship between the number of ARGs carried by MAGs, the number of phages hosted by them and the number of ARGs present in those phages.** Linear regression model with a two-sided test adjusted using the Benjamini-Hochberg was employed for statistical analysis. The shaded areas represent the 95% confidence interval, while centre lines indicate the line of best fit.

https://doi.org/10.1038/s44284-026-00433-z



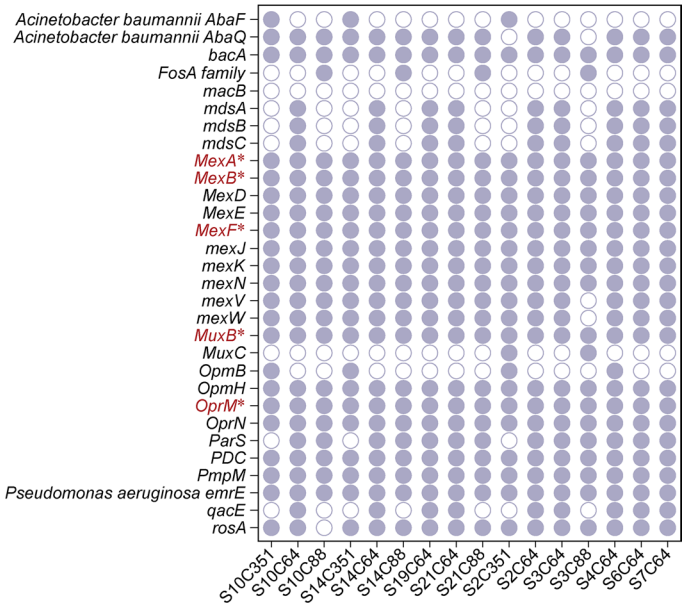

**Extended Data Fig. 5 | Host status of *Pseudomonas* by phages.** Heatmap showing the infection profiles of phages across the *Pseudomonas* genus.

https://doi.org/10.1038/s44284-026-00433-z



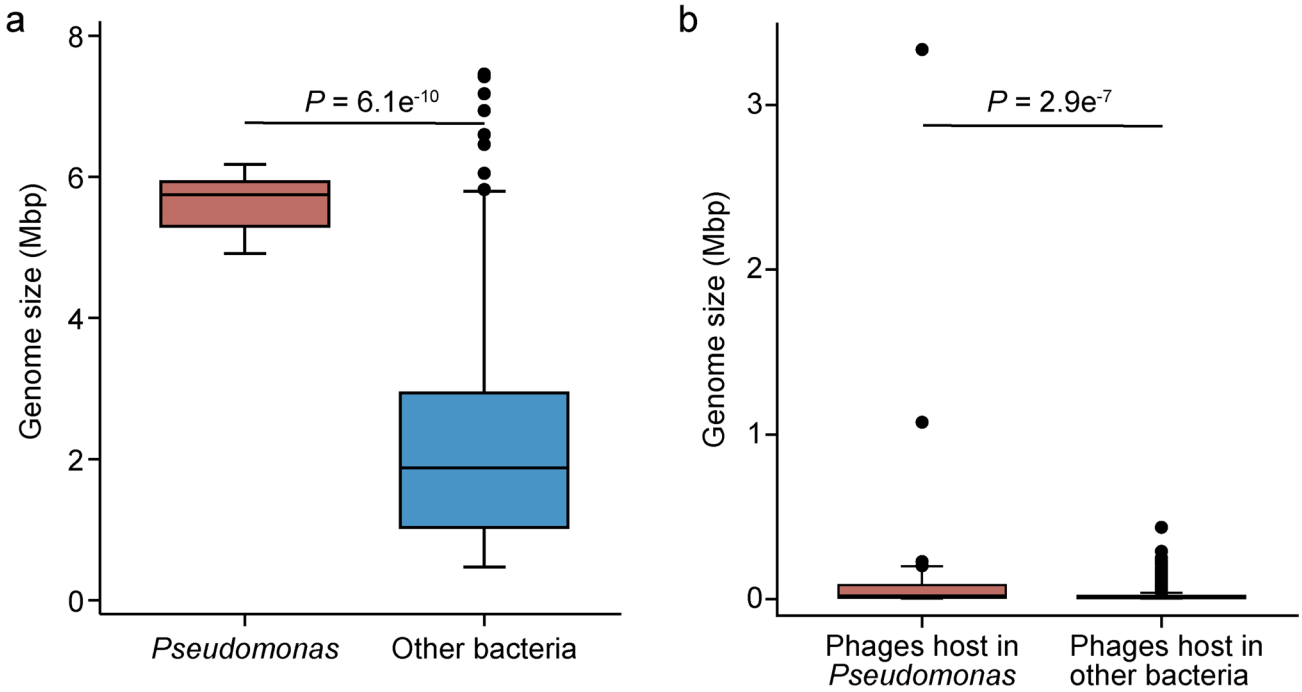

**Extended Data Fig. 6 | Interaction network between phages and their key hosts (*Pseudomonas* genus) in urban wetlands.** Green nodes represent bacterial hosts belonging to the *Pseudomonas* genus, while pink nodes represent phages. Node size indicates the degree of connectivity.

https://doi.org/10.1038/s44284-026-00433-z




**Extended Data Fig. 7 | The ARGs carry by *Pseudomonas*.** The asterisk and red mark indicating that these ARGs are also carried by phages.

https://doi.org/10.1038/s44284-026-00433-z




**Extended Data Fig. 8 | Traits of *Pseudomonas* and their associated phages.** (**a**) Comparison of genome sizes between *Pseudomonas* and other bacteria (n = 274); (**b**) Comparison of genome sizes between viruses hosted by the *Pseudomonas* and other phages (n = 1690). For box plots, the tops of the boxes represent the 75th percentile, the bottoms indicate the 25th percentile and the centre lines denote the median. The whiskers extend to the maximum and minimum non-outlier values. Since the data did not follow a normal distribution as shown in Supplementary Table 15, *P* values were calculated using two-sided Mann-Whitney U test adjusted with the false discovery rate (FDR).