nature portfolio

Peer Review File

Ecological Connectivity of Anthropogenic Antimicrobial Resistance Genes across Urban Water Systems and Drinking Water Risk

Corresponding Author: Professor Liping Ma

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
I was happy to receive the invitation to review this comprehensive and extremely data rich study of an urban water system resistome. I thank the authors for their attention to detail and their extensive efforts to generate and analyze the data.

Despite this, I do question the novelty of the study. While the scope is impressive, it seems to me not much new information was learned from the study. I try to address some areas where additional exploration could improve novelty. In addition, I find certain aspects of the paper are overstated and could use clarification in reporting and suggest some ways the language should be changed.

"487 Our uncovered genetic evidence of co-occurring ARGs and MGEs, and identified exact 488 HGT events through gene flow analysis, collectively demonstrate the critical role of MGEs in 489 contributing to the prosperity of ARGs throughout the UWS."

1. Here and throughout, its important to soften the language regarding HGT and to be a bit more specific about what you found. WAAFLE will provide evidence of HGT that occurred at some time prior to when you sampled; this could be because, simply, microbes which exchanged genes sometime in the distant past happen to co-exist in the same treatment plant. This is a very common issue with HGT from metagenomes.
a. In addition, detection of HGT as is done here typically cannot readily differentiate between alternative pathways.
b. For example,
Gene 1 is found in Taxa A, B, C, and D and is identical in all four taxa.
HGT prediction should support connections between some or all of the four taxa. But, the potential paths for transfer (which could have happened in the distant past or present) are A->B, A->C,A->D, etc. the implication of this is that it is not possible to conclude a real directionality without additional scrutiny and considering longitudinal variation (and even in this case, it is uncertain). Thus, saying “exact HGT events” is probably not accurate. Instead, I would reframe throughout to highlight that HGT could be occurring in the WWTP, but it is unclear when/where specifically the HGT occurred.
2. The nanopore sequencing of integron amplicons is really interesting. The results you display seem to indicate a high frequency of gene duplication in the cassette regions (Fig 5d, for example, many aadA are encoded in tandem). This is not something I had seen reported before—and in general—gene copy dynamics are poorly understood.
a. Can you estimate how common these duplications are? Which ARGs are most common?
b. One would also want to consider technical artefacts that might look like biology in this context and try to rule them out with additional analyses if possible.
c. Do you find these integrons on assembled plasmids or LFE isolate genomes?
d. Do you find these exact integron casettes downstream? (i.e., is it the same MGE-ARG combinations occurring throughout the system or are they different?)
3. For the HGT analysis, are there major donors or recipients for ARGs? You could try to estimate how central different taxa are as donors or recipients either through some network based analysis or by simply counting the number of times each taxa served as a donor or a recipient.
4. The discussion of quantitative risk should also be modified – it is too close currently to the framing of quantitative microbial risk assessment, which is a different class of analyses as the authors are aware.

5. In addition, I find the risk 1 ARGs to be a bit problematic given their dubious relevance to clinical antibiotic use. While I understand the scheme formerly proposed by Tong Zhang's lab, the distinction is very important considering the need to effectively communicate to (and not frighten) the public:
a. Three of the four risk 1 ARGs detected are related to membrane dynamics and/or peptide antibiotics, which are not extensively employed in humans to my knowledge. Instead, they are ARGs commonly detected in chiefly environmental taxa. My interpretation of this is that the analysis may be limited by the ability to assemble/bin/taxonomically classify some of the prevalent clinically relevant ARG scaffolds in wastewater. Often, these are encoded on plasmids or other MGEs which can be extremely difficult to bin into MAGs/may lack sufficient phylogenetic signal to taxonomically classify. Could the authors check for clinically relevant ARGs in unbinned contigs? E.g., CTX-M, NDM, PER, other ESBLs? This might give a sense of what is being missed in their workflow and not appearing in their highest risk ARGs.
b. fosA might be appropriately labeled a high risk ARG.

Minor comments
Please report in the text or supplementary the details around the construction of the integrase database. These genes are sometimes tricky to annotate given that, in some cases, they are only differentiable on the basis of a NT 15aa arm domain which is not in PFAM.
What is meant by "diversity of rearrangements" ? in Figure 5
There needs to be more reporting of assembly completeness- size, contiguity, etc. These aspects are important for being able to detect HGT.
In addition, a correlation between assembly N50 and/or completeness with HGT event rates could demonstrate independence between assembly completeness and the frequency of detected HGTs-

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)
This study by Liu et al. investigates antimicrobial resistance in four urban water systems, including wastewater treatment plants, receiving river water, drinking water reservoirs, and tap water. The objective is to explore the potential connectivity of antimicrobial resistance across these systems. The authors collected 32 water samples and performed several analyses, including metagenomic sequencing, culturing and genome sequencing of lactose-fermenting Enterobacteriaceae, 16S rRNA gene amplicon sequencing, and chemical analysis of water quality.

The study focuses on identifying antimicrobial resistance genes to infer resistance patterns within and across the sampled systems. The authors also attempt to evaluate potential risks using a risk score they developed. While the research question is relevant and timely, there are substantial concerns about the validity and interpretation of the results. The findings are mainly descriptive and lack mechanistic insight, offering limited advancement beyond existing studies that have already examined antimicrobial resistance in each of these environments. The investigation into connectivity, which could have been the most valuable contribution of the study, is not effectively addressed. The data from different systems are presented separately, and the suggested connections are speculative and insufficiently supported by analysis.

This reviewer identifies three major flaws in the study:

1. The Methods section states that wastewater effluent is discharged into the receiving river, suggesting a direct connection between these two systems. The authors also analyzed water from two reservoirs and tap water; however, no samples were collected from drinking water treatment plants. It is assumed that reservoir water undergoes treatment before distribution. One important gap in the study's logic is the lack of a clearly established hydrological connection between the receiving river and the reservoirs. Without this connection, inferences about AMR transmission across these systems are not scientifically justified. Furthermore, the absence of data from the drinking water treatment process limits the ability to assess how treatment may influence AMR levels. Disinfection steps such as chlorination or UV treatment can significantly reduce microbial load and potentially remove ARGs. Additionally, tap water is not an ideal proxy for treated water, as household plumbing can alter the microbial profile through biofilm formation and water stagnation. These limitations in the sampling design significantly weaken the central claim of AMR connectivity across the four urban water systems. As currently presented, the study does not provide sufficient evidence to support this conclusion.

2. A second major concern is that, unlike wastewater treatment plants, the receiving river and reservoirs are open systems that likely receive inputs from a variety of sources, including urban runoff and other non-point pollution. Consequently, it is not possible to attribute the presence of AMR exclusively to discharged wastewater. Moreover, key factors such as water residence time, environmental conditions, and the growth or decay of bacteria carrying ARGs are likely to influence AMR dynamics in these systems. These factors are not considered in the study. By oversimplifying the complexity of the aquatic environment, the study draws conclusions that are premature and not adequately supported by the data.

3. The authors state that samples were collected in June 2023; however, it is unclear whether the sampling strategy was designed to follow the same parcel of water through the different components of the urban water system. Please provide a clearer description of the timeline and coordination of the 32 sample collections. Without this information, it is difficult to assess whether the data represent a continuous flow through the system or isolated, independent snapshots. This lack of temporal alignment raises concerns about the validity of the conclusions regarding connectivity among the water systems.

Other comments:

1. The title implies that the study addresses mechanisms and mitigation strategies, but neither is clearly presented. It is unclear what mechanisms are being referred to, and no specific mitigation approach is described. Furthermore, the claim of amplified health risks appears overstated, as the risk assessment is based solely on bioinformatic predictions without experimental validation or supporting measured data.

2. The Introduction is relatively general and would benefit from greater specificity. In particular, it would be helpful for the authors to clarify whether there have been prior efforts to investigate the fate and transport of AMR across different water compartments, whether in natural or urban systems.

3. Line 107. The authors describe the study as a case study, and the manuscript is written as such. However, this framing limits the broader scientific value of the work. The manuscript lacks a discussion of how the findings contribute to generalizable knowledge or advance the broader understanding of AMR dynamics in urban water systems.

4. Line 132. It is unclear what the underlying assumption is when representing the number of ARG copies per cell. Since microbial communities are heterogeneous, with some cells carrying multiple ARGs and others carrying none, it is important to explain what this metric represents and how it should be interpreted.

5. Line 149. Please clarify what constitutes a “strain” of lactose-fermenting Enterobacteriaceae in this study. Is this classification based on genetic criteria such as average nucleotide identity (ANI), or another method? This point is not clearly described in the Methods. Additionally, among the 3,040 isolated strains, how many are genetically unique?

6. Line 240. The use of 16S rRNA gene amplicon data to associate microbial taxa with ARGs provides only correlative evidence and does not allow for reliable identification of true ARG hosts. This method does not confirm whether the taxa are actually carrying the ARGs, and therefore referring to them as “hosts” is both inaccurate and potentially misleading.

7. Line 262. While the use of MAGs to link ARGs to potential hosts is a more robust approach, it is unclear whether dereplication of MAGs was performed, as this step is not explicitly mentioned in the Methods. Dereplication is important to eliminate redundancy and ensure that the analysis is based on representative, non-redundant genomes. Without this step, the inclusion of closely related or duplicate MAGs could bias the results and lead to inflated or misleading interpretations.

8. Line 319. The manuscript would benefit from stronger evidence linking ARGs to the measured antibiotics. For instance, is there a correlation between tetracycline concentrations and the abundance of the tetA gene? According to the data in Line 331, tetA does not appear to correlate with tetracycline levels.

9. Line 350. The Methods and Results sections do not clearly articulate the novelty of the health risk assessment used in this study. It is unclear how this approach differs from existing methods, or how its novelty is reflected in the equation presented in Line 381.

10. Line 393. Is there a theoretical basis for the risk-based prioritization described in this section? The criteria used do not appear to be grounded in an established theoretical framework, and the rationale behind their selection is not clearly explained.

11. The Discussion section largely restates the results without providing meaningful mechanistic interpretation. Strengthening this section with deeper analysis and a clearer theoretical framework is necessary to improve the scientific rigor of the study. As noted above, this is essential for moving the work beyond a descriptive case study and highlighting its broader relevance and contribution to the field.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)
The authors thoroughly investigate the prevalence of ARGs across the urban water system from sewage to drinking water. While a case study, the intense interrogation of 32 water samples across this spectrum using 16S rRNA sequencing, metagenomics, and culturomics, reveals that the some treatment steps may select for ARGs. Several types of ARGs were also uniquely appearing only after treatment. The enrichment of several type of ARGs could be attributed to both biotic and abiotic factors, and drivers of the resistome were identified. The most concerning ARGs were also identified, setting a potential framework for ARG monitoring in urban water streams. This is a valuable study, especially as ARG monitoring is of increasing popularity. However, some of the conclusions may be overstated, and some methods can be clarified.

Relative vs Absolute:
The authors rely on quantification of ARGs in copies/cell, which is a measure of relative abundance. This metric does not take into account the absolute quantification of cells or ARGs in each water tyle (e.g., copies/mL, cells/mL of water), and thus misses a major contextual point that needs to be discussed around water treatment. Do the authors have the ability to make such quantifications? If yes, it is helpful to include this context, even as a separate discussion, with enrichment amongst surviving cells (relative abundance) and overall exposure reduction (absolute abundance) as unique topics. Readers may

take away from the manuscript that exposure to ARGs in untreated (or non-disinfected) water is less than that of treated/disinfected water, which may not be true from the perspective of total exposure, given the orders of magnitude reduction in cell counts during treatment. When discussing risk, it is paramount to include absolute quantification.

In defending the measure ARG copies/cell, the authors cite a study which also states that “Both absolute and relative measures are essential for a comprehensive evaluation of antibiotic resistance in environmental settings and provide complimentary insights, e.g., regarding mass flows and exposure risks, on one hand, and processes of resistance selection and comparison of widely different environments (e.g., dense gut microbiomes versus water), on the other.” This study also mentions that transformation is easy, and thus, presenting transformed data (e.g., even roughly by multiplying by 16S rRNA gene copies/L) should be feasible. As raw wastewater and clean drinking water are vastly different in absolute density, it is critical to present absolute measures to support the broad and, at present, overstated conclusions

The inclusion of this data will verify that water treatment is working as intended across each treatment step, and will help some of the discussion of results make more sense. While it is understandable that both drinking and wastewater treatment select for microbes that are resistant to disinfectants and antibiotics via relative abundance (by removing less resistant cells), the actual concentration of total cells should concurrently decrease, and thus total potential exposure to ARGs may also decrease. Further, the increase in resistance between treatment and tap is potentially even more concerning, as this likely accompanies an increase in total cell concentration.

As the authors only present relative abundance, care should also be taken discussion of increases in relative abundance vs that of absolute abundance. For example L226-228 “... with its abundance further elevated by 7.0 fold following drinking water disinfection” is referring to an increase in relative abundance, and thus proportion or prevalence may be a better word than abundance. In another example L 260 – “...between HGT event rates and ARG abundance across samples” – this is ARG prevalence or ARG relative abundance. This nuance must be cared for throughout the text, as it was confusing to read as is.

L 234 – can the authors expand on what they mean by etc.? Typically an e.g., for example list would not include this denotation.

L 323 – Define BAC, PCMX at first mention in the text. Please note that BAC can be confused for biological activated carbon in the drinking water context.

L357 – “First ARGs...drinking water.” This sentence can likely be rearranged and/or shortened to make more sense. For example...: We defined sewage-derived ARGs as those that are discharged into natural waters through effluent, are subsequently enriched in the receiving waters, and ultimately persist into tap water. We used the following criteria:...”. The first sentences of each of these sections have similar issues with long sentences that can be made more direct.

L362 – can the authors define a what “detectable” in household drinking water means? Given the unique collection methods from this study (>2,000 L) compared to typical drinking water monitoring methods (with defined volumes that are often lower), is it likely that ARGs are often undetectable?

L370 – “...based on which to develop...” should potentially be “which was used to develop”

L 381... - Does the quantitative health risk assessment take into account any absolute quantification, or only relative? The equation seems to indicate relative abundances, but the figure does not specify. If only relative, is this typical? Is exposure taken into account? (e.g., how are people typically exposed to drinking water vs wastewater [is this only workers]??) Further, can the authors expand on the pathogens in drinking water that were found and at what concentrations? More discussion is likely needed on this QMRA metric.

L390 – “While subsequent chlorination...” – is this mitigation of risks due to the overall reduction in cell counts? Is the conclusion ultimately that treatment is effective at reducing risks? If yes, this may need to be emphasized in final discussion/conclusions.

L408 – a subsequent 3.3-fold increase in tap water relative to source drinking water – is this an increase in relative abundance? Or does absolute concentration also increase?

L454-L470 – can the authors comment on the prevalence of pathogens in this drinking water? The authors at some point mention E. coli and P. aeruginosa, but often in developed countries’ urban water systems, E. coli is often no longer prevalent in treated (disinfected) water, and P. aeruginosa is inhaled as a primary infection route, which may change ARG implications contrary to the discussion. If invoking consumption of pathogens containing ARGs as a risk, it is important to also divulge the quantification of pathogens in this water system, as it may be unique. The authors further mention boiling water, which is not often a requirement in many developed countries. The level of disinfection, and maintenance of disinfectant needs to be further explained, as this context may have strong implications on the results. This is also where absolute quantification of cells will help put these results into a larger context.

Methods:
Can the authors clarify the quantity of samples from each location? From descriptions, it seems like 32 samples were analyzed – but n is difficult to follow. Figure 1 may be an opportunity to clarify n.

Can the authors clarify the use of a filtration device for concentration of over 2,000 L? Was this device an R.O. membrane, or a cartridge filtration device? For either option, what was the time frame, and did these samples potentially enrich for filter-living microbes? How did the authors quantify total bacteria present in these samples, and can that be normalized to a water basis?

L583 – Tap waters were “designed”? Word choice here is likely wrong.

Figure 1a: Defining n for samples may also be helpful in this figure.

Figure 2: please define n clearly.
Figure 2a – Can the authors expand on the meaning of: the ellipses (are these statistically determined), the numbers in circles, and arrows in the caption? As this figure is busy, it is suggested to remove the numbers in circles, which could be confused for another sample type, and potentially vary symbols to offer more distinction (e.g., squares for WW samples, circles for rivers, triangles for drinking water), which could also reinforce the classes used throughout.
Figure 2b/c – The establishment of baseline is somewhat unclear for the arrowed bars. It is suggested to add ends to the bars, or, better yet, define the range for the change with text (e.g. RU->RS, 8.4%^). Is there a reason that for RD the axis color changes (2c)?

Figure 7/b: It may be useful to clarify which of these inputs was theoretical vs from data in the paper.

(Remarks on code availability)
Code not analyzed.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Thank you to the authors for their thorough and detailed response to my review. I feel they have sufficiently addressed my concerns.

Additional comments regarding the previous comments of reviewer #3:
The reviewer correctly points out that reduced cell density (eg in tap water vs source water) can likely influence the “risk” interpretations. The authors again estimate the absolute abundance of the pathogens in drinking water and find low concentrations (~3 cells per 100L) overall. The relevant response:

Absolute quantification of pathogens in drinking water. In response to the reviewer’s Comment #9, in addition to metagenomic relative quantification, we further estimated the absolute concentrations as suggested. Four bacterial species were detected, including Acinetobacter johnsonii, Escherichia coli, Mycobacterium phocaicum, and Pseudomonas aeruginosa, with at least one pathogenic species present in each sample. The average total concentration of these pathogens in tap water was 2.67 cells/100 L, and all carried diverse ARGs. Notably, P. aeruginosa, which harbored resistance genes to five antibiotic classes, was found in all samples.

The major finding of the study regards the ecological connectivity between the different compartments, so I think the read mapping here is worthy of some scrutiny. Namely- what are you actually detecting in drinking water?
At such low concentrations, and with the high detection limit of metagenomics, it is quite possible that reads aligning to related bacterial species are being erroneously mapped to the MAGs constructed.

First, the detection of E. coli in drinking water is concerning and a potential indication that spurious read mapping is occurring.

Second, there is no reason to suspect that the same P. aeruginosa strain(s) represented by the MAG would (1) be found in all compartments or (2) retain the same ARGs throughout. It is not generally possible from the approach taken here (read mapping to 95% ANI dereplicated MAGs) to differentiate populations of PA or closely related species in the different compartments.
For example, premise plumbing or distribution system biofilms could plausibly harbor stable and phenotypically distinct populations of PA. Therefore, they do not really prove connectivity through this method at least for drinking water.

In sum, it would be good to have additional details regarding
1) The read mapping strategy, especially whether read mapping was performed by aligning short reads to a combined index (so that reads are aligned against the pathogens AND alternative references which might be better matches- see item 2 in this link: https://instrain.readthedocs.io/en/latest/important_concepts.html)

2) The breadth of coverage for each pathogen genome in the drinking water samples.
3) If the authors would like to support the notion that the MDR PA is found in drinking water, some indication that the ARGs and/or relevant MGEs are present in that compartment.
Last, it would be good to be explicit about the impact of absolute abundance on the conclusions. I noticed the abstract didn't change much, despite that much of the framing has theoretically shifted given the drastic decrease in absolute abundances of ARGs.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)
The revised manuscript has improved substantially and addresses many of the concerns raised in the previous round, which this reviewer appreciates. While most points have been clarified, two issues still require attention because they are central to the manuscript's main claims and interpretation.

First, the claim that the different water systems are hydrologically connected, despite lacking physical connectivity, needs clearer justification. To support this, the authors conducted SourceTracker analyses and examined shared ARG subtypes, reporting substantial resistome overlap. However, this evidence alone does not necessarily demonstrate hydrologic connectivity or transfer between systems. For example, river water and reservoir source water are not physically connected, so shared ARGs could reflect widespread background occurrence, common upstream sources, or similar selective pressures rather than direct exchange. Because this point is fundamental to the study, the authors should strengthen and clarify the inference and explicitly distinguish resistome similarity from physical transfer.

Second, additional clarification is needed regarding the fate of ARGs in open systems such as rivers and reservoirs. Although the authors considered multiple factors in the Mantel tests and the PLS-SEM model, key confounders related to loss processes appear to be missing. In addition to host proliferation, it is important to consider bacterial mortality in the water column and decay of extracellular DNA or ARG signals due to natural processes (e.g., dilution, sunlight exposure, temperature, predation, settling). These processes should be incorporated where feasible or, at minimum, discussed explicitly as limitations. Without addressing them, the interpretation of ARG persistence and transport in open systems remains incomplete.

Additional comments regarding previous comments of reviewer #3:
1.) Overall, the authors have largely addressed Reviewer #3's comments and revised the manuscript accordingly.

2.) With respect to Comment 14 (filtration method), I believe the authors should provide additional clarification and, where possible, further justification. Reviewer #3's central concern relates to whether the sampled microbial community is representative of the water column (i.e., source water) or whether the approach may bias the sample toward organisms associated with filters.

Although I agree with the authors that substantive microbial growth on the filters during sampling is unlikely under typical conditions, the reported 24-hour sampling period still raises two related questions:

a.) Potential contribution from distribution system biofilms/pipe-associated communities: Over a prolonged sampling duration, the collected material may reflect not only microorganisms suspended in the water source but also organisms intermittently shed from pipes or other infrastructure, which could affect representativeness.

b.) Comparability across sampling locations and study inferences about connectivity: If other matrices in the study (e.g., wastewater treatment plant effluent, river water, source water) were collected as grab samples or over much shorter time windows, the markedly different temporal integration (24-hour composite vs. grab/short-duration sampling) could introduce a mismatch in sampling timescales. Given that the manuscript emphasizes ecological connectivity across compartments, the authors should more explicitly discuss how differences in sampling duration may influence comparability and interpretation of connectivity signals.

In my view, the response and the revised manuscript would be strengthened by a clearer explanation of: (i) the rationale for the 24-hour filtration duration; (ii) the extent to which this approach is expected to capture water-column communities versus potential distribution-system contributions; and (iii) how differences in temporal integration across sample types are addressed analytically or, at minimum, explicitly acknowledged as a limitation.

(Remarks on code availability)
N/A

Version 2:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
I thank the authors for their exemplary response to my comments and feel they have addressed my concerns sufficiently.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)
I appreciate the authors' efforts in revising the manuscript to address the comments raised.

(Remarks on code availability)

Responses to the Comments of Reviewers

Dear Reviewers,

Thank you for your precious comments on our paper entitled **"Ecological Connectivity of Anthropogenic Antimicrobial Resistance in Urban Water Systems and Its Implications for Drinking Water Risk Assessment"** (Manuscript Number: **NCOMMS-25-80871**). We have modified the manuscript according to your comments. Please see the revision marked in red in the revised version of the manuscript and below the point-by-point responses to the comments.

Response to Reviewers' Comments

Reviewer #1:

I was happy to receive the invitation to review this comprehensive and extremely data rich study of an urban water system resistome. I thank the authors for their attention to detail and their extensive efforts to generate and analyze the data.
Despite this, I do question the novelty of the study. While the scope is impressive, it seems to me not much new information was learned from the study. I try to address some areas where additional exploration could improve novelty. In addition, I find certain aspects of the paper are overstated and could use clarification in reporting and suggest some ways the language should be changed.

Response: We sincerely thank the reviewer for thorough review of the manuscript and for providing insightful and constructive comments. We also appreciate the positive assessment of the study's scope and acknowledgment of the considerable effort invested in generating and rigorously analyzing this comprehensive dataset.

According to the reviewer's suggestions, we have performed additional analyses and substantially revised the manuscript to more clearly articulate its novel contributions. Specifically, we (*i*) expanded the quantitative assessment of integron cassette duplication patterns, (*ii*) refined the HGT-related analyses, and (*iii*) revised several conclusions to avoid overstatement. These revisions strengthen both the methodological rigor and the scholarly presentation of this study, thereby clarifying its substantive contribution to the field. Point-by-point revisions are provided below in response to the reviewer's specific comments.

Together, this study provides the first systematic investigation of anthropogenic AMR migration across interconnected urban water compartments in a megacity. It elucidates the mechanisms of AMR transmission enabled by ecological connectivity, identifies sewage-derived ARGs that persist throughout the system, and offers a scientific foundation for source-oriented interventions to disrupt critical transmission links and mitigate associated health risks in drinking water.

1. *Line 487, "Our uncovered genetic evidence of co-occurring ARGs and MGEs, and identified exact HGT events through gene flow analysis, collectively demonstrate the*

critical role of MGEs in contributing to the prosperity of ARGs throughout the UWS." Here and throughout, its important to soften the language regarding HGT and to be a bit more specific about what you found. WAAFLE will provide evidence of HGT that occurred at some time prior to when you sampled; this could be because, simply, microbes which exchanged genes sometime in the distant past happen to co-exist in the same treatment plant. This is a very common issue with HGT from metagenomes.
a. In addition, detection of HGT as is done here typically cannot readily differentiate between alternative pathways.
b. For example,
Gene 1 is found in Taxa A, B, C, and D and is identical in all four taxa.
HGT prediction should support connections between some or all of the four taxa. But, the potential paths for transfer (which could have happened in the distant past or present) are A->B, A->C, A->D, etc. the implication of this is that it is not possible to conclude a real directionality without additional scrutiny and considering longitudinal variation (and even in this case, it is uncertain). Thus, saying "exact HGT events" is probably not accurate. Instead, I would reframe throughout to highlight that HGT could be occurring in the WWTP, but it is unclear when/where specifically the HGT occurred.

Response: Thanks for the thoughtful, detailed, and constructive comments. We fully agree with the reviewer that WAAFLE and associated similarity-based HGT events identification tools provide signatures of HGT that occurred at some time point prior to sampling, and it is unclear when/where specifically the HGT occurred. As noted by the reviewer, gene flow analysis cannot reliably resolve directionality or specific transmission pathways (e.g., donor-recipient relationships among multiple potential hosts), nor distinguish between alternative routes or multi-step transfer scenarios.

To address this concern, we have revised the language throughout the manuscript to be more precise. Specifically, "exact HGT events" in the Discussion (as noted by the reviewer) and other instances of "HGT events" have been changed to "putative HGT events" (**Lines 293**, **296**, **488**). Furthermore, we have added a statement to the Discussion acknowledging these methodological limitations: "the precise timing, spatial origin, and molecular pathways of these HGT events remain unresolved, owing to inherent limitations in current similarity-based approaches for HGT inference" (**Lines 489**–**491**).

Lines 292-295: "Further contig-based genetic full-scanning captured a total of 4,097 putative HGT events occurring among bacteria (Fig. 5a), and results indicated the natural aquatic environment provides a relatively favorable platform for HGTs."
Lines 295-297: "Further correlation analysis revealed a significant positive relationship between putative HGT event rates and ARG abundance across samples ($r^2 = 0.74$, $p < 0.001$), indicating the patterns of MGEs-driven resistome dynamics (Fig. 5b)."
Lines 487-491: "Our uncovered genetic co-occurrence of ARGs and MGEs, along with putative HGT events inferred from gene flow analysis, underscores the critical role of MGEs in contributing to the prosperity of ARGs throughout the UWS. However, the precise timing, spatial origin, and molecular pathways of these HGT events remain unresolved, owing to

inherent limitations in current similarity-based approaches for HGT inference."

2. *The nanopore sequencing of integron amplicons is really interesting. The results you display seem to indicate a high frequency of gene duplication in the cassette regions (Fig 5d, for example, many aadA are encoded in tandem). This is not something I had seen reported before—and in general—gene copy dynamics are poorly understood.*

Response: Thanks for highlighting this observation and for recognizing the potential novelty of our findings on tandem of resistance gene cassette duplications revealed by Nanopore sequencing. We completely agree that the dynamics of gene copy number variation within integron cassette arrays remain poorly characterized, and we appreciate the opportunity to further clarify and strengthen this aspect of the study in response to your insightful comment.

a. Can you estimate how common these duplications are? Which ARGs are most common?

To address this issue, we performed additional quantitative analyses of the integron amplicon sequencing data to estimate the prevalence of tandem cassette duplications. To ensure the reliability of these analyses, only open reading frames (ORFs) exhibiting intact start and stop codons and exceeding a predefined minimum coding sequence length were retained. Our analysis revealed that 6.48% ± 0.35% of all integron cassette arrays in each sample contained at least one tandemly duplicated ARG. This proportion increased with the number of ARGs harbored within each integron cassette array. The dominant ARGs in tandem is *aad*A2, followed by *aad*A24 and *aad*A5, conferring resistance to aminoglycosides.

Accordingly, the additional analyses results were supplemented in the revised manuscript (**Lines 316–318**).

Lines 316-318: "Notably, among all integron cassette arrays, 6.5% contained at least one tandemly duplicated ARG, predominantly *aad*A2, *aad*A24, and *aad*A5."

b. One would also want to consider technical artefacts that might look like biology in this context and try to rule them out with additional analyses if possible.

To rigorously exclude technical artefacts, we implemented a multi-tiered quality control strategy encompassing experimental, sequencing, and bioinformatic stages: **(1) Experimental validation**: In addition to strict adherence to standardized experimental protocols, PCR amplicons were verified by agarose gel electrophoresis and subsequently purified using a commercial silica-membrane-based DNA purification kit (QIAGEN). **(2) Stringent sequencing quality control**: Nanopore sequencing was conducted with rigorous library preparation and base calling. Quality control of Nanopore raw reads was performed using chopper (v0.8.0) [1]. Low-quality reads, chimeric sequences, and reads with excessive errors were removed prior to downstream analyses to minimize sequencing-induced mismatches and artefactual variants. **(3) Conservative gene prediction**: ORFs

were identified using Prodigal (v2.6.3)[2] with stringent criteria. Only ORFs exhibiting intact start and stop codons, consistent coding frames, and high confidence scores were retained; sequences flagged as partial or truncated were excluded to eliminate artefacts arising from fragmented DNA templates. Collectively, these measures systematically minimized the impact of potential errors introduced during experimental handling, sequencing, and annotation.

Consistent with our findings, the biological plausibility of tandem ARG duplications is well documented, with examples including tandem *aad*A2 genes within an integron in an *Escherichia coli* isolate[3] and, more broadly, tandem ARGs predominantly from the *aad*A and *dfr*A families identified on integrons in urban sewage through PCR combined with third-generation sequencing[4,5]. This structural arrangement serves as an important mechanism for ARG enrichment and is mediated by IntI integrase through a "cut-paste-reinsert" model of site-specific recombination[6,7].

For completeness, the relevant technical details have now been explicitly described in the revised **Supplementary Methods 5**.

> **Supplementary Methods Lines 186-198**: "Following amplification, a series of quality control steps were implemented to ensure the reliability of integron gene cassette sequences prior to downstream analyses. First, PCR amplicons were examined by agarose gel electrophoresis to verify amplification specificity and confirm expected fragment sizes, and subsequently purified using a commercial silica membrane-based DNA purification kit (QIAGEN). The purified amplicons were then subjected to Nanopore sequencing following a stringent workflow for library preparation, sequencing, and quality control. Low-quality reads, chimeric sequences, and reads with excessive base-calling errors were filtered using chopper (v0.8.0)[12] prior to further analyses. Open reading frames (ORFs) were predicted using Prodigal (v2.6.3), with only complete and high confidence gene predictions retained. ORFs lacking intact start or stop codons, or those flagged as partial or truncated, were excluded to minimize artefacts arising from fragmented DNA sequences."

Reference:

[1] De Coster W, Rademakers R. NanoPack2: population-scale evaluation of long-read sequencing data. *Bioinformatics* 39, btad311 (2023).

[2] Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinform* 11, 119 (2010).

[3] An R, Qi Y, Zhang XX, Ma L. Xenogenetic evolutionary of integrons promotes the environmental pollution of antibiotic resistance genes — Challenges, progress and prospects. *Water Res* 231, 119629 (2023).

[4] An XL, Abass OK, Zhao CX, Xu MR, Pan T, Pu Q, et al. Nanopore sequencing analysis of integron gene cassettes in sewages and soils. *Sci Total Environ* 817, 152766 (2022).

[5] Yang Y, Zhang A, Che Y, Liu L, Deng Y, Zhang T. Underrepresented high diversity of class 1 integrons in the environment uncovered by PacBio sequencing using a new primer. *Sci Total Environ* 787, 147611 (2021).

[6] Collis CM and Hall RM. Site-specific deletion and rearrangement of integron insert genes catalyzed by the integron DNA integrase. *J Bacteriol* 174, 1574-1585 (1992).

[7] Escudero JA, Loot C, Nivina A, Mazel D. The integron: Adaptation on demand. *Microbiol Spectr*

3, MDNA3-0019-2014 (2015).

c. Do you find these integrons on assembled plasmids or LFE isolate genomes?

For assembled plasmids. Accordingly, assembled metagenomics contigs obtained using MEGAHIT (v1.1.3)[1] were classified as plasmid-derived using PlasFlow (v1.1)[2] with default parameters. All ORFs from these putative plasmid contigs were subsequently extracted and annotated for the presence of the integron marker gene *int*I1 and ARGs harbored. As a result, no ARG-carrying integrons were detected on these assembled plasmid sequences, a finding that may be attributable to the inherent limitations of short-read metagenomic assembly. In the future, complete plasmid genomes could be obtained through plasmid extraction combined with hybrid assembly of second- and third-generation sequencing data, thereby enabling definitive identification.

For LFE isolate genomes. Based on LFE isolates (3,040 strains) combined with PCR amplification of integron variable regions followed by Nanopore long-read sequencing, we obtained a comprehensive profile of integron cassette arrays present in LFE isolate genomes. Overall, 5.19% ± 0.78% of integron cassette arrays in the LFE contained at least one tandemly duplicated ARG, predominantly *aad*A, *cat*B, and *dfr*A, which confer resistance to aminoglycosides, chloramphenicol, and trimethoprim, respectively. Following the reviewer's suggestion, we assembled the quality-controlled LFE metagenomic datasets using MEGAHIT. All ORFs predicted from the assembled contigs were extracted and annotated to identify the integron marker gene *int*I1 and associated ARGs. Unlike the findings from amplicons, no tandemly duplicated ARGs were identified in the metagenomic data, although arrays with repeated ARGs (e.g., *dfr*A14–*aad*A–*cml*A1–*aad*A–*dfr*A27, containing a duplicated *aad*A) were observed. This difference is likely due to the inherent limitations of metagenomic assembly efficiency.

We thank the reviewer for raising these important points. We recognize that future research determining whether integrons are plasmid-borne and elucidating their association with potential pathogens is critical for a comprehensive understanding of integron mobility and dissemination potential. The corresponding discussion has now been incorporated into the revised manuscript (**Lines 510–514**).

> **Lines 510-514**: "Consequently, the xenogenetic evolution of ARG-associated integrons within the UWS, particularly mediated by selective pressures from physicochemical parameters and pollutants, plasmids-borne localization, and carriage by potential pathogens, necessitates investigation beyond conventional *int*I gene-based quantitation."

Reference:

[1] Li DH, Liu CM, Luo RB, Sadakane K, Lam TW. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly *via* succinct graph. *Bioinformatics* 31, 1674-1676 (2015).

[2] Krawczyk PS, Lipinski L, Dziembowski A. PlasFlow: Predicting plasmid sequences in metagenomic data using genome signatures. *Nucleic Acids Res* 46, e35 (2018).

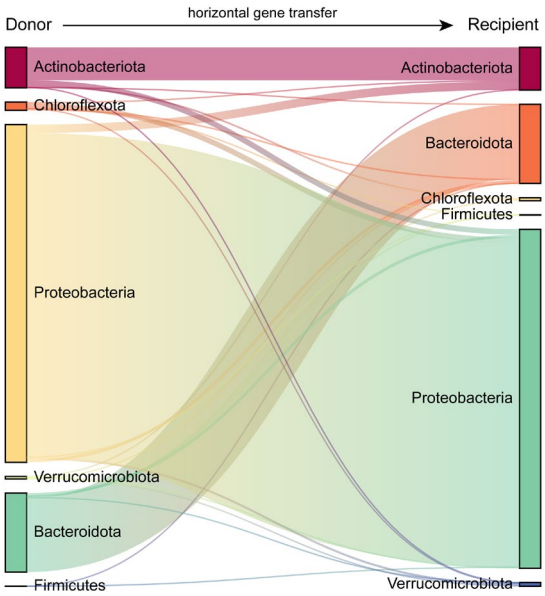
d. Do you find these exact integron cassettes downstream? (i.e., is it the same MGE-ARG combinations occurring throughout the system or are they different?)

We re-examined all identified integrons carrying ARG-associated cassettes across water compartments, with particular attention to those harboring duplicated ARGs. As a result, diverse ARG arrays were detected not only downstream but also throughout the system (e.g., *aad*A–*aad*A24–*aad*A22, *aac*(6')-Ib′–OXA-1, and *aad*A2–*aad*A25–*dfr*A12–*dfr*A12), while "*arr*–*aad*A–*cat*B–OXA-1–*aac*(6')-I–*aac*(6')-I" emerged downstream. This site-specific occurrence could be attributed to the intrinsic mechanism of integrons to capture exogenous ARGs, which contributes to the emergence of multidrug-resistant bacteria and is potentially shaped by aquatic environmental factors and microbial communities. We sincerely thank the reviewer for the insightful perspective, and the findings have been integrated into the revised manuscript (**Lines 508–510**).

> **Lines 508-510**: "The observation of both conserved and compartment-specific integron-associated ARG arrays across the UWS indicates concurrent downstream dissemination of stable MGE–ARG combinations and local cassette rearrangements mediated by HGT."

3. *For the HGT analysis, are there major donors or recipients for ARGs? You could try to estimate how central different taxa are as donors or recipients either through some network based analysis or by simply counting the number of times each taxa served as a donor or a recipient.*

Response: We thank the reviewer for this constructive suggestion regarding the identification of major donors and recipients in the HGT analysis. To address this point, we leveraged MetaCHIP (v1.10.6)[1], which infers putative donor-recipient relationships for HGT events based on a combination of phylogenetic incongruence, sequence similarity, and genomic context. As noted in the reviewer's comment #1, the inferred directionality is informed by evolutionary and comparative genomic signals, and does not imply precise temporal or spatial resolution of the HGT events. Using these directional assignments, we summarized HGT events across major bacterial phyla and visualized the results using a directional Sankey diagram. Six distinct bacterial phyla were implicated in HGT events. Among these, phyla *Proteobacteria*, *Bacteroidota*, and *Actinobacteriota* served as both primary donors and recipients of ARGs in the urban water system. Transfer events occurred both within and across these phyla. For optimization, the figure has been supplemented as **Supplementary Fig. 14**, and the corresponding methodology and conclusions have been integrated into the revised manuscript (**Lines 678–680**, **301–303**).



Supplementary Fig. 14 | Phylum-level summary of inferred donor-recipient relationships in ARG-associated horizontal gene transfer. Directional Sankey diagram summarizing horizontal gene transfer events inferred by MetaCHIP at the phylum level. Each flow represents inferred HGT events from donor phyla (left) to recipient phyla (right), with link widths proportional to the number of detected events.

Lines 678-680: "MetaCHIP (v1.10.6)[72] was used to identify the primary bacterial taxa serving as donors and recipients in ARG-associated HGT events."

Lines 301-303: "These putative HGT events predominantly involved six bacterial phyla, mainly *Proteobacteria*, *Bacteroidota*, and *Actinobacteriota*, with transfers occurring both within and across phyla (Supplementary Fig. 14)."

Reference:

[1] Song W, Wemheuer B, Zhang S, Steensen K, Thomas T. MetaCHIP: Community-level horizontal gene transfer identification through the combination of best-match and phylogenetic approaches. *Microbiome* 7, 36 (2019).

4. The discussion of quantitative risk should also be modified – it is too close currently to the framing of quantitative microbial risk assessment, which is a different class of analyses as the authors are aware.

Response: Thanks for this thoughtful comment. As suggested, we have revised our discussion of risk assessment to clearly distinguish it from quantitative microbial risk assessment (QMRA). Our research provides a genome-informed, system-scale comparative risk indicator based on 61 sewage-derived ARGs identified in our dataset that evaluates the relative hazard associated with sewage-derived resistomes across urban water compartments, by incorporating three genetic risk determinants: (*i*) the relative abundance of ARGs, (*ii*) the co-occurrence of ARGs with mobile genetic elements (MGEs) as an indicator of horizontal transfer potential, and (*iii*) the association of ARGs with pathogenic bacterial hosts. These revisions are detailed in **Lines 382–391** of the revised manuscript.

Lines 382-391: "**Sewage-derived ARG-associated health risk assessment.** To evaluate the health hazards associated with sewage-derived AMR along the urban water transmission pathway, we developed a genome-informed, system-scale comparative risk indicator (*Q*). This approach is based on the dataset of 61 identified sewage-derived ARGs and integrates three genetic risk determinants including ARG abundance, mobility potential (colocalization with MGEs), and association with pathogenic hosts into a multidimensional framework, which was adapted from the MetaCompare pipeline[32]. This enables direct comparison of resistome-associated hazards across different water compartments by quantifying the relative hazard potential based on contig-level genetic features, independent of absolute abundance or human exposure pathways (Fig. 7b; detailed calculation provided in Methods)."

5. *In addition, I find the risk 1 ARGs to be a bit problematic given their dubious relevance to clinical antibiotic use. While I understand the scheme formerly proposed by Tong Zhang's lab, the distinction is very important considering the need to effectively communicate to (and not frighten) the public:*
a. Three of the four risk 1 ARGs detected are related to membrane dynamics and/or peptide antibiotics, which are not extensively employed in humans to my knowledge. Instead, they are ARGs commonly detected in chiefly environmental taxa. My interpretation of this is that the analysis may be limited by the ability to assemble/bin/taxonomically classify some of the prevalent clinically relevant ARG scaffolds in wastewater. Often, these are encoded on plasmids or other MGEs which can be extremely difficult to bin into MAGs/may lack sufficient phylogenetic signal to taxonomically classify. Could the authors check for clinically relevant ARGs in unbinned contigs? E.g., CTX-M, NDM, PER, other ESBLs? This might give a sense of what is being missed in their workflow and not appearing in their highest risk ARGs.
b. fosA might be appropriately labeled a high risk ARG.

Response: Thanks for this thought-provoking comment. We concur with the reviewer that the clinical relevance of ARGs classified at high risk level warrants particularly rigorous interpretation, especially in public communication of antimicrobial resistance risks, where precision is essential to avoid unwarranted extrapolation of clinical implications.

First, ***fos*A** was retained as a high-risk ARG due to its documented clinical relevance, in accordance with the reviewer's suggestion.

Second, to enhance the identification of clinically relevant ARGs, we conducted **additional analyses of unbinned contigs** focusing on annotation of clinical-relevant ARGs and their associated bacterial taxa, the detailed procedure is provided in Supplementary Method 6. The analysis revealed four sewage-derived, clinically relevant ARGs—*sul*1, *lsa*B and OXA-10, and *tet*A(48)—newly detected in the unbinned contigs. Among these, ***sul*1** and **OXA-10** were identified carried by pathogens (*Vibrio cholerae*), and both were subsequently classified as Risk Rank I.

Consequently, three ARGs clinical relevance—*fos*A, OXA-10, and *sul*1—were ultimately assigned to Risk Rank I and integrated into the risk-ranking framework (modified **Fig. 7c**). These changes have been incorporated into the revised

manuscript (**Lines 406–408**).

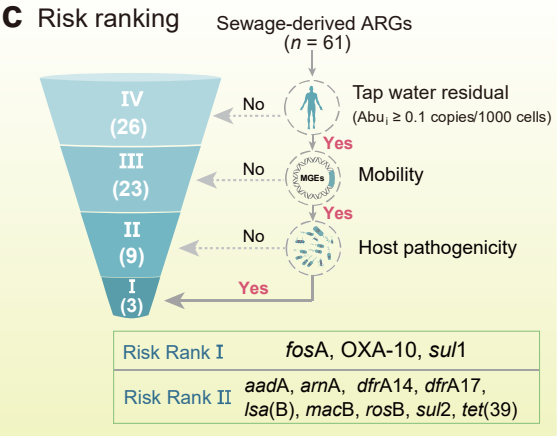


Fig. 7c | Risk stratification of sewage-derived ARGs through assessment of abundance, mobility potential, and host pathogenicity, yielding a tiered classification (I–IV).

Lines 406-408: "As result, three ARG subtypes were attributed to Risk Rank I, including *fos*A, OXA-10 and *sul*1, encoding resistance to fosfomycin, beta-lactam and sulfonamide, respectively."

Supplementary Method 6 | Detection of high-risk ARGs on unbinned metagenomic contigs. "To complement the MAG-based resistome risk assessment and to evaluate the potential underrepresentation of clinically relevant ARGs caused by assembly and taxonomic binning limitations, unbinned contigs (≥1,000 bp) were extracted from each metagenomic assembly. ORFs on these unbinned contigs were predicted using Prodigal (v2.6.3)[13]. The predicted ORFs were then searched against the SARG (v3.0) database using DIAMOND (v2.0.15), with an e-value threshold of $\leq 1 \times 10^{-10}$, a minimum amino acid identity of 80%, and a query coverage of ≥ 70%, to identify ARG-like ORFs[14]. To assess the potential association between ARGs and mobile genetic elements (MGEs), ORFs annotated as ARGs were further examined for spatial proximity to MGE-related genes on the same contig. All ORFs were screened against the MobileGeneticElement database[15] using DIAMOND with the same alignment criteria to identify MGE-related genes. An ARG was considered potentially mobilizable if an MGE-related gene was located within a 5 kb flanking region upstream or downstream on the same contig[16]. Subsequently, these unbinned contigs harboring both ARGs and nearby MGE-related genes were compared against the NCBI non-redundant (nr) database to provide additional functional annotation and taxonomic context, to identify those high-risk ARGs associated with genetic mobility and pathogenic bacterial hosts."

Minor comments:

6. *Please report in the text or supplementary the details around the construction of the integrase database. These genes are sometimes tricky to annotate given that, in some cases, they are only differentiable on the basis of a NT 15aa arm domain which is not in PFAM.*

Response: Thank you for highlighting this important point. Integron integrase candidates were identified and rigorously curated through a multi-stage

bioinformatic pipeline comprising: **(*i*) expansion** of the reference integrase database—initially established in our previously published work[1] (148 *int*I1 sequences)—by integrating newly available *int*I1 sequences from NCBI and INTEGRALL database[2] using iterative keyword-based retrieval and homology-based searching; **(*ii*) sequence dereplication** to remove exact duplicates; and **(*iii*) domain-based functional validation**, wherein all candidate sequences were screened against Pfam and CDsearch to confirm the presence and integrity of tyrosine recombinase core domain. This process yielded a non-redundant, functionally verified *int*I1 sequence dataset of 424 entries for downstream analysis and database construction. For clarification, we have now supplemented the detailed procedure in **Supplementary Method 4** and cited it in the main text (**Line 691**).

Supplementary Method 4 | Construction of the Class 1 Integrase Database

"The class 1 integrase (*int*I1) reference database was constructed based on publicly available sequences deposited in the National Center for Biotechnology Information (NCBI) database and the INTEGRALL database[10]. Briefly, annotated *int*I1-related sequences were retrieved from NCBI using the keywords "*int*I1" and "class 1 integrase", with searches restricted to the Gene and Protein categories, and complemented with curated class 1 integrase sequences obtained from INTEGRALL. After further dereplication and validation by comparing against Pfam and CDsearch database, 148 sequences were retained, composing the preliminary *int*I1 database[11]. To further expand sequence diversity, the preliminary *int*I1 database was used as a query to search against the NCBI non-redundant protein database (nr) using BLASTP, with an E-value cutoff of 1×10^{-5}. Candidate sequences were filtered based on sequence similarity (≥ 80%), and alignment coverage (≥ 50%). Duplicate or partial sequences were excluded. The retained candidates were subjected to an additional round of validation against the Pfam and CDsearch databases to exclude non-integron tyrosine recombinases. The resulting high-confidence sequences (n = 424) constituted the final class 1 integrase (*int*I1) reference database used for downstream analyses."

Reference:

[1] Ma LP, Li AD, Yin XL, Zhang T. The prevalence of integrons as the carrier of antibiotic resistance genes in natural and man-made environments. *Environ Sci Technol* 51, 5721-5728 (2017).

[2] Moura A, Soares M, Pereira C, Leitao N, Henriques I, Correia A. INTEGRALL: A database and search engine for integrons, integrases and gene cassettes. *Bioinformatics* 25, 1096-1098 (2009).

7. *What is meant by "diversity of arrangements" ? in Figure 5*

Response: Thanks for the kind reminder for pointing out the vague definition. The "diversity of arrangements" refers to "the number of unique, non-redundant ARG cassette arrays identified within integrons". For clarification, the definition was supplemented in legends of Figure 5 (**Lines 1036–1040**).

Lines 1036-1040: "Paired bar plots show the diversity of ARGs harbored by integrons (left), and the diversity of ARG cassette arrangements (right) quantified as the number of unique, non-redundant ARG cassette arrays identified within integrons. Concentric pie charts depict the relative distribution of ARG subtypes among the total integrons."

8. *There needs to be more reporting of assembly completeness- size, contiguity, etc. These aspects are important for being able to detect HGT.*
In addition, a correlation between assembly N50 and/or completeness with HGT event rates could demonstrate independence between assembly completeness and the frequency of detected HGTs.

Response: We thank the reviewer for highlighting the importance of assembly quality in HGT detection. In response, we have added comprehensive assembly metrics, including number of contigs, mean contig length, GC content, N50, N90, auN, L50, and L90 to **Supplementary Table 6**.

Supplementary Table 6. Summary statistics characterizing metagenomic assembly quality across samples.

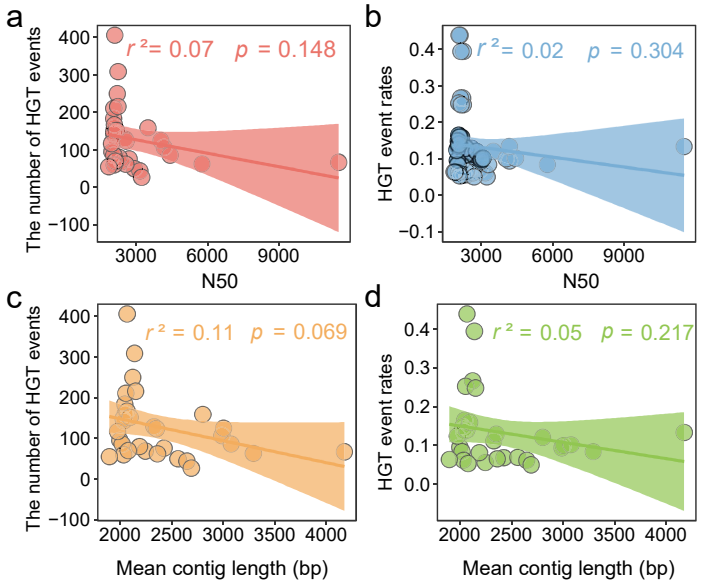
Sample	# contigs	Mean contig length (bp)	GC (%)	N50 (bp)	N90 (bp)	auN (bp)	L50 (bp)	L90 (bp)
INF1	160,141	2,068	44	2,108	1,121	4,724	38,692	128,826
INF2	153,601	2,122	50	2,210	1,127	5,019	36,105	122,842
INF3	206,954	2,245	56	2,393	1,135	8,715	43,192	163,254
INF4	144,850	2,140	46	2,246	1,138	4,856	34,592	115,736
AS1	192,201	2,425	57	2,706	1,156	12,113	35,167	148,729
AS2	225,183	2,328	58	2,541	1,160	7,053	47,196	176,415
AS3	165,778	2,188	56	2,282	1,131	6,961	36,126	131,611
AS4	180,355	2,352	60	2,584	1,159	8,783	36,700	140,871
EFF1	131,586	2,562	52	2,930	1,186	15,753	22,857	100,537
EFF2	127,835	2,648	55	3,160	1,191	16,409	20,663	96,723
EFF3	170,573	2,362	53	2,570	1,156	11,024	33,485	133,011
EFF4	102,415	2,690	52	3,233	1,206	15,988	16,623	77,254
RU1	154,110	1,998	48	1,988	1,106	5,305	37,922	124,794
RU2	202,533	2,048	50	2,097	1,122	4,639	50,206	163,300
RU3	176,188	2,020	49	2,064	1,118	4,213	44,601	142,480
RU4	173,509	2,045	51	2,059	1,113	6,254	41,625	139,823
RS1	178,776	2,035	47	2,068	1,119	4,595	44,377	144,315
RS2	178,347	2,027	48	2,060	1,119	4,663	44,811	144,116
RS3	159,202	2,034	48	2,078	1,116	4,393	39,443	128,494
RS4	130,474	2,055	50	2,066	1,112	5,638	30,654	104,994
RD1	144,006	1,975	47	1,967	1,108	4,614	36,600	116,941
RD2	148,592	2,150	52	2,234	1,131	5,770	33,893	118,491
RD3	228,808	2,079	48	2,129	1,123	4,898	55,066	183,828
RD4	191,638	2,067	50	2,093	1,116	7,531	45,231	154,083
SW1	167,243	2,096	51	2,173	1,125	5,151	40,025	134,140
SW2	112,026	1,896	47	1,854	1,095	4,127	29,641	91,689
TW1	119,904	3,292	61	5,766	1,226	55,873	8,379	84,056
TW2	74,050	4,174	58	11,510	1,338	75,284	3,561	47,097
TW3	236,885	2,802	60	3,501	1,200	21,714	32,141	176,087
TW4	198,966	2,987	59	4,174	1,216	27,750	23,107	144,853
TW5	217,467	3,002	60	4,041	1,228	46,072	26,687	158,312

TW6	148,254	3,077	60	4,430	1,214	54,513	13,453	106,710

Notes: N50 and N90 represent the contig lengths at which 50% and 90% of the total assembly size are reached, respectively. L50 and L90 indicate the number of contigs contributing to 50% and 90% of the total assembly length, respectively. auN (area under the Nx curve) represents an integrated contiguity metric that summarizes the entire contig length distribution, with higher values indicating more contiguous assemblies. These metrics were calculated based on contigs ≥1,000 bp, which were retained for downstream analyses.

To address whether assembly quality influenced HGT inference, we performed Spearman's rank correlation analysis between assembly metrics (N50 and contig length) and HGT event rates across all samples (**Supplementary Fig. 13**). As a result, no significant correlations were observed ($p > 0.05$), indicating that inter-sample variation in HGT frequency reflects genuine biological differences rather than technical artifacts attributable to assembly quality. These results have been incorporated into the revised Results section (**Lines 298–300**).

Lines 298-300: "Importantly, HGT event frequency showed no significant correlation with assembly quality metrics (Supplementary Fig. 13), indicating that the observed inter-sample variation in HGT frequency reflects genuine biological differences rather than assembly-related technical artefacts."



Supplementary Fig. 13 | Correlation between assembly contiguity and detected HGT events. Scatter plots show the correlations between assembly quality metrics and HGT detection outcomes across all samples. Panels depict the relationships between (**a**) N50 and the number of HGT events, (**b**) N50 and HGT event rates, (**c**) mean contig length and the number of HGT events, and (**d**) mean contig length and HGT event rates. HGT event rates were normalized by total assembly size. Each panel displays the coefficient of determination (r^2) and corresponding *p*-values from Spearman's rank correlation.

Reviewer #2:

1. *This study by Liu et al. investigates antimicrobial resistance in four urban water systems, including wastewater treatment plants, receiving river water, drinking water reservoirs, and tap water. The objective is to explore the potential connectivity of antimicrobial resistance across these systems. The authors collected 32 water samples and performed several analyses, including metagenomic sequencing, culturing and genome sequencing of lactose-fermenting Enterobacteriaceae, 16S rRNA gene amplicon sequencing, and chemical analysis of water quality.*
The study focuses on identifying antimicrobial resistance genes to infer resistance patterns within and across the sampled systems. The authors also attempt to evaluate potential risks using a risk score they developed. While the research question is relevant and timely, there are substantial concerns about the validity and interpretation of the results. The findings are mainly descriptive and lack mechanistic insight, offering limited advancement beyond existing studies that have already examined antimicrobial resistance in each of these environments. The investigation into connectivity, which could have been the most valuable contribution of the study, is not effectively addressed. The data from different systems are presented separately, and the suggested connections are speculative and insufficiently supported by analysis.

Response: Thank you for the detailed summary and for confirming that our research question is both relevant and timely. Your insightful comments have substantially strengthened the scientific rigor and clarity of the work. The manuscript has been carefully revised in response to all the comments and suggestions.

With respect to the two principal concerns raised by the reviewer, we have substantially strengthened the mechanistic depth and connectivity evidence of our study through the following key advancements:

(1) Multi-level evidence for hydrological connectivity. We performed complementary analyses across three hierarchical biological levels—resistome, subtype, and sequence—using machine-learning source-tracking (80% of downstream resistome from wastewater), shared ARG profiling (97 subtypes across all compartments), and phylogenetic homology (high sequence identity of high-risk ARGs). Collectively, these analyses provide robust, multi-dimensional evidence for inherent hydrological connectivity along the urban water continuum (Supplementary Fig. 7; Lines 236–238, 705–715).

(2) Study scope and sampling rationale. Extending beyond single-compartment studies, our study provides the first systematic investigation across the interconnected urban water system in a megacity, from sewage to receiving rivers, source water, and tap water, integrating metagenomics and culturomics to track AMR dynamics. The primary focus is the public health risk posed by tap water, specifically to characterize sewage-derived contamination and inform source-oriented interventions and precision management. We have also provided detailed rationale for our sampling design, acknowledging practical constraints while demonstrating that our multi-level connectivity evidence supports robust

conclusions (Lines 125–128, 428–432, 575–577).

(3) Systematic analysis of influencing factors and mechanistic drivers of ARG transmission and proliferation. We systematically examined potential influencing factors, including spatial heterogeneity, microbial diversity, biotic factors (bacterial host proliferation, HGT), and abiotic factors (water quality parameters, antibiotics, disinfectants, trace metals, and non-antibiotic pharmaceuticals). Mantel tests and PLS-SEM models were used to identify primary and direct/indirect drivers of ARGs. Additionally, in the revised manuscript, we have incorporated MetaCHIP analysis to identify major donor and recipient phyla in HGT networks (Supplementary Fig. 14; Lines 301–303, 678–680) and characterized tandem array features in integron-mediated ARG transfer (Lines 316–318, 508–510), offering deeper mechanistic insights into HGT as a primary driving factor of ARG dynamics.

(4) Enhanced mechanistic interpretation in the Discussion. In response to the reviewer's suggestion that the Discussion could benefit from stronger mechanistic interpretation, we have substantially revised this section. Redundant descriptions were removed, paragraphs were merged, and in-depth mechanistic interpretations regarding ecological connectivity and horizontal gene transfer were added (Lines 428–432, 466–470, 487–494, 508–514).

Together, these revisions strengthen both the analytical depth and mechanistic rigor of this study. It elucidates the mechanisms of AMR transmission enabled by ecological connectivity, identifies sewage-derived ARGs that persist throughout the system, and offers a scientific foundation for source-oriented interventions to mitigate associated health risks in drinking water.

The detailed analytical results and point-by-point responses are provided as follows.

This reviewer identifies three major flaws in the study:

2. *The Methods section states that wastewater effluent is discharged into the receiving river, suggesting a direct connection between these two systems. The authors also analyzed water from two reservoirs and tap water; however, no samples were collected from drinking water treatment plants. It is assumed that reservoir water undergoes treatment before distribution. One important gap in the study's logic is the lack of a clearly established hydrological connection between the receiving river and the reservoirs. Without this connection, inferences about AMR transmission across these systems are not scientifically justified. Furthermore, the absence of data from the drinking water treatment process limits the ability to assess how treatment may influence AMR levels. Disinfection steps such as chlorination or UV treatment can significantly reduce microbial load and potentially remove ARGs. Additionally, tap water is not an ideal proxy for treated water, as household plumbing can alter the microbial profile through biofilm formation and water stagnation. These limitations in the sampling design significantly weaken the central claim of AMR connectivity across the four urban water systems. As currently presented, the study does not provide sufficient evidence to support this conclusion.*

Response: We thank the reviewer for the careful assessment of the hydrological

connection and sampling design. We fully agree that these two points are central to the exploration and significance of this study. In response, we have supplemented extensive analyses to further support hydrological connection based on machine-learning-based source tracking and genetic phylogenetic analysis.

(1) Regarding the concern on drinking water treatment plants not included

First, we agree with the reviewer that tap water cannot be equivalent to treated water, as household plumbing can alter the microbial profile through biofilm formation and water stagnation. In fact, the primary focus of this study is the public health risk posed by tap water, specifically to characterize the categories of sewage-derived contamination present therein to inform source-oriented intervention and precision management, rather than the drinking water treatment process itself. This focus is motivated by our prior source-tracking analysis across 47 cities[1], which applied a Bayesian-based machine learning algorithm and revealed that ARGs in global tap water largely originate from urban wastewater (Introduction, Lines 58–62). This finding expands our understanding that domestic sewage is a major contributor to elevated AMR risk in drinking water, indicating that human activities exacerbate the ongoing input of such pollution. However, based on current knowledge, the transmission pathways from sewage to natural waters and ultimately to tap water remain unclear. Moreover, categories of sewage-derived contaminants relevant to human health are still largely unregulated in drinking water, posing significant public health risks in regions with inadequate water treatment and where tap water is consumed directly. Therefore, this study aims to elucidate the underlying transmission mechanisms of AMR determinants and to identify sewage-derived risk indicators, thereby supporting evidence-based, source-oriented intervention and precision management of drinking water.

Second, we agree with the reviewer that disinfection steps implemented in drinking water treatment plants can substantially reduce microbial load and influence the ARG compositions. We have, in fact, conducted extensive research on the effects of drinking water disinfection on ARGs and antibiotic-resistant bacteria, both in full-scale treatment processes[2] and lab-based simulations[3]. These studies collectively indicate that the removal efficiency of pathogen-mediated ARGs is limited, and some related publications even report increased relative abundance of ARGs in bacterial communities following disinfection[4, 5]. Thus, drinking water treatment plants were not covered in this study, as our focus lies on the major compartments of the urban water system instead of on specific treatment processes.

We acknowledge the importance of disinfection dynamics and have incorporated a discussion of this aspect in the revised manuscript, suggesting that future studies integrate disinfection intensity and residual chlorine effects with upstream source control for a more comprehensive risk management strategy (**Lines 466–470**).

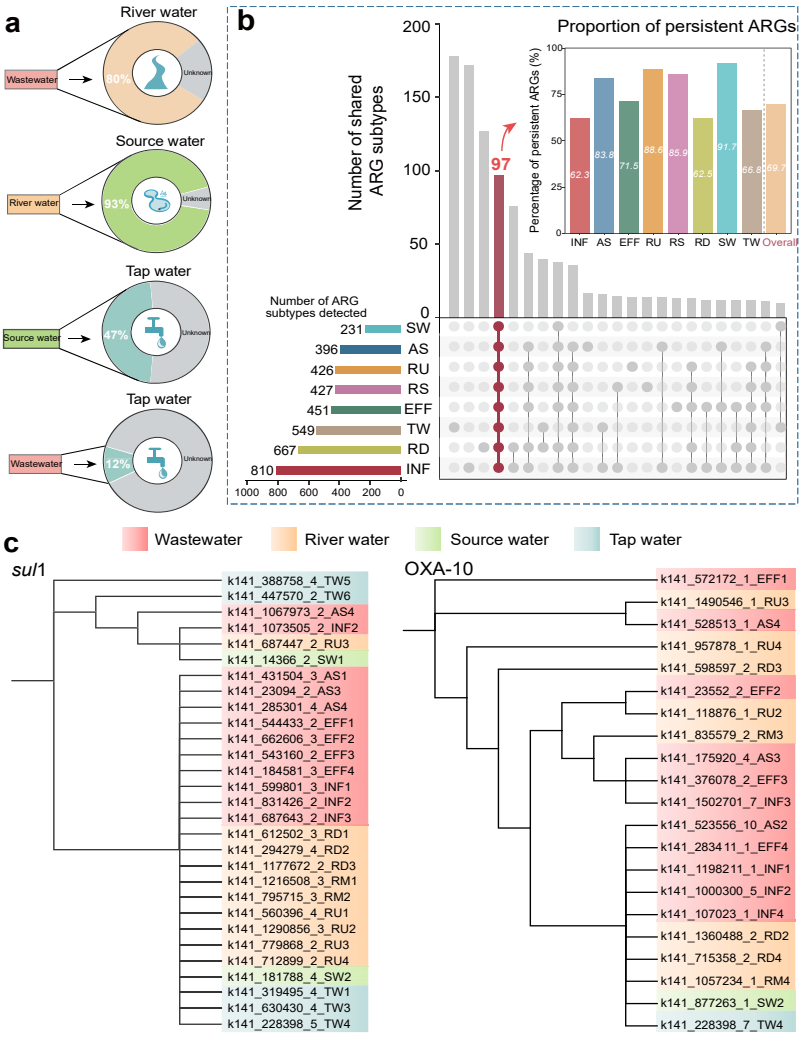
> **Lines 466-470**: "Future integration of disinfection dynamics, including disinfection intensity and residual chlorine in distribution systems, with upstream source control may further strengthen a coordinated dual-end strategy, offering a more comprehensive paradigm for safeguarding public health from AMR-related hazards across the urban water system."

(2) Regarding the concern on hydrological connection

We concur with the reviewer that endogenous hydrological connectivity serves as a key conceptual foundation of this study. As articulated above, the primary objective is to elucidate the transmission mechanisms of AMR determinants behind concerns over drinking water safety driven by domestic sewage. The pathway "sewage discharge → receiving water → source water → tap water" provides an ideal model for tracing dynamics and assessing associated risks, as stated in the ***Discussion*** (Lines 536–540): "The focused water pollution pathway in a megacity can be regarded as a conceptual model for AMR transmission, representatively reflecting its dynamic features; the sewage receiving, treatment and discharge of the large populations of a megacity could represent the pollution emission characteristics of the surrounding regions, a collection of populations under relatively low density in the upstream."

To further substantiate hydrological connection despite the inherent complexity of urban water systems, we performed complementary analyses across three hierarchical biological levels—**resistome, subtype, and sequence**[6, 7]. The results are as follows: **(1) Resistome level**: SourceTracker analysis[8] based on ARG compositions revealed extensive resistome sharing across the main water compartments (80%, 93%, and 47%, respectively), supporting hydrological connectivity (Supplementary Fig. 7a). **(2) Subtype level**: A total of 97 ARG subtypes (representing 69.7% of total abundance on average) were shared across all compartments (Supplementary Fig. 7b), highlighting their potential for dissemination via hydrological connections. **(3) Sequence level**: Phylogenetic analysis based on nucleotide sequence homology for Risk Rank I ARGs revealed high sequence identity across compartments (Supplementary Fig. 7c). Collectively, these multi-level analyses provide robust evidence for inherent hydrological connectivity among the sampled water compartments.

We sincerely thank the reviewer for the insightful comments, which enabled us to strengthen our analysis and interpretation of hydrological connections. The corresponding methodological details and results have been comprehensively updated in the revised manuscript (**Lines 705–715**, **236–238**).



Supplementary Fig. 7 | Ecological connectivity analyses of ARGs across three hierarchical levels (resistome, subtype, sequence) in the urban water system. a, Bayesian source-tracking analysis showing the proportional contributions of sewage, receiving river, and source water to downstream ARG profiles, respectively. ARGs not assigned to any predefined source are shown as "Unknown". **b**, UpSet plot depicting the distribution of persistent ARG subtypes across water compartments. The adjacent bar chart indicates their relative proportion. **c**, Maximum-likelihood phylogenetic tree of *sul*1 and OXA-10 based on their complete ORF sequences recovered from metagenomic assemblies across urban water compartments, illustrating their sequence-level connectivity**.**

Lines 705-715: "**Ecological connectivity analysis of antibiotic resistance**

To characterize AMR connectivity across the UWS, we performed complementary analyses across three hierarchical biological levels "resistome-subtype-sequence" [78,79], respectively. First, Bayesian-based SourceTracker (v0.9.1)[80] was applied to characterize the probabilistic contribution of resistance gene sources across urban water compartments. Second, a custom script was developed to identify the shared ARG subtypes consistently detected across all urban water compartments; the summed relative abundance of these subtypes across samples was calculated as a prevalence index. Third, sequences annotated as the same high-risk ARGs were extracted from metagenomic assemblies, followed by dereplication of redundant sequences to retain unique gene variants. The remaining sequences were aligned using

MUSCLE (v5.3)[81], and maximum-likelihood phylogenetic trees were constructed using FastTree (v2.1.11)[82]."

Lines 236-238: "Moreover, systematic homology-based analysis across three hierarchical biological levels "resistome-subtype-sequence" provides evidence for their hydrological connectivity within the UWS (Supplementary Fig. 7)."

Reference:

[1] Wang C, Yang HY, Liu HF, Zhang XX, Ma LP. Anthropogenic contributions to antibiotic resistance gene pollution in household drinking water revealed by machine-learning-based source-tracking. *Water Res* 246, 120682 (2023).

[2] Chao YQ, Ma LP, Yang Y, Ju F, Zhang XX, Wu WM, Zhang T. Metagenomic analysis reveals significant changes of microbial compositions and protective functions during drinking water treatment. *Sci Rep* 3, 3550 (2013).

[3] Ma LP, Yang H, Guan L, Liu X, Zhang T. Risks of antibiotic resistance genes and antimicrobial resistance under chlorination disinfection with public health concerns. *Environ Int* 158, 106978 (2022).

[4] Zhao W, Hou Y, Wei L, Wei W, Zhang K, Duan H, Ni BJ. Chlorination-induced spread of antibiotic resistance genes in drinking water systems. *Water Res* 274, 123092 (2025).

[5] Meng Q, Wang J, Li K, Zhang Y, Hu Z, Wang F, Pan F, Fu J, Dang C. Low-dose chlorine disinfection poses a greater potential risk of antibiotic resistance genes and their pathogenic hosts. *Water Res* 289, 124895 (2026).

[6] Zhao YX, Li LG, Huang Y, Xu XQ, Liu ZS, Li SX, Zhu LZ, Hu BL, Zhang T. Global soil antibiotic resistance genes are associated with increasing risk and connectivity to human resistome. *Nat Commun* 16, 7141 (2025).

[7] Xu XQ, Lin Y, Deng Y, Liu L, Wang D, Tang Q, et al. Ecological connectivity of genomic markers of antimicrobial resistance in *Escherichia coli* in Hong Kong. *Nat Commun* 16, 7319 (2025).

[8] Knights D, Kuczynski J, Charlson ES, Zaneveld J, Mozer MC, Collman RG, Bushman FD, Knight R, Kelley ST. Bayesian community-wide culture-independent microbial source tracking. *Nat Methods* 8, 761-763 (2011).

3. *A second major concern is that, unlike wastewater treatment plants, the receiving river and reservoirs are open systems that likely receive inputs from a variety of sources, including urban runoff and other non-point pollution. Consequently, it is not possible to attribute the presence of AMR exclusively to discharged wastewater. Moreover, key factors such as water residence time, environmental conditions, and the growth or decay of bacteria carrying ARGs are likely to influence AMR dynamics in these systems. These factors are not considered in the study. By oversimplifying the complexity of the aquatic environment, the study draws conclusions that are premature and not adequately supported by the data.*

(1) Regarding the open nature of receiving rivers and reservoirs

We fully concur with the reviewer that the sources are highly diverse, and the ARGs detected in the receiving river and reservoirs should not be solely attributed to discharged wastewater. As assessed in our response to the above comments, Bayesian machine learning-based source tracking analysis of ARG compositions indicated that approximately 80% of the resistome in the downstream river originated from wastewater, and 67% in reservoirs. In addition, three primary

motivations underpin this work: **First**, based on large-scale survey, drinking water samples with human health risks harbor diverse ARGs, with a 100% detection frequency across samples; notably, these ARGs were frequently identified in pathogens[1, 2, 3]. This underscores the necessity of tracing and managing upstream pathways that shape resistome risks at the point of consumption. **Second**, source-tracking analysis across 47 cities worldwide revealed that urban wastewater is the key anthropogenic source of ARGs in global tap water, contributing an average of 36% of the total load[1]. This recurring pattern across regions provides a strong rationale for prioritizing wastewater as a key control node within integrated urban water networks. **Third**, both field investigations and laboratory-scale chlorination experiments have demonstrated that conventional disinfection processes can enhance the relative abundance of certain high-risk ARGs and antibiotic-resistant bacteria[4, 5, 6]. Under such conditions, source-directed intervention represents the most direct and effective strategy for mitigating the risk posed by ARGs in drinking water. However, the transmission pathways of ARGs in the drinking water system remain poorly characterized. Thus, this study was designed to address this critical knowledge gap by generating mechanistic insights and quantitative evidence to inform targeted, source-directed intervention strategies.

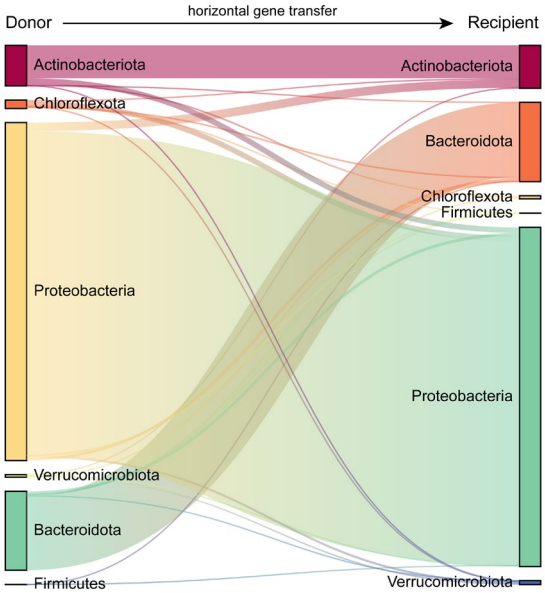
(2) Regarding environmental and biological processes

We thank the reviewer for raising this important point. We would like to clarify that these key processes were explicitly considered and systematically analyzed in our study. Specifically, we considered spatial heterogeneity, microbial diversities, both biotic factors (i.e., bacterial host proliferation, horizontal gene transfer) and abiotic factors (water quality parameters, antibiotics, disinfectants, trace metals, and non-antibiotic pharmaceuticals). These were analyzed using Mantel tests and PLS-SEM models to identify primary as well as direct/indirect drivers of ARGs. Thus, our manuscript addresses the environmental factors and host bacterial proliferation characteristics noted by the reviewer.

Regarding the additional factor of water residence time (WRT) the reviewer mentioned, it was not explicitly included in the current analysis. This was primarily due to: (1) methodological constraints from field sampling—to ensure sufficient biomass for metagenomic sequencing, sample volumes varied substantially across different water environments, which precluded the standardized flow/volume calculations required for WRT estimation; and (2) the general lack of established WRT parameter data in the specific natural water bodies investigated in this study. We acknowledge this as a valuable suggestion and fully agree that WRT is a critical parameter for understanding pollutant fate. In our subsequent research focused on source control strategies, we will certainly seek to incorporate hydraulic retention time (HRT) in engineered systems and, where feasible, collaborate with hydrologists to estimate WRT in natural water bodies. We sincerely appreciate the reviewer's insightful comment, which has helped us identify an important direction for future work.

To further strengthen mechanistic insights into HGT identified as a primary driving

factor of ARG dynamics, we additionally incorporated MetaCHIP analysis to identify major donor and recipient phyla in HGT networks (**Supplementary Fig. 14**; **Lines 301–303**, **678–680**), and characterized tandem arrays in integron-mediated ARG transfer and their dynamics in urban water system (**Lines 316–318**, **508–510**).



Supplementary Fig. 14 | Phylum-level summary of inferred donor-recipient relationships in ARG-associated horizontal gene transfer. Directional Sankey diagram summarizing horizontal gene transfer events inferred by MetaCHIP at the phylum level. Each flow represents inferred HGT events from donor phyla (left) to recipient phyla (right), with link widths proportional to the number of detected events.

Lines 301-303: "These putative HGT events predominantly involved six bacterial phyla, mainly *Proteobacteria*, *Bacteroidota*, and *Actinobacteriota*, with transfers occurring both within and across phyla (Supplementary Fig. 14)."

Lines 678-680: "MetaCHIP (v1.10.6)[72] was used to identify the primary bacterial taxa serving as donors and recipients in ARG-associated HGT events."

Lines 316-318: "Notably, among all integron cassette arrays, 6.5% contained at least one tandemly duplicated ARG, predominantly *aad*A2, *aad*A24, and *aad*A5."

Lines 508-510: "The observation of both conserved and compartment-specific integron-associated ARG arrays across the UWS indicates concurrent downstream dissemination of stable MGE–ARG combinations and local cassette rearrangements mediated by HGT."

Reference:

[1] Wang C, Yang HY, Liu HF, Zhang XX, Ma LP. Anthropogenic contributions to antibiotic resistance gene pollution in household drinking water revealed by machine-learning-based source-tracking. *Water Res* 246, 120682 (2023).

[2] Liu HF, Jiao PB, Guan L, Wang C, Zhang XX, Ma LP. Functional traits and health implications of the global household drinking-water microbiome retrieved using an integrative genome-centric approach. *Water Res* 250, 121094 (2024).

[3] Ma LP, Li B, Zhang T. New insights into antibiotic resistome in drinking water and management perspectives: A metagenomic based study of small-sized microbes. *Water Res* 152, 191-201 (2019).

[4] Ma LP, Yang HY, Guan L, Liu X, Zhang T. Risks of antibiotic resistance genes and antimicrobial resistance under chlorination disinfection with public health concerns. *Environ Int* 158, 106978 (2022).
[5] Jia SY, Wang S, Zhuang Y, Gao L, Zhang X, Ye L, Zhang XX, Shi P. Free-living lifestyle preferences drive the antibiotic resistance promotion during drinking water chlorination. *Water Res* 249, 120922 (2024).
[6] Zhang HC, Chang FY, Shi P, Ye L, Zhou Q, Pan Y, Li AM. Antibiotic resistome alteration by different disinfection strategies in a full-scale drinking water treatment plant deciphered by metagenomic assembly. *Environ Sci Technol* 53, 2141-2150 (2019).

4. *The authors state that samples were collected in June 2023; however, it is unclear whether the sampling strategy was designed to follow the same parcel of water through the different components of the urban water system. Please provide a clearer description of the timeline and coordination of the 32 sample collections. Without this information, it is difficult to assess whether the data represent a continuous flow through the system or isolated, independent snapshots. This lack of temporal alignment raises concerns about the validity of the conclusions regarding connectivity among the water systems.*

Response: We thank the reviewer for this thought-provoking comment. "Tracing the same parcel of water through the different components" sounds interesting, and we fully agree with the reviewer that this idea is indeed an ideal approach for directly demonstrating hydrological connectivity. We acknowledge that our sampling strategy was not designed as a Lagrangian tracking of a single water mass, and we appreciate the opportunity to clarify the rationale and constraints of our field sampling design:

(1) Practical challenges in implementing a "same parcel" tracking approach

a. Disparity in microbial biomass across compartments: Due to the drastically different microbial abundances across the investigated water compartments, the sample volumes required to obtain sufficient DNA for metagenomic sequencing varied substantially, ranging from 1 liter for wastewater to up to 2,000 liters for tap water. **b. Inaccessible flow paths**: A substantial portion of the urban water system, particularly the segments connecting the source water reservoir to the drinking water treatment plant, and the distribution network to taps, consists of enclosed pipelines or restricted facilities. Therefore, the variability in sample volumes and the inaccessibility of enclosed pathways render the attempts to follow a single, tagged water parcel impracticable.

(2) Our sampling strategy: a representative "snapshot" under stable conditions

In light of above constraints, we designed our sampling to capture a representative snapshot of the system under stable, baseline conditions. All 32 samples were collected within a concentrated two-week window in June 2023, during normal system operation, without major rainfall events, process upsets, or changes in treatment configuration. Samples were collected sequentially from wastewater and its receiving water bodies, followed by source water and then household tap water.

This approach minimizes the influence of short-term fluctuations in climate, flow, and human activities, allowing us to focus on the intrinsic hydrological linkages reflected in the shared resistome.

(3) Connectivity conclusions are supported by multi-level evidence

Importantly, our conclusion regarding hydrological connectivity does not rely solely on the timing of collection. It is substantiated by multiple independent lines of evidence: **(a)** SourceTracker analysis quantitatively attributed resistome signatures from upstream to downstream compartments; **(b)** Shared ARG subtypes (n = 97) were consistently detected across all compartments; **(c)** High sequence homology of ARGs across compartments strongly suggests recent common origins and transmission via water flow. Detailed analyses have been provided in our response to the reviewer's comment #2. These complementary analyses performed at three hierarchical biological **"resistome-subtype-sequence" levels** collectively substantiate the role of hydrological connectivity in ARG dissemination.

We have now clarified these methodological considerations in the revised manuscript (**Lines 575–577**). We fully agree with the reviewer that future work should aim to include high-frequency sampling or tracer studies to quantitatively link water residence time to ARG fate. We sincerely appreciate this insightful comment, which has helped us articulate the rationale, constraints, and strengths of our field sampling design more clearly.

> **Lines 575-577**: "All samples were collected sequentially along the hydrological pathway during a two-week period in June 2023 under stable operating conditions, without major rainfall, process upsets, or treatment changes."

Other comments:

5. *The title implies that the study addresses mechanisms and mitigation strategies, but neither is clearly presented. It is unclear what mechanisms are being referred to, and no specific mitigation approach is described. Furthermore, the claim of amplified health risks appears overstated, as the risk assessment is based solely on bioinformatic predictions without experimental validation or supporting measured data.*

Response: Thank you for this thoughtful comment. "Mechanisms" refers to the transmission dynamics and driving factors of ARGs along urban water continuum revealed in this study. "Mitigation strategies" refers to the established risk assessment framework and identified high-risk indicators that can inform future source control interventions. Regarding "amplified health risks", our intent was to highlight the contribution of wastewater to tap water ARGs. To avoid overstatement, we have removed these terms and revised the title to: "Ecological Connectivity of Anthropogenic Antimicrobial Resistance in Urban Water Systems and Its Implications for Drinking Water Risk Assessment" (**Lines 1–3**).

We appreciate the reviewer's guidance in improving our manuscript's precision.

6. *The Introduction is relatively general and would benefit from greater specificity. In particular, it would be helpful for the authors to clarify whether there have been prior efforts to investigate the fate and transport of AMR across different water compartments, whether in natural or urban systems.*

Response: Thanks for this thoughtful suggestion. Prior efforts have been paid on AMR transmission across specific compartments of UWS, such as from wastewater effluents to receiving rivers[1, 2], or from source water to finished drinking water[3, 4]. These provide preliminary insights into both the dissemination of anthropogenic pollution to natural environments, and persistent AMR risks within drinking water distribution systems. However, systematic research encompassing the key compartments of UWS remains scarce, especially for identifying key sewage-derived indicators to enable targeted source control and ultimately ensure drinking water safety. As suggested, the contents have been supplemented in the Introduction (**Lines 70–74**).

> **Lines 70-74**: "Previous research has advanced our understanding of AMR transmission across specific compartments of UWS, such as from wastewater effluents to receiving rivers[15,16], or from source water to finished drinking water[11,17]. Nevertheless, systematic research encompassing the key compartments of UWS remains scarce."

Reference:

[1] Knight ME, Farkas K, Kiss A, Jones DL. National-scale insights into AMR transmission along the wastewater-environment continuum. *Water Res* 282, 123603 (2025).

[2] Wang N, Li S, Shi M, Ni N, Zhang X, Guo X, Lin H, Luo Y. Trajectory of antibiotic resistome response to antibiotics gradients: A comparative study from pharmaceutical and associated wastewater treatment plants to receiving river. *Water Res* 266, 122444 (2024).

[3] Ke YC, Sun WJ, Xue YE, Zhu Y, Yan S, Xie SG. Effects of treatments and distribution on microbiome and antibiotic resistome from source to tap water in three Chinese geographical regions based on metagenome assembly. *Water Res* 249, 120894 (2024).

[4] Han ZM, Zhang Y, An W, Lu J, Hu J, Yang M. Antibiotic resistomes in drinking water sources across a large geographical scale: Multiple drivers and co-occurrence with opportunistic bacterial pathogens. *Water Res* 183, 116088 (2020).

7. *Line 107. The authors describe the study as a case study, and the manuscript is written as such. However, this framing limits the broader scientific value of the work. The manuscript lacks a discussion of how the findings contribute to generalizable knowledge or advance the broader understanding of AMR dynamics in urban water systems.*

Response: Thank you for this thoughtful suggestion, which has helped enhance the scientific value of our study. As suggested, we have added a discussion at the end of this paragraph in the ***Introduction*** "The novel knowledge on transmission dynamics, critical control nodes, and risk assessment framework incorporating key indicators, enabled by ecological connectivity, provides a scientific foundation for targeted interventions to disrupt critical transmission links and mitigate health risks in urban water systems." (**Lines 125–128**).

Additionally, associated discussion had been included at the end of the ***Discussion*** "Our systematic assessment of AMR risks across the UWS provides novel knowledge on transmission dynamics, critical control nodes, and risk assessment framework incorporating key indicators. These findings establish a foundation for future research prioritizing high-resolution tracking of wastewater-derived ARG pollutants with drinking water contamination potential, and the development of precision intervention strategies, particularly for mobile ARGs associated with pathogenic hosts. These efforts will substantially advance mitigation of anthropogenic health risks stemming from waterborne AMR dissemination." (Lines 540–546), and the ***Conclusion*** "Collectively, our research advances the systematic studies on waterborne AMR dissemination, while providing novel insights into critical control points, risk assessment paradigms, and source-targeted strategies for system-wide AMR risk management. This establishes a scientific foundation for mitigating anthropogenic threats to ecological and public health." (Lines 565–569).

Lines 125-128: "The novel knowledge on transmission dynamics, critical control nodes, and risk assessment framework incorporating key indicators, enabled by ecological connectivity, provides a scientific foundation for targeted interventions to disrupt critical transmission links and mitigate health risks in urban water systems."

8. *Line 132. It is unclear what the underlying assumption is when representing the number of ARG copies per cell. Since microbial communities are heterogeneous, with some cells carrying multiple ARGs and others carrying none, it is important to explain what this metric represents and how it should be interpreted.*

Response: We appreciate the considerate comment and agree that microbial communities are highly heterogeneous, with some cells carrying multiple ARGs and others carrying none. The metric "ARG copies per cell" is based on the assumption of "one genome per cell"[1, 2]. It reflects the average load or relative density at the community level by normalizing for sequencing depth, gene length, and prokaryotic cells/genomes (equation listed as follows, Supplementary Method 3). As a collective measure of resistance potential within a population, it enables standardized comparisons across different environments[3]. As suggested, we have clarified the definition of this metric in the Methods (**Lines 654–656**).

$$ARG\ Abundance\ (copy/cell) = \sum_{1}^{n} \frac{N_{ARG-like\ reads} \times L_{read}/L_{ARG\ reference\ sequence}}{Cell\ number}$$

where $N_{ARG\text{-}like\ reads}$ is the number of ARG-like reads annotated to one specific ARG reference sequence, L_{read} indicates the clean read length, $L_{ARG\ reference\ sequence}$ is the nucleotide sequence length of the corresponding ARG reference sequence, n is the number of mapped ARG reference sequences attributed to the ARG type or subtype, and *Cell number* is the cell number estimated by mapping against an essential single-copy marker gene database[7].

Lines 654-656: "ARG abundance was expressed as copies per cell, representing the relative abundance normalized to the estimated total microbial cell count[49], as detailed in Supplementary Method 3."

Reference:

[1] Akerlund T, Nordstrom K, Bernander R. Analysis of cell size and DNA content in exponentially growing and stationary-phase batch cultures of *Escherichia coli*. *J Bacteriol* 177, 6791-6797 (1995).

[2] Yin XL, Zheng XW, Li LG, Zhang AN, Jiang XT and Zhang T. ARGs-OAP v3.0: Antibiotic-resistance gene database curation and analysis pipeline optimization. *Engineering* 27, 234-241 (2023).

[3] Yin XL, Chen X, Jiang XT, Yang Y, Li B, Shum MH, et al. Toward a universal unit for quantification of antibiotic resistance genes in environmental samples. *Environ Sci Technol* 57, 9713-9721 (2023).

9. *Line 149. Please clarify what constitutes a "strain" of lactose-fermenting Enterobacteriaceae in this study. Is this classification based on genetic criteria such as average nucleotide identity (ANI), or another method? This point is not clearly described in the Methods. Additionally, among the 3,040 isolated strains, how many are genetically unique?*

Response: We sincerely thank the reviewer for raising this important point. All 3,040 lactose-fermenting *Enterobacteriaceae* (LFE) strains—95 per sample, individually isolated for phenotypic characterization of antibiotic resistance and pooled in equimolar proportions for ARG profiling—were obtained *via* pure-culture isolation using MacConkey agar as a selective and differential medium. The procedure was detailed in Lines 624–635 of the ***Methods***. MacConkey agar selectively isolates gram-negative enteric bacteria through the inhibitory action of crystal violet and bile salts against gram-positive organisms. It differentiates LFE strains, which produce pink colonies *via* acidification of the medium. Strain identity was confirmed at the community level using MetaPhlAn4[1] applied to metagenomic sequencing data generated from the pooled LFE isolates. Notably, our objective was not to enumerate genetically distinct strains but **to characterize and compare the collective resistome features of culturable LFE across the urban water system**.

Nevertheless, the genetic diversity and strain-level heterogeneity of LFE isolates from aquatic environments remain underexplored and warrant dedicated investigation. In future work, we plan to perform whole-genome sequencing of representative isolates to resolve strain-level phylogeny, track ARG carriage dynamics, and elucidate environmental persistence and transmission mechanisms at the single-strain level.

To avoid ambiguity, **the term "strain" has been consistently replaced with "isolate" throughout the manuscript** to accurately reflect that these units correspond to independent culture-derived isolates obtained *via* colony purification on MacConkey agar, not genetically defined strains established through whole-genome sequencing or other high-resolution phylogenomic criteria.

Reference:

[1] Blanco-Miguez A, Beghini F, Cumbo F, McIver LJ, Thompson KN, Zolfo M, et al. Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn

4. *Nat Biotechnol* 41, 1633-1644 (2023).

10. *Line 240. The use of 16S rRNA gene amplicon data to associate microbial taxa with ARGs provides only correlative evidence and does not allow for reliable identification of true ARG hosts. This method does not confirm whether the taxa are actually carrying the ARGs, and therefore referring to them as "hosts" is both inaccurate and potentially misleading.*

Response: We thank the reviewer for this important methodological clarification and fully agree that co-occurrence analysis based on 16S rRNA gene amplicon data provides only correlative evidence and cannot reliably identify true ARG hosts. In the present study, we therefore employed a two-pronged approach. **First**, 16S-based network analysis was used to identify bacterial taxa that were statistically associated with ARGs across samples, revealing potential associations at the community level. **Second**, we applied metagenomic binning to more definitively assign ARGs to their bacterial hosts based on metagenome-assembled genomes (MAGs).

To avoid any misinterpretation, we have revised the description of the co-occurrence analysis. The term "potential bacterial hosts" has been removed, and framed as "taxa-ARG association patterns" (**Lines 251**–**253**).

> **Lines 251-253**: "First, we characterized the microbial community dynamics in UWS by employing high-throughput sequencing of 16S rRNA genes, and then identified taxa-ARG association patterns through co-occurrence network analysis."

11. *Line 262. While the use of MAGs to link ARGs to potential hosts is a more robust approach, it is unclear whether dereplication of MAGs was performed, as this step is not explicitly mentioned in the Methods. Dereplication is important to eliminate redundancy and ensure that the analysis is based on representative, non-redundant genomes. Without this step, the inclusion of closely related or duplicate MAGs could bias the results and lead to inflated or misleading interpretations.*

Response: We thank the reviewer for highlighting the importance of dereplication. To clarify, MAGs were dereplicated after refinement using dRep (v3.3.0)[1] with parameters -sa 0.95 -nc 0.30. This step removed redundant or highly similar genomes, ensuring that subsequent identification of ARG-host associations was based on representative, non-redundant MAGs. We have now explicitly included this detail in the Methods section to enhance clarity and reproducibility (**Lines 666**–**668**).

> **Lines 666-668**: "The recovered MAGs were subsequently dereplicated using dRep (v3.3.0)[63] with the parameters '-sa 0.95 -nc 0.30'."

Reference:

[1] Olm MR, Brown CT, Brooks B, Banfield JF. dRep: A tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J* 11, 2864-2868 (2017).

12. *Line 319. The manuscript would benefit from stronger evidence linking ARGs to the measured antibiotics. For instance, is there a correlation between tetracycline concentrations and the abundance of the tetA gene? According to the data in Line 331, tetA does not appear to correlate with tetracycline levels.*

Response: Thanks for this insightful observation. Indeed, the relationship between specific ARGs and measured antibiotics reveals distinct patterns. As noted, *tet*A showed no significant correlation with tetracycline levels. In contrast, *sul*2 (r = 0.632, $p < 0.001$) and *aad*A (r = 0.635, $p < 0.001$) exhibited statistically significant positive correlations with sulfamethoxazole concentrations (Figure 6b). The results have been reported in Lines 344–346 of the ***Results*** section.

13. *Line 350. The Methods and Results sections do not clearly articulate the novelty of the health risk assessment used in this study. It is unclear how this approach differs from existing methods, or how its novelty is reflected in the equation presented in Line 381.*

Response: We thank the reviewer for this thoughtful comment. In the Results section, we present a systematic framework for "AMR-associated health risk assessment and monitoring indicators in urban water systems". This framework integrates three key components: (1) identification of sewage-derived ARGs, (2) quantitative assessment of associated health risks, and (3) risk-based prioritization, corresponding respectively to contaminant fingerprinting, comprehensive risk analysis, and targeted monitoring of high-priority pollutants. First, ***sewage-derived ARGs*** were identified based on their observed dynamics across the urban water continuum. Specifically, ARGs were classified as sewage-derived if they met the following criteria: (*i*) maintaining a relatively high abundance in WWTP effluent (≥ 0.1 copies/1000 cells); (*ii*) exhibiting proliferation of ≥ 50% or emerging up to ≥ 0.1 copies/1000 cells in downstream receiving water; and (*iii*) being detectable in household drinking water (Fig. 7a). This process yielded a set of 61 sewage-derived ARGs. Second, ***health risk assessment*** was performed on these 61 sewage-derived ARGs using a quantitative structural risk score (Q) adapted from the MetaCompare framework[1]. This score integrates genetic risk metrics considering the co-occurrence of ARGs, MGEs, and pathogens (Fig. 7b). Third, ***risk-based prioritization*** was conducted using an established ranking framework that integrates three sequence-based criteria—abundance, mobility, and pathogenicity—to identify high-risk ARGs for targeted monitoring (Fig. 7c).

To avoid any misinterpretation and enhance clarity, we have relocated the health risk assessment equation to the Methods section (**Lines 716–731**).

Lines 716-731: "**Quantitative model for sewage-derived ARG-associated health hazard assessment.** To quantify health hazards associated with sewage-derived resistomes across different urban water compartments, we developed a quantitative hazard score (Q) model adapted from the MetaCompare framework[32]. The 61 identified anthropogenic ARGs were included to construct a sewage-derived ARGs dataset, based on which to integrate their genetic risk determinants at the contig level: (*i*) ARG relative abundance, (*ii*) ARG mobility

potential indicated by physical colocalization with MGEs within a 5 kb distance on the same contig[83], and (*iii*) ARG association with pathogenic hosts identified by comparison against a clinically-relevant pathogen list comprising 1,005 species[70]. Based on this dataset, the hazard score (Q) for each sample was calculated using multidimensional risk vector normalization in Euclidean space[32] as follows:

$$Q = \left[1 - \frac{\sqrt{\left(1-\frac{N_{ARG}}{N_{Contig}}\right)^2+\left(1-\frac{N_{ARG,MGE}}{N_{Contig}}\right)^2+\left(1-\frac{N_{ARG,MGE,PAT}}{N_{Contig}}\right)^2+\left(1-\frac{N_{ARG,PAT}}{N_{Contig}}\right)^2}}{2}\right] \times 10^2$$

Where N_{Contig} denotes the total number of contigs in the sample; N_{ARG} refers to the number of contigs containing only ARGs; $N_{ARG,MGE}$ represents the number of contigs containing both ARGs and MGEs (physical co-localization of ARGs and MGEs within a 5 kb distance on the same contig); $N_{ARG,MGE,PAT}$ indicates the number of contigs harboring ARGs and MGEs within genomes of pathogens; $N_{ARG,PAT}$ represents the number of contigs containing ARGs within pathogen genomes."

Reference:

[1] Rumi MA, Oh M, Davis BC, Brown CL, Juvekar A, Vikesland PJ, Pruden A, Zhang L. MetaCompare 2.0: Differential ranking of ecological and human health resistome risks. *FEMS Microbiol Ecol* 100, fiae155 (2024).

14. *Line 393. Is there a theoretical basis for the risk-based prioritization described in this section? The criteria used do not appear to be grounded in an established theoretical framework, and the rationale behind their selection is not clearly explained.*

Response: Thanks for this considerate question. Yes, our risk-based prioritization is grounded in established theoretical and empirical foundations. Martínez et al.[1] demonstrated that the health risk posed by ARGs depends not only on their detection but on the integrated assessment of three interdependent factors: prevalence, dissemination potential, and clinical relevance—such that high-risk ARGs are characterized by high abundance, genetic mobility (e.g., *via* MGEs), and association with clinically significant pathogens. Building on this conceptual framework, Zhang et al.[2] translated these principles into a computationally tractable, sequence-based scoring system comprising three quantifiable criteria: (*i*) ARG abundance, (*ii*) co-localization of ARGs with MGEs, and (*iii*) taxonomic association with known pathogenic hosts. Our study applies this validated tripartite criterion to prioritize ARGs for targeted surveillance, thereby ensuring methodological consistency with current best practices in environmental resistome risk assessment. The theoretical foundation for this risk-based prioritization framework has been explicitly introduced at the outset of the section (**Lines 396–400**).

Lines 396-400: "Given that the prevalence and dynamic characteristics of sewage-derived ARGs in the UWS may pose varying degrees of public health threats, we established a risk ranking framework aimed at identifying high-risk ARGs for targeted monitoring, grounded in a conceptual AMR risk paradigm in which health risk is determined jointly by ARG abundance, transmissibility, and clinical relevance[33, 34]."

Reference:

[1] Martínez JL, Coque TM, Baquero F. What is a resistance gene? Ranking risk in resistomes. *Nat Rev Microbiol* 13, 116-123 (2015).

[2] Zhang AN, Gaston JM, Dai CL, Zhao S, Poyet M, Groussin M, et al. An omics-based framework for assessing the health risk of antimicrobial resistance genes. *Nat Commun* 12, 4765 (2021).

15. *The Discussion section largely restates the results without providing meaningful mechanistic interpretation. Strengthening this section with deeper analysis and a clearer theoretical framework is necessary to improve the scientific rigor of the study. As noted above, this is essential for moving the work beyond a descriptive case study and highlighting its broader relevance and contribution to the field.*

Response: We thank the reviewer for this insightful comment. To address this, the manuscript has been revised accordingly: (1) Redundant results have been removed and detailed data presentations condensed; (2) The original second and third paragraphs have been merged and refined to improve logical flow; (3) More in-depth mechanistic interpretations regarding ecological connectivity and horizontal gene transfer have been added (**Lines 428–432**, **466–470**, **487–494**, **508–514**). These revisions have strengthened the theoretical framework and scientific rigor of the study. We are grateful for the reviewer's constructive guidance.

Lines 428-432: "The ecological connectivity of AMR contaminants along the urban water continuum was demonstrated by systematic analyses conducted across three hierarchical biological levels "resistome-subtype-sequence", indicating that continuous sewage discharge introduces contaminants into natural waters which can subsequently enter drinking water supplies through ecological connectivity."

Lines 466-470: "Future integration of disinfection dynamics, including disinfection intensity and residual chlorine in distribution systems, with upstream source control may further strengthen a coordinated dual-end strategy, offering a more comprehensive paradigm for safeguarding public health from AMR-related hazards across the urban water system."

Lines 487-494: "Our uncovered genetic co-occurrence of ARGs and MGEs, along with putative HGT events inferred from gene flow analysis, underscores the critical role of MGEs in contributing to the prosperity of ARGs throughout the UWS. However, the precise timing, spatial origin, and molecular pathways of these HGT events remain unresolved, owing to inherent limitations in current similarity-based approaches for HGT inference. The presence of such transferable ARGs in clinically relevant contexts raises concerns, as the WHO-listed priority pathogens observed in this study were found to harbor multiple ARGs and MGEs, which may significantly constrain therapeutic options for severe infections and increase the risks of morbidity and mortality[46]."

Lines 508-514: "The observation of both conserved and compartment-specific integron-associated ARG arrays across the UWS indicates concurrent downstream dissemination of stable MGE–ARG combinations and local cassette rearrangements mediated by HGT. Consequently, the xenogenetic evolution of ARG-associated integrons within the UWS, particularly mediated by selective pressures from physicochemical parameters and pollutants, plasmids-borne localization, and carriage by potential pathogens, necessitates investigation beyond conventional *int*I gene-based quantitation."

Reviewer #3:

1. *The authors thoroughly investigate the prevalence of ARGs across the urban water system from sewage to drinking water. While a case study, the intense interrogation of 32 water samples across this spectrum using 16S rRNA sequencing, metagenomics, and culturomics, reveals that some treatment steps may select for ARGs. Several types of ARGs were also uniquely appearing only after treatment. The enrichment of several type of ARGs could be attributed to both biotic and abiotic factors, and drivers of the resistome were identified. The most concerning ARGs were also identified, setting a potential framework for ARG monitoring in urban water streams. This is a valuable study, especially as ARG monitoring is of increasing popularity. However, some of the conclusions may be overstated, and some methods can be clarified.*

> **Response:** We sincerely appreciate the positive comments and valuable suggestions, all of which have significantly contributed to strengthening the rigor and clarity of our manuscript. In response, we have carefully refined the conclusions to ensure they are precisely supported by the evidence presented and have expanded the methodological section to enhance transparency, methodological rigor, and reproducibility. A comprehensive, point-by-point response to each comment is provided below.

2. Relative vs Absolute:

The authors rely on quantification of ARGs in copies/cell, which is a measure of relative abundance. This metric does not take into account the absolute quantification of cells or ARGs in each water tyle (e.g., copies/mL, cells/mL of water), and thus misses a major contextual point that needs to be discussed around water treatment. Do the authors have the ability to make such quantifications? If yes, it is helpful to include this context, even as a separate discussion, with enrichment amongst surviving cells (relative abundance) and overall exposure reduction (absolute abundance) as unique topics. Readers may take away from the manuscript that exposure to ARGs in untreated (or non-disinfected) water is less than that of treated/disinfected water, which may not be true from the perspective of total exposure, given the orders of magnitude reduction in cell counts during treatment. When discussing risk, it is paramount to include absolute quantification.

In defending the measure ARG copies/cell, the authors cite a study which also states that "Both absolute and relative measures are essential for a comprehensive evaluation of antibiotic resistance in environmental settings and provide complimentary insights, e.g., regarding mass flows and exposure risks, on one hand, and processes of resistance selection and comparison of widely different environments (e.g., dense gut microbiomes versus water), on the other." This study also mentions that transformation is easy, and thus, presenting transformed data (e.g., even roughly by multiplying by 16S rRNA gene copies/L) should be feasible. As raw wastewater and clean drinking water are vastly different in absolute density, it is critical to present absolute measures to support the broad and, at present, overstated conclusions.

The inclusion of this data will verify that water treatment is working as intended across

each treatment step, and will help some of the discussion of results make more sense. While it is understandable that both drinking and wastewater treatment select for microbes that are resistant to disinfectants and antibiotics via relative abundance (by removing less resistant cells), the actual concentration of total cells should concurrently decrease, and thus total potential exposure to ARGs may also decrease. Further, the increase in resistance between treatment and tap is potentially even more concerning, as this likely accompanies an increase in total cell concentration.

Response: We sincerely thank the reviewer for this insightful comment. We fully agree with the reviewer that expressing ARG abundances as gene copies per cell reflects relative enrichment rather than absolute exposure levels—a distinction of particular importance in water treatment contexts, where microbial biomass and viable cell counts are reduced by several orders of magnitude. Furthermore, we acknowledge that exclusive reliance on relative abundance metrics may engender interpretive ambiguity. Accordingly, we performed three complementary analyses to address this concern: (1) transformation of relative ARG abundances to absolute concentrations (copies/mL); (2) quantification of absolute ARG dynamics across water compartments; and (3) revision of risk interpretation to emphasize persistence rather than amplification.

(1) Unit Transformation. We performed a sample volume- and biomass-normalized transformation to estimate the absolute abundance of ARGs, expressed as gene copies per milliliter (copies/mL), using a mass-conservation-constrained methodology[1, 2]. The detailed procedure and associated equations have been supplemented (**Supplementary Method 7**), as follows.

Absolute ARG concentrations (C_{abs}, copies·mL^{-1}) were calculated from relative abundances using a mass conservation-based approach[17,18]:

$$C_{abs} = \frac{C_{rel} \times N_{cell}}{V_{sample}} \times \frac{M_{extract}}{M_{seq}}$$

where C_{rel} is the relative ARG abundance (copies per cell) calculated based on metagenomic sequencing datasets, N_{cell} represents the cell number in the sequenced reads quantified by ARGs-OAP on the alignment of essential single-copy marker genes, V_{sample} is the volume of water filtered (mL), $M_{extract}$ is the total DNA mass extracted from the sample, and M_{seq} is the DNA mass represented by the sequencing reads (gram, as double-stranded DNA), calculated from the nucleotide composition as:

$$M_{seq} = \frac{2 \times \sum(n_i \times M_i)}{N_A}$$

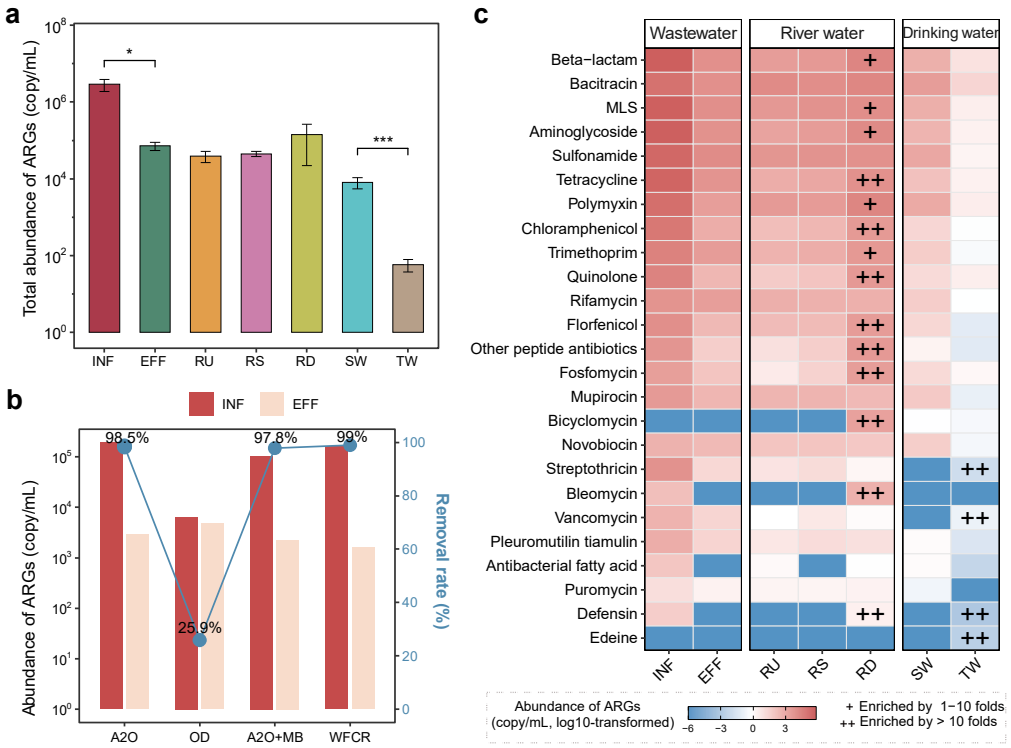
with n_i being the total count of nucleotide i (A, T, C, G) in the sequencing reads, M_i represents its molecular weight (313.2, 304.2, 289.2, 329.2 g·mol^{-1}), and N_A is Avogadro's constant.

(2) Absolute abundance dynamics of ARGs. After conversion, we determined the absolute abundances of ARGs (copies/mL) across all water compartments. Please see the following figure. The results showed that wastewater treatment processes efficiently removed ARGs by an average of 97.5%. Despite this, effluent discharge into receiving water caused a 2.69-fold increase in absolute ARG abundances in

downstream river water, with 14 subtypes showing particularly elevated levels. In contrast with relative abundance, drinking water disinfection process significantly decreased the absolute abundances of ARGs by 99.3%, while four types of ARGs (streptothricin, vancomycin, defensin and edeine) were significantly enriched by > 1-fold. As the reviewer pointed out, the actual concentration of total cells in drinking water should concurrently decrease, and thus total potential exposure to ARGs may also decrease; and the absolute abundance results can be incorporated into the text, which indicates the overall reduction in exposure, complementary to the relative abundance indicating the enrichment among the surviving cells (**Supplementary Fig. 2**, **Lines 148–151**, **178–181**).

(3) Concerns regarding increased risk in drinking water. The primary focus of this study is to identify sewage-derived ARGs that persist through the urban water pathway and remain detectable in drinking water, thereby informing source control and targeted treatment strategies. To avoid overinterpretation, we have reduced emphasis on risk amplification and instead highlight ARG persistence, and all associated comparisons now explicitly report abundance as relative abundance or on a per-cell basis to ensure accurate interpretation. These revisions have been incorporated throughout the manuscript.

We sincerely thank the reviewer for this insightful guidance, which has significantly strengthened our work.



Supplementary Fig. 2 | Absolute abundance dynamics of ARGs across urban water compartments. a, Shifts in the absolute abundance of ARGs (copies/mL) within the total microbial community across different water compartments. **b**, Absolute abundances and removal efficiencies of ARGs across wastewater treatment processes. *A2O: anaerobic-anoxic-oxic; OD: oxidation ditch process; MBR: membrane bio-reactor; FCR: food chain reactor.* **c**, Changes in

the absolute abundance of major ARG types across water compartments. Symbols '+' and '++' denote the magnitude of ARG enrichment in key comparisons: downstream *vs.* upstream receiving water (RD/RU), and tap water *vs.* source water (TW/SW). *INF: influent; EFF: effluent; RU: upstream of receiving water; RS: outlet site of receiving water; RD: downstream of receiving water; SW: source water; TW: tap water.*

Lines 148-151: "Notably, while the disinfection process drastically reduced the absolute abundances of ARGs (copies/mL) in tap water by 99.3% (Supplementary Fig. 2), it concurrently led to a statistically significant 26.1% increase in their relative abundance (copies/cell)."

Lines 178-181: "Thus, while water treatment processes substantially reduce total bacterial populations, the selective pressures exerted may disproportionately eliminate susceptible bacteria, enriching resistant phenotypes and ARG-harboring taxa among survivors across the urban water continuum."

Reference:

[1] Yin XL, Yang Y, Deng Y, Huang Y, Li L, Chan LYL, Zhang T. An assessment of resistome and mobilome in wastewater treatment plants through temporal and spatial metagenomic analysis. *Water Res* 209, 117885 (2022).

[2] Jiang XT, Yu K, Li LG, Yin XL, Li AD, Zhang T. AQMM: Enabling absolute quantification of metagenome and metatranscriptome. *bioRxiv* 218347 (2017).

3. *As the authors only present relative abundance, care should also be taken discussion of increases in relative abundance vs that of absolute abundance. For example L226-228 "... with its abundance further elevated by 7.0 fold following drinking water disinfection" is referring to an increase in relative abundance, and thus proportion or prevalence may be a better word than abundance. In another example L260 – "...between HGT event rates and ARG abundance across samples" – this is ARG prevalence or ARG relative abundance. This nuance must be cared for throughout the text, as it was confusing to read as is.*

Response: Thanks for highlighting this terminological imprecision. Accordingly, throughout the manuscript, the descriptions of increases or abundance have been clarified or revised to "ARG prevalence" and "ARG relative abundance", respectively.

Lines 234-236: "Particularly, the polymyxin resistance gene *pmr*F showed a 643.5-fold enrichment downstream and a 7.0-fold increase in relative prevalence following drinking water disinfection."

Lines 295-297: "Further correlation analysis revealed a significant positive relationship between putative HGT event rates and ARG relative abundance across samples ($r^2 = 0.74$, $p < 0.001$), indicating patterns consistent with MGE-driven resistome dynamics."

4. *L 234 – can the authors expand on what they mean by etc.? Typically an e.g., for example list would not include this denotation.*

Response: We thank the reviewer for pointing out this issue. Yes, "e.g." (such as) should not be followed by "etc." (and so on). For clarification, we have removed

"etc." in the revised manuscript (**Line 243**).

> **Lines 242-244**: "Both biotic (e.g., proliferation of bacterial hosts or HGTs mediated by MGEs) and abiotic factors (e.g., temperature, pH, and pollutants) may contribute to the development and dissemination of ARGs in UWS."

5. *L 323 – Define BAC, PCMX at first mention in the text. Please note that BAC can be confused for biological activated carbon in the drinking water context.*

Response: Thanks for this insightful comment. To avoid misleading, we now define BAC as benzalkonium chloride and PCMX as para-chloro-meta-xylenol at first mention in the revised manuscript (**Line 336**).

> **Lines 333-338**: "Mantel test analyses revealed significant associations ($p < 0.05$) between ARG abundance profiles and five water quality parameters (pH, DO, COD, TP, and NH_3-N), six types of trace metals (Mg, Al, Mn, Fe, Cu, and Zn), two common disinfectants (benzalkonium chloride, BAC; para-chloro-meta-xylenol, PCMX), and two widely-used antibiotics (sulfamethazine, SMZ; oxytetracycline, OTC), while no significant correlation was observed for antidepressants."

6. *L357 – "First ARGs...drinking water." This sentence can likely be rearranged and/or shortened to make more sense. For example...: We defined sewage-derived ARGs as those that are discharged into natural waters through effluent, are subsequently enriched in the receiving waters, and ultimately persist into tap water. We used the following criteria:...". The first sentences of each of these sections have similar issues with long sentences that can be made more direct.*

Response: We thank the reviewer for the constructive suggestion. Following this advice, we have revised the indicated sections to improve conciseness and clarity (**Lines 370–372, 382–384**, **396–400**).

> **Lines 370-372**: "We defined sewage-derived ARGs as those that are discharged into natural waters through effluent, are subsequently enriched in the receiving waters, and ultimately persist into tap water. We used the following criteria:..."
> **Lines 382-384**: "To evaluate the health hazards associated with sewage-derived AMR along the urban water transmission pathway, we developed a genome-informed, system-scale comparative risk indicator (Q)."
> **Lines 396-400**: "Given that the prevalence and dynamic characteristics of sewage-derived ARGs in the UWS may pose varying degrees of public health threats, we established a risk ranking framework aimed at identifying high-risk ARGs for targeted monitoring, grounded in a conceptual AMR risk paradigm in which health risk is determined jointly by ARG abundance, transmissibility, and clinical relevance[33,34]."

7. *L362 – can the authors define a what "detectable" in household drinking water means? Given the unique collection methods from this study (>2,000 L) compared to typical drinking water monitoring methods (with defined volumes that are often lower), is it likely that ARGs are often undetectable?*

Response: Thanks for this thoughtful question. The term "detectable" reflects the probabilistic detection of ARG signals in our metagenomic dataset. Given a sequencing depth of 10 Gb per sample (corresponding to approximately 5.5 ng of sequenced DNA), detection is inherently subject to stochastic variation. In practice, an average of 1,530 ng DNA was extracted from these household drinking water samples. This suggests that our observed ARG prevalence corresponds to a sample volume of approximately 7.2 L of drinking water (assuming 100% extraction efficiency for the 2,000-liter water sample). Therefore, the actual sample volume in this study falls within the higher range commonly used for pathogen detection in routine drinking water monitoring. We intentionally adopted this strategy to maximize sensitivity for capturing trace-level, sewage-derived ARG contamination, thereby providing a robust basis for source tracking and informing future source control measures.

Meanwhile, the pronounced heterogeneity in relative abundance among detected ARG categories underscores their differential ecological relevance and public health significance. In Figure 7a, we therefore highlighted ARG types with relative abundances exceeding 0.1 copies/1000 cells. A total of 35 such ARG subtypes were identified, accounting for 57% of all sewage-derived ARGs detected in this study. We hope these findings can provide a reference for the future development of standardized AMR detection and monitoring protocols in drinking water.

For clarification, we have appended the definition of "detectable" with the parenthetical note "ARG signals in metagenomic datasets" in the revised manuscript (**Line 375–376**).

> **Lines 375-376**: "(3) being detectable in household drinking water, indicated by ARG signals in metagenomic datasets."

8. *L370 – "...based on which to develop..." should potentially be "which was used to develop"*

Response: Thanks for this considerate comment. As suggested by the 6th comment of the reviewer, the first sentence of this section has been thoroughly optimized as follows:

> **Lines 382-384**: "To evaluate the health hazards associated with sewage-derived AMR along the urban water transmission pathway, we developed a genome-informed, system-scale comparative risk indicator (Q)."

9. *L 381... - Does the quantitative health risk assessment take into account any absolute quantification, or only relative? The equation seems to indicate relative abundances, but the figure does not specify. If only relative, is this typical? Is exposure taken into account? (e.g., how are people typically exposed to drinking water vs wastewater [is this only workers]??) Further, can the authors expand on the pathogens in drinking water that were found and at what concentrations? More discussion is likely needed on this QMRA metric.*

Response: Thanks for these insightful and constructive questions.

(1) Does the risk assessment use absolute or relative ARG abundance?

This metric, adapted from the widely used MetaCompare framework[1], represents a genetic co-occurrence-based risk potential index that integrates the physical linkage of ARGs with mobile genetic elements (MGEs) and their presence within pathogen genomes at the contig level. It quantifies the horizontal transfer potential and pathogen-associated health relevance of ARGs, independent of absolute or relative abundance, enabling direct comparisons across samples. Meanwhile, our primary focus is on quantifying the inherent hazard (or risk potential) of these sewage-derived ARGs along the pollution transmission pathway. Thus, as the reviewer correctly noted, exposure is not incorporated. For clarification, we have reorganized this section (**Lines 382–391**) and provided a more detailed description of the calculation model in the Methods (**Lines 716–731**).

Lines 382-391: "**Sewage-derived ARG-associated health risk assessment.** To evaluate the health hazards associated with sewage-derived AMR along the urban water transmission pathway, we developed a genome-informed, system-scale comparative risk indicator (Q). This approach is based on the dataset of 61 identified sewage-derived ARGs and integrated three genetic risk determinants including ARG abundance, mobility potential (colocalization with MGEs), and association with pathogenic hostsinto a multidimensional framework, which was adapted from the MetaCompare pipeline[32]. This enables direct comparison of resistome-associated hazards across different water compartments by quantifying the relative hazard potential based on contig-level genetic features, independent of absolute abundance or human exposure pathways (Fig. 7b; detailed calculation provided in Methods)."

Lines 716-731: "**Quantitative model for sewage-derived ARG-associated health hazard assessment.** To quantify health hazards associated with sewage-derived resistomes across different urban water compartments, we developed a quantitative hazard score (Q) model adapted from the MetaCompare framework[32]. The 61 identified anthropogenic ARGs were included to construct a sewage-derived ARGs dataset, based on which to integrate their genetic risk determinants at the contig level: (*i*) ARG relative abundance, (*ii*) ARG mobility potential indicated by physical colocalization with MGEs within a 5 kb distance on the same contig[83], and (*iii*) ARG association with pathogenic hosts identified by comparison against a clinically-relevant pathogen list comprising 1,005 species[70]. Based on this dataset, the hazard score (Q) for each sample was calculated using multidimensional risk vector normalization in Euclidean space[32] as follows:

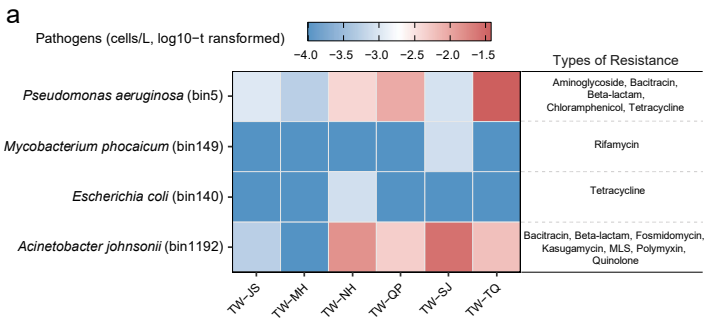
$$Q = \left[1 - \frac{\sqrt{\left(1-\frac{N_{ARG}}{N_{Contig}}\right)^2+\left(1-\frac{N_{ARG,MGE}}{N_{Contig}}\right)^2+\left(1-\frac{N_{ARG,MGE,PAT}}{N_{Contig}}\right)^2+\left(1-\frac{N_{ARG,PAT}}{N_{Contig}}\right)^2}}{2}\right] \times 10^2$$

Where N_{Contig} denotes the total number of contigs in the sample; N_{ARG} refers to the number of contigs containing only ARGs; $N_{ARG,MGE}$ represents the number of contigs containing both ARGs and MGEs (physical co-localization of ARGs and MGEs within a 5 kb distance on the same contig); $N_{ARG,MGE,PAT}$ indicates the number of contigs harboring ARGs and MGEs within genomes of pathogens; $N_{ARG,PAT}$ represents the number of contigs containing ARGs within pathogen genomes."

(2) Quantification of pathogens in drinking water

In the original manuscript, we employed the widely used metagenomic relative quantification approach to quantify the abundance of metagenomic-assembled genomes (MAGs) classified as pathogens, with unit expressed as RPKM (Fig. 4a). Using this method, we detected an average total abundance of 3.63 RPKM in tap water samples, encompassing four bacterial species, including *Acinetobacter johnsonii*, *Escherichia coli*, *Mycobacterium phocaicum*, *Pseudomonas aeruginosa*.

In response to the reviewer's suggestion, we further estimated the absolute concentrations of these pathogens in tap water using the adapted mass conservation-based absolute quantification framework described in our response to your Comment #2 and Supplementary Methods. The resulting absolute concentrations for each sample are shown in the figure below. On average, the total concentration of these species in tap water was 0.027 cells/L, all of which carried diverse ARGs. Notably, *P. aeruginosa* which harbored ARGs conferring resistance to five classes of antibiotics was detected in all samples. The widespread presence of these pathogens and the sewage-derived ARGs they carry warrants particular attention and targeted control measures. For optimization, the associated results are now supplemented as **Supplementary Fig. 17a** in the revised manuscript.



Supplementary Fig. 17a | Pathogenic hosts and associated ARGs detected in drinking water. Heatmap illustrating the cell-equivalent concentrations of pathogen-associated metagenome-assembled genomes (MAGs) in tap water samples. Values are expressed as cells/L and $\log_{10}$-transformed. The right panel lists the ARG classes identified in each MAG by genome annotation.

Reference:

[1] Rumi MA, Oh M, Davis BC, Brown CL, Juvekar A, Vikesland PJ, Pruden A, Zhang L. MetaCompare 2.0: Differential ranking of ecological and human health resistome risks. *FEMS Microbiol Ecol* **100**, fiae155 (2024).

10. *L390 – "While subsequent chlorination..." – is this mitigation of risks due to the overall reduction in cell counts? Is the conclusion ultimately that treatment is effective at reducing risks? If yes, this may need to be emphasized in final discussion/conclusions.*

Response: Thanks for this insightful comment. As noted in the above response, the MetaCompare-based risk score (Q) captures the structural hazard potential of the resistome by integrating the physical linkage of ARGs with MGEs and their presence within pathogen genomes at the contig level. Therefore, this observed reduction in risk score suggests that treatment is effective at mitigating the MGE-

and pathogens-associated AMR risks posed by sewage-derived ARGs. Importantly, as the reviewer noted, this reduction is accompanied by a substantial decrease in total cell counts. Together, these complementary lines of evidence demonstrate that treatment is effective at mitigating both the mobility- and pathogen-associated hazard. Following the reviewer's suggestion, we have now emphasized this conclusion in the revised Discussion associated with drinking water (**Lines 394–395, 461–466**).

> **Lines 394-395**: "Subsequent chlorination further mitigated AMR-associated health hazards in tap water, though residual hazards remained at 0.61 ± 0.05."
>
> **Lines 461-466**: "The framework developed here for monitoring AMR dynamics within the UWS enables the precise identification of prevalent or terminally enriched sewage-derived risk-associated ARGs, as well as assessment of treatment efficacy in mitigating mobility- and pathogen-associated AMR hazard of anthropogenic ARGs, providing valuable insights for source-oriented intervention strategies and regulatory standards aimed at ensuring drinking water safety."

11. *L408 – a subsequent 3.3-fold increase in tap water relative to source drinking water – is this an increase in relative abundance? Or does absolute concentration also increase?*

Response: We thank the reviewer for highlighting this ambiguity in data description. Yes, the observed increase refers to relative abundance, suggesting that disinfection treatment may select for its bacterial hosts or facilitate associated HGTs. As the reviewer suggested, we also examined its absolute concentration, which revealed a substantial reduction in exposure levels. As stated in the end of Discussion (Lines 532–535), this study primarily adopts "copies per cell" as the unit of measurement because it is widely applied and recommended standard by researchers globally for quantification of AMR risks[1]; More importantly, it is particularly valuable for investigating the microscale mechanisms underlying AMR proliferation and dissemination within urban water system, including vertical transmission and HGT.

To avoid misinterpretation, we have revised the manuscript to emphasize ARG persistence in tap water rather than risk amplification, and all associated comparisons now explicitly report abundance as relative abundance or on a per-cell basis. These revisions have been incorporated throughout the manuscript.

Reference:

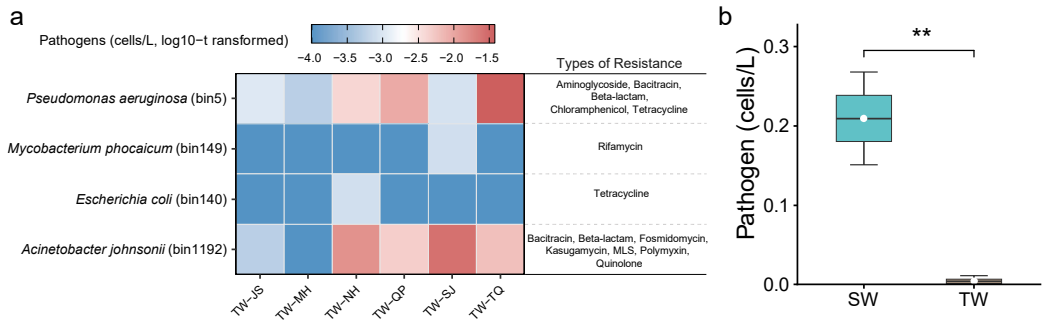
[1] Yin XL, Chen X, Jiang XT, Yang Y, Li B, Shum MH, et al. Toward a universal unit for quantification of antibiotic resistance genes in environmental samples. *Environ Sci Technol* 57, 9713-9721 (2023).

12. *L454-L470 – can the authors comment on the prevalence of pathogens in this drinking water? The authors at some point mention E. coli and P. aeruginosa, but often in developed countries' urban water systems, E. coli is often no longer prevalent in treated (disinfected) water, and P. aeruginosa is inhaled as a primary infection route, which may change ARG implications contrary to the discussion. If invoking*

consumption of pathogens containing ARGs as a risk, it is important to also divulge the quantification of pathogens in this water system, as it may be unique. The authors further mention boiling water, which is not often a requirement in many developed countries. The level of disinfection, and maintenance of disinfectant needs to be further explained, as this context may have strong implications on the results. This is also where absolute quantification of cells will help put these results into a larger context.

Response: We thank the reviewer for this thoughtful and important comment regarding the prevalence of pathogens, exposure pathways, and the role of disinfection in the drinking water system. In response, we have **(*i*)** supplemented the absolute quantification of pathogens in drinking water, which was estimated based on metagenomic relative abundance, sample volume and extracted DNA biomass. We have also **(*ii*)** added a discussion on the disinfection process, including disinfection intensity and residual disinfectant maintenance, and **(*iii*)** removed the mention of boiling water to improve the generalizability of the discussion.

Absolute quantification of pathogens in drinking water. In response to the reviewer's Comment #9, in addition to metagenomic relative quantification, we further estimated the absolute concentrations as suggested. Four bacterial species were detected, including *Acinetobacter johnsonii*, *Escherichia coli*, *Mycobacterium phocaicum*, and *Pseudomonas aeruginosa*, with at least one pathogenic species present in each sample. The average total concentration of these pathogens in tap water was 2.67 cells/100 L, and all carried diverse ARGs. Notably, *P. aeruginosa*, which harbored resistance genes to five antibiotic classes, was found in all samples. Among the four opportunistic pathogens detected in tap water, *A. johnsonii* mainly threatens immunocompromised individuals and may cause skin, urinary tract, or bloodstream infections[1]. *E. coli* indicates potential fecal contamination and may induce gastroenteritis and diarrhea, especially in susceptible populations[2]. *M. phocaicum*, a nontuberculous mycobacterium resistant to disinfection, can cause pulmonary infections *via* inhalation[3]. *P. aeruginosa* poses a considerable infection risk to immunocompromised individuals, particularly when carrying multidrug resistance[4]. These opportunistic pathogens in drinking water pose a threat to human health through exposure pathways such as ingestion, aerosols, and skin contact[5]. Accordingly, the absolute quantification results have been supplemented in **Supplementary Fig. 17**.



Supplementary Fig. 17 | (Opportunistic) pathogens detected in drinking water. a, Heatmap illustrating the cell-equivalent concentrations of pathogen-associated metagenome-assembled

genomes (MAGs) in tap water samples. Values are expressed as cells/L and $\log_{10}$-transformed. The right panel lists ARG classes identified in each MAG by genome annotation. **b**, Boxplots comparing total pathogen concentrations in source water (SW) and tap water (TW). Boxes represent the interquartile range, center lines indicate medians, whiskers denote the data range, and white dots indicate mean values. Statistical significance was determined by Student's t-test (**$p \leq 0.01$).

Risks associated with the detection level of pathogens in drinking water. As noted in our response to the reviewer's Comment #7, the actual screened volume per sample for metagenomic analysis was approximately 7.2 L. Therefore, for routine culture-based monitoring of culturable pathogens in drinking water (typically 100 mL, or 1 L for targeting specific pathogenic species), the failure to detect pathogens, as mentioned by the reviewer, may be attributed mainly to two reasons: **(1) Detection stochasticity:** at low cell densities, sampling a smaller volume increases the probability of false negatives. **(2) Microbial physiology:** chlorination may induce a viable but nonculturable (VBNC) state in some bacteria[6], rendering them undetectable by culture-based methods while still detectable *via* DNA-based approaches like metagenomics.

Although the present study focused on a specific region to characterize resistome dynamics in UWS, our prior work has documented the presence of opportunistic pathogens (including *A. pittii*, *A. junii*, and *M. gordonae*) in drinking water supplies of developed countries[7, 8]. Notably, the methodology employed in this study, integrating large-volume water sampling with metagenomic sequencing, achieves an enhanced limit of detection, enables early risk identification, and provides a scientific basis for routine environmental monitoring. Furthermore, to enhance the generalizability of the discussion, we have removed the text regarding "boiling water", as noted by the reviewer (**Line 458**).

Disinfection intensity and residual disinfectant maintenance. We agree with the reviewer that the level of disinfection and the maintenance of disinfectant are important factors that may have strong implications for our findings. We acknowledge that although the present study is mainly focused on the systematic concept of AMR dynamics in key water compartments of urban water system, attempting to identify sewage-derived AMR indicators for future source control to ensure drinking water safety, the disinfection intensity played a key role in controlling these target indicators. Specifically, drinking water treatment plants in the studied megacity typically employ liquid chlorine (or sodium hypochlorite) as the primary disinfectant, with typical dosages ranging from 1.0 to 3.0 mg/L (expressed as free chlorine) depending on raw water quality and seasonal variations. This is complemented by advanced treatment processes such as ozonation and activated carbon filtration. In accordance with the national standard for drinking water quality (GB 5749–2022), a minimum contact time of ≥ 30 minutes is required to ensure effective disinfection. The residual chlorine concentration in finished water is typically maintained at $\geq$ 0.3 mg/L, while at the distal ends of the distribution network, it must remain $\geq$ 0.05 mg/L to ensure continuous microbial inactivation and prevent regrowth or contamination.

Although chlorine dosage and residual chlorine levels were not directly measured in the present study, which instead focused on the microbial community and resistome in tap water to link with dynamics in water system, in future research, we will specifically extend to investigate the effects of disinfection intensity and distribution system residual chlorine on sewage-derived ARGs and their bacterial hosts (particularly pathogens) identified in this study. We greatly appreciate this valuable comment raised by the reviewer, and believe that such follow-up research will help improve the effectiveness of terminal control. When combined with source-control strategies at the wastewater front end, these dual-end approaches hold promise for more effectively mitigating the potential health risks posed by these contaminants. This discussion has now been incorporated into the revised manuscript (**Lines 466–470**).

> **Lines 466-470**: "Future integration of disinfection dynamics, including disinfection intensity and residual chlorine in distribution systems, with upstream source control may further strengthen a coordinated dual-end strategy, offering a more comprehensive paradigm for safeguarding public health from AMR-related hazards across the urban water system."

Reference:

[1] Aguilar-Vera A, Bello-Lopez E, Pantoja-Nunez GI, Rodriguez-Lopez GM, Morales-Erasto V, Castillo-Ramirez S. *Acinetobacter junii*: An emerging One Health pathogen. *mSphere* 9, e0016224 (2024).

[2] Anastasi EM, Matthews B, Stratton HM, Katouli M. Pathogenic *Escherichia coli* found in sewage treatment plants and environmental waters. *Appl Environ Microbiol* 78, 5536-5541 (2012).

[3] Shachor-Meyouhas Y, Geffen Y, Arad-Cohen N, Zaidman I, Ben-Barak A, Davidson S, Kassis I. *Mycobacterium phocaicum* bacteremia: An emerging infection in pediatric hematology-oncology patients. *Pediatr Infect Dis J* 33, 1299-1301 (2014).

[4] Hernandez-Jimenez P, Lopez-Medrano F, Fernandez-Ruiz M, Silva JT, Corbella L, San-Juan R, et al. Risk factors and outcomes for multidrug resistant *Pseudomonas aeruginosa* infection in immunocompromised patients. *Antibiotics* 11, 1459 (2022).

[5] Hull NM, Ling F, Pinto AJ, Albertsen M, Jang HG, Hong PY, et al. Drinking water microbiome project: Is it time? *Trends Microbiol* 27, 670-677 (2019).

[6] Jiang Q, Li H, Wan K, Ye C, Yu X. Quantification and antibiotic resistance risk assessment of chlorination-residual viable/VBNC *Escherichia coli* and *Enterococcus* in on-site hospital wastewater treatment system. *Sci Total Environ* 872, 162139 (2023).

[7] Ma LP, Li B, Jiang XT, Wang YL, Xia Y, Li AD, Zhang T. Catalogue of antibiotic resistome and host-tracking in drinking water deciphered by a large scale survey. *Microbiome* 5, 154 (2017).

[8] Liu HF, Jiao PB, Guan L, Wang C, Zhang XX and Ma LP. Functional traits and health implications of the global household drinking-water microbiome retrieved using an integrative genome-centric approach. *Water Res* 250, 121094 (2024).

Methods:

13. *Can the authors clarify the quantity of samples from each location? From descriptions, it seems like 32 samples were analyzed – but n is difficult to follow. Figure 1 may be an opportunity to clarify n.*

Response: Thanks for the thoughtful suggestion. We then clarified the quantity (*n*)

of samples both in revised Methods (**Lines 578–588**) and **Figure 1a**, as follows. Water/sludge samples were collected in biological triplicates over a 24 h period, and subsequently pooled into a single composite sample to ensure representativeness. For tap water, which required sufficient biomass for metagenomic sequencing, each composite sample was obtained by continuous filtration over 24 hours, resulting in a total filtered volume exceeding 2,000 L.

Lines 578-588: "Specifically, water/sludge samples were collected from the influent (INF, *n* = 4), activated sludge (AS, *n* = 4), and effluent (EFF, *n* = 4) of four primary full-scale WWTPs with a total treatment capacity of 375,000 m³/day that employ distinct treatment processes (anaerobic-anoxic-oxic, oxidation ditch, membrane bio-reactor, and food chain reactor). Meanwhile, the receiving rivers include three sampling points: 300 m upstream (RU, *n* = 4), 10 m downstream site (RS, *n* = 4), and 300 m downstream (RD, *n* = 4) of the discharge of WWTP effluent, representing areas unaffected by sewage, directly impacted by sewage, and subject to long-distance sewage influence, respectively. Source water (SW, *n* = 2) samples were collected from two main local source water reservoirs. Above samples were collected in biological triplicates over a 24 h period, and subsequently mixed as a representative composite sample. Tap waters (TW, *n* = 6) were collected from six administration districts across the city."

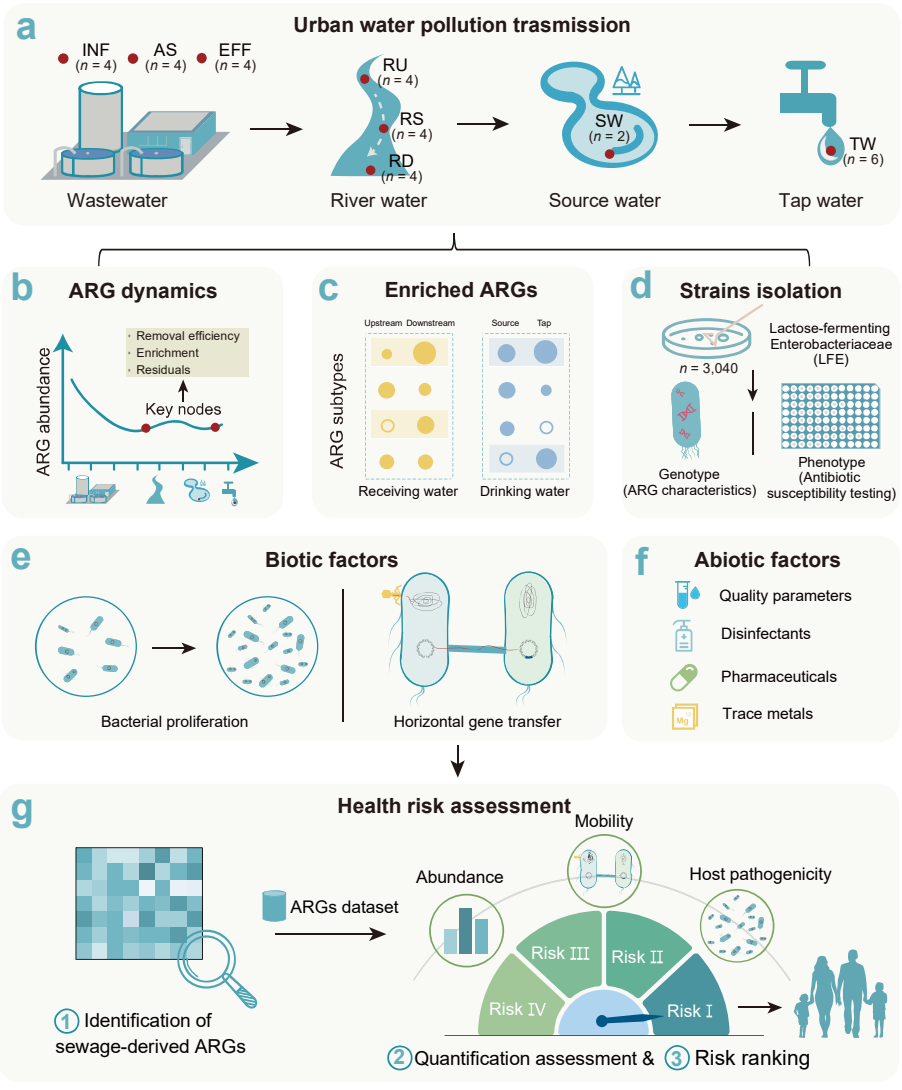


Fig. 1 | Schematic analysis of AMR risks in urban water system. a, The study focused on key water compartments within a megacity, including wastewater treatment plants, receiving river systems, drinking water reservoirs, and tap water distribution networks. Samples were collected in biological triplicates over a 24 h period, and subsequently pooled into a single composite sample to ensure representativeness. For tap water, which required sufficient biomass for metagenomic sequencing, each composite sample was obtained by continuous filtration over 24 hours, resulting in a total filtered volume exceeding 2,000 L. *n* represents the number of composite samples analyzed. *INF: influent, AS: activated sludge, EFF: effluent, RU: upstream, RS: receiving site, RD: downstream, SW: source water, TW: tap water.* **b**, Metagenomic-based profiling revealed resistome dynamics in urban water system, and spatial critical points were identified based on characterized patterns of removal efficiency, enrichment and residuals. **c**, Enriched ARG subtypes were discerned in critical points. **d**, Metagenomic ARG profiling and intrinsic antibiotic susceptibility testing were performed using isolated 3,040 isolates of fecal-associated lactose-fermenting *Enterobacteriaceae* (LFE). **e**, Biotic factors that could drive ARG dynamics across urban water system were examined, including bacterial hosts and horizontal gene transfer mechanisms. **f**, Abiotic factors were identified from diverse water quality parameters, antibiotics, trace metals, disinfectants, non-antibiotic pharmaceuticals, based on Spearman correlation and Partial least squares structural equation modeling (PLS-SEM) analyses. **g**, We proposed a health risk assessment framework for AMR pollution in urban water systems, encompassing identification of sewage-derived ARGs, quantification assessment and risk ranking, aimed at source control and mitigation of health threats.

14. *Can the authors clarify the use of a filtration device for concentration of over 2,000 L? Was this device an R.O. membrane, or a cartridge filtration device? For either option, what was the time frame, and did these samples potentially enrich for filter-living microbes? How did the authors quantify total bacteria present in these samples, and can that be normalized to a water basis?*

Response: We sincerely thank the reviewer for raising these important questions.

(1) Methodological information clarified

The used filtration device is an R.O. membrane system (A.O. Smith, USA). The sampling duration was approximately 24 hours, with a constant flow rate of 1.5 $L \cdot min^{-1}$. Detailed procedure has been supplemented in Methods (**Lines 588–590**).

Lines 588-590: "Microorganisms were collected by filtering over 2,000 L of water using a domestic reverse osmosis (R.O.) purification system (A.O. Smith, USA) operated at a constant flow rate of 1.5 $L \cdot min^{-1}$ for approximately 24 h."

(2) Whether it could potentially enrich filter-living microbes

Compared with our prior sampling protocols—where cartridge-type water purifiers equipped with hollow-fiber membranes (pore size: 0.1 μm; Torayvino, Toray Industries Inc., Japan) required ~48 hours to process comparable volumes[1,2]—this RO-based approach enables both shorter sampling duration and significantly greater water volume processed, thereby enhancing microbial biomass recovery. Nevertheless, we acknowledge that even with this improved setup, the potential enrichment of filter-associated microbes raised by the reviewer, cannot be completely excluded. However, such concern could be negligible in this study for two main reasons: (1) The relatively short filtration period (~24 hours) and the oligotrophic conditions of the tap water substantially limit microbial growth during

filtration; (2) More importantly, this study focuses primarily on the relative abundance of both the resistome and associated bacterial hosts, which is less sensitive to minor variations in absolute biomass recovery. Large-volume filtration was essential to overcome the inherently low biomass in tap water and to ensure sufficient DNA yield for downstream metagenomic sequencing.

(3) Quantification of total bacteria

In the present study, microbial cell abundance was estimated from metagenomic datasets based on genome-equivalent counts inferred from essential single-copy marker genes, as detailed in the Supplementary Methods. In response to the reviewer's comment, the total microbial abundance inferred from metagenomic data is now theoretically converted to a per-volume water basis using a mass conservation-based approach[3,4] assuming 100% DNA extraction efficiency and neglecting potential enrichment effects during filtration. The conversion formula is provided below. Using this method, the estimated microbial cell count was obtained, with an average of 2.03×10^2 cells/mL in tap water.

$$Cell\ concentration\ (cells/mL) = \frac{N_{cell}}{V_{sample}} \times \frac{M_{extract}}{M_{seq}}$$

where $M_{extract}$ is the total mass DNA extracted from the sample (g), and M_{seq} indicates the mass of DNA represented by the sequencing reads (g), calculated based on the total base pairs sequenced and the average molecular weight of double-stranded DNA. N_{cell} represents the number of cell equivalents detected in the sequenced reads, estimated from the coverage of essential single-copy marker genes using ARGs-OAP. V_{sample} is the volume of water filtered (mL).

Reference:

[1] Liu HF, Jiao P, Guan L, Wang C, Zhang XX, Ma LP. Functional traits and health implications of the global household drinking-water microbiome retrieved using an integrative genome-centric approach. *Water Res* 250, 121094 (2024).

[2] Ma LP, Li B, Jiang XT, Wang YL, Xia Y, Li AD, Zhang T. Catalogue of antibiotic resistome and host-tracking in drinking water deciphered by a large scale survey. *Microbiome* 5, 154 (2017).

[3] Yin XL, Yang Y, Deng Y, Huang Y, Li L, Chan LYL, Zhang T. An assessment of resistome and mobilome in wastewater treatment plants through temporal and spatial metagenomic analysis. *Water Res* 209, 117885 (2022).

[4] Jiang XT, Yu K, Li LG, Yin XL, Li AD, Zhang T. AQMM: Enabling absolute quantification of metagenome and metatranscriptome. *bioRxiv* 218347 (2017).

15. *L583 – Tap waters were "designed"? Word choice here is likely wrong.*

Response: We thank the reviewer for the correction. "Designed" has been replaced with "collected" accordingly (**Lines 587–588**).

> **Lines 587-588**: "Tap water samples (TW, *n* = 6) were collected from six administrative districts across the city."

16. *Figure 1a: Defining n for samples may also be helpful in this figure.*

Response: Thanks for this helpful comment. Yes! In response to Comment #13, we have revised the manuscript accordingly. Please refer to the updated Figure 1a; the

definition of *n* for samples has also been added to the Methods section.

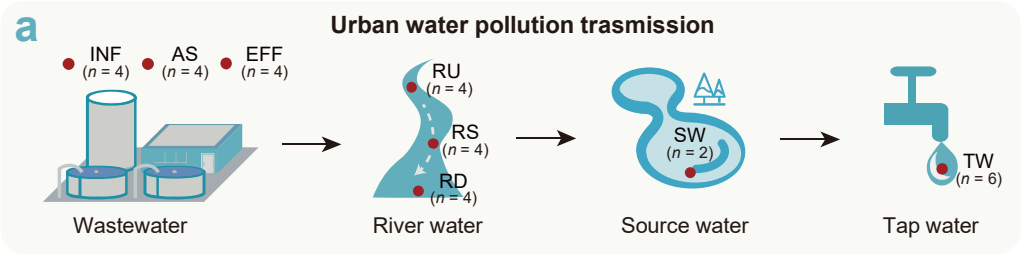


Fig. 1a | The study focused on key water compartments within a megacity, including wastewater treatment plants, receiving river systems, drinking water reservoirs, and tap water distribution networks. Samples were collected in biological triplicates over a 24 h period, and subsequently pooled into a single composite sample to ensure representativeness. For tap water, which required sufficient biomass for metagenomic sequencing, each composite sample was obtained by continuous filtration over 24 hours, resulting in a total filtered volume exceeding 2,000 L. *n* represents the number of composite samples analyzed. *INF: influent, AS: activated sludge, EFF: effluent, RU: upstream, RS: receiving site, RD: downstream, SW: source water, TW: tap water.*

17. *Figure 2: please define n clearly.*
Figure 2a – Can the authors expand on the meaning of: the ellipses (are these statistically determined), the numbers in circles, and arrows in the caption? As this figure is busy, it is suggested to remove the numbers in circles, which could be confused for another sample type, and potentially vary symbols to offer more distinction (e.g., squares for WW samples, circles for rivers, triangles for drinking water), which could also reinforce the classes used throughout.
Figure 2b/c – The establishment of baseline is somewhat unclear for the arrowed bars. It is suggested to add ends to the bars, or, better yet, define the range for the change with text (e.g. RU->RS, 8.4%^). Is there a reason that for RD the axis color changes (2c)?

Response: We thank the reviewer for the detailed and constructive comments on Figure 2, which have helped us substantially improve the clarity and interpretability of the results.

Following the reviewer's detailed suggestions, we have revised **Figure 2a** accordingly. Specifically, we have clearly defined *n* for samples, removed the numbers in circles, and introduced distinct symbols to improve differentiation: squares represent wastewater samples, circles denote river samples, and triangles indicate drinking water. This indeed greatly reinforce the classes used throughout. In addition, the extra explanations of "ellipses" and "arrows" were supplemented in the legends (**Lines 997–999**): the ellipses represent the 95% confidence intervals for each water compartment, calculated based on sample dispersion in the ordination space. The arrows indicate the inferred directional transitions of ARG profiles along the urban water system. Regarding **Figure 2b/c**, we have revised both figures and their captions to clarify the baseline for each arrowed bar by explicitly indicating the comparison ranges (e.g., RU→RD, SW→TW) and labeling the magnitude of change directly. The axis color in Figure 2c has been standardized for consistency. These improvements are reflected in the updated Figure 2.

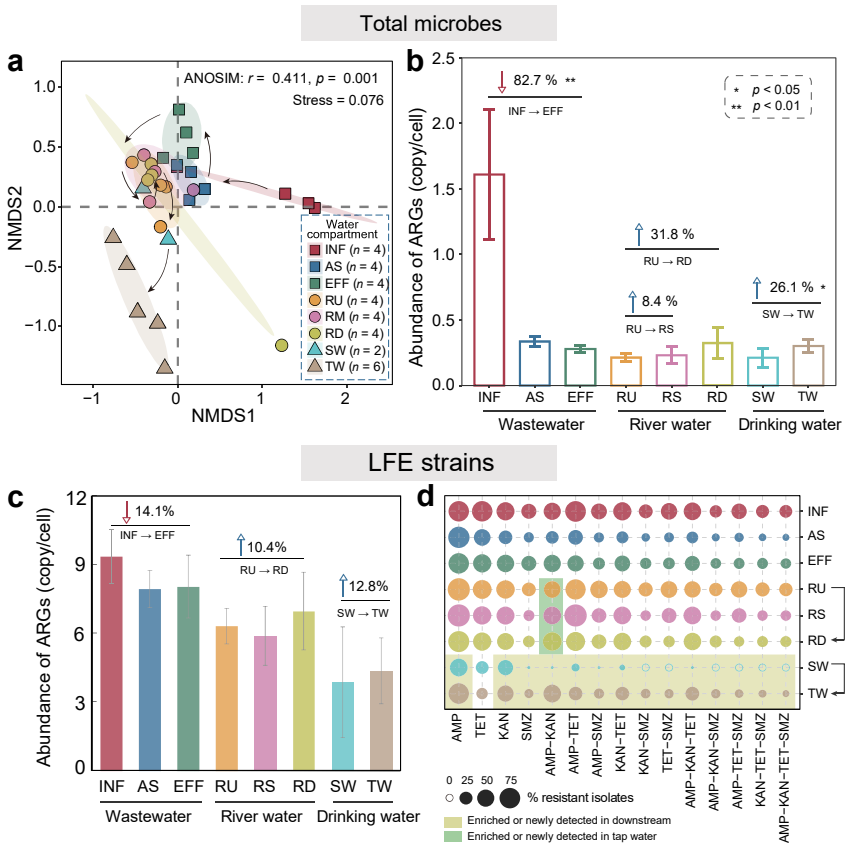


Fig. 2 | Dynamics of resistome across the urban water system. a, Non-metric multidimensional scaling (NMDS) analysis revealed statistically significant differences in ARG profiles of total microbes across water compartments, based on Bray-Curtis dissimilarity. Ellipses represent 95% confidence intervals for each water compartment. The arrows indicate the inferred directional transitions of ARG profiles along the urban water system. **b**, Shifts of ARG relative abundances (copy/cell) within total microbial community were observed across water compartment. Error bars represent the standard deviation. Arrows denote an increase/decrease relative to the baseline condition within each group, corresponding to sewage treatment efficiency, enrichment in the receiving water, and drinking water disinfection effects, respectively. **c**, Shifts of ARG relative abundances within LFE isolates were observed across water compartment. **d**, Multi-resistance of LFE strains in the urban water system. *AMP: ampicillin, TET: tetracycline, KAN: kanamycin, SMZ: sulfamethoxazole.*

18. *Figure 7/b: It may be useful to clarify which of these inputs was theoretical vs from data in the paper.*

Response: We thank the reviewer for this thoughtful comment. Accordingly, we have re-organized **Figure 7b** to enhance its conceptual clarity and structural transparency. Specifically, the figure is now explicitly partitioned into three logically distinct components: (*i*) the theoretical framework (top panel); (*ii*) the empirical data inputs (indicated alongside the arrows), including the genetic prevalence of ARGs, their inferred horizontal mobility potential, and the pathogenicity profiles of their bacterial hosts; and (*iii*) the resulting risk assessment outcomes (bottom panel). All data inputs refer to the 61 sewage-derived ARGs identified in this study. For clarification, the above explanation has also been incorporated into the figure legend (**Lines 1067–1069**).

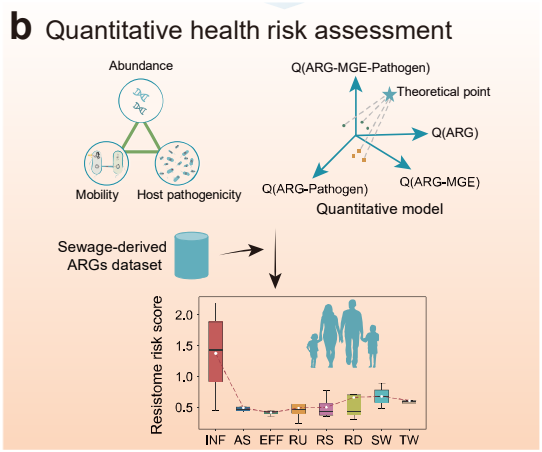


Fig. 7b | Risk assessment integrating the abundance, mobility, and host pathogenicity of sewage-derived ARGs. The top panel shows the theoretical framework, arrows indicate empirical data inputs, and the bottom panel presents the risk assessment outcomes.

Responses to the Comments of Reviewers

Dear Reviewers,

Thank you very much for your thoughtful and constructive comments on our manuscript entitled **"Ecological Connectivity of Anthropogenic Antimicrobial Resistance in Urban Water Systems and Its Implications for Drinking Water Risk Assessment"** (Manuscript Number: **NCOMMS-25-80871A**). We have carefully considered all the comments and further revised the manuscript accordingly. The changes made in this second revision are marked in red in the revised manuscript, and our detailed point-by-point responses are provided below.

Response to Reviewers' Comments

Reviewer #1:

Thank you to the authors for their thorough and detailed response to my review. I feel they have sufficiently addressed my concerns.

> **Response:** We sincerely thank ***Reviewer #1*** for the positive evaluation of our revised manuscript and for acknowledging that our responses have sufficiently addressed all previous concerns. We greatly appreciate the time and effort invested in reviewing our work and for constructive feedback, which has been instrumental in improving the quality of the manuscript.

Reviewer #1's comments regarding the previous comments of Reviewer #3:

1. The reviewer correctly points out that reduced cell density (eg in tap water vs source water) can likely influence the "risk" interpretations. The authors again estimate the absolute abundance of the pathogens in drinking water and find low concentrations (~3 cells per 100L) overall. The relevant response:
"Absolute quantification of pathogens in drinking water. In response to the reviewer's Comment #9, in addition to metagenomic relative quantification, we further estimated the absolute concentrations as suggested. Four bacterial species were detected, including Acinetobacter johnsonii, Escherichia coli, Mycobacterium phocaicum, and Pseudomonas aeruginosa, with at least one pathogenic species present in each sample. The average total concentration of these pathogens in tap water was 2.67 cells/100 L, and all carried diverse ARGs. Notably, P. aeruginosa, which harbored resistance genes to five antibiotic classes, was found in all samples."

The major finding of the study regards the ecological connectivity between the different compartments, so I think the read mapping here is worthy of some scrutiny. Namely- what are you actually detecting in drinking water?

At such low concentrations, and with the high detection limit of metagenomics, it is quite possible that reads aligning to related bacterial species are being erroneously mapped to the MAGs constructed.

First, the detection of E. coli in drinking water is concerning and a potential indication that spurious read mapping is occurring.

Second, there is no reason to suspect that the same P. aeruginosa strain(s) represented by the MAG would (1) be found in all compartments or (2) retain the same ARGs throughout. It is not generally possible from the approach taken here (read mapping to 95% ANI dereplicated MAGs) to differentiate populations of PA or closely related species in the different compartments.

For example, premise plumbing or distribution system biofilms could plausibly harbor stable and phenotypically distinct populations of PA. Therefore, they do not really prove connectivity through this method at least for drinking water.

In sum, it would be good to have additional details regarding:

1) The read mapping strategy, especially whether read mapping was performed by aligning short reads to a combined index (so that reads are aligned against the pathogens AND alternative references which might be better matches- see item 2 in this link: https://instrain.readthedocs.io/en/latest/important_concepts.html)

2) The breadth of coverage for each pathogen genome in the drinking water samples.

3) If the authors would like to support the notion that the MDR PA is found in drinking water, some indication that the ARGs and/or relevant MGEs are present in that compartment.

Last, it would be good to be explicit about the impact of absolute abundance on the conclusions. I noticed the abstract didn't change much, despite that much of the framing has theoretically shifted given the drastic decrease in absolute abundances of ARGs.

Response: We thank the reviewer for this insightful and constructive comment. We fully agree that read mapping specificity in low-biomass metagenomic datasets represents a key methodological challenge, particularly when interpreting low-abundance pathogen-associated signals in drinking water, and therefore requires careful evaluation. To address the reviewer's concerns, we have performed additional rigorous analyses as summarized below.

1) Re-evaluation of read mapping using competitive alignment (inStrain)

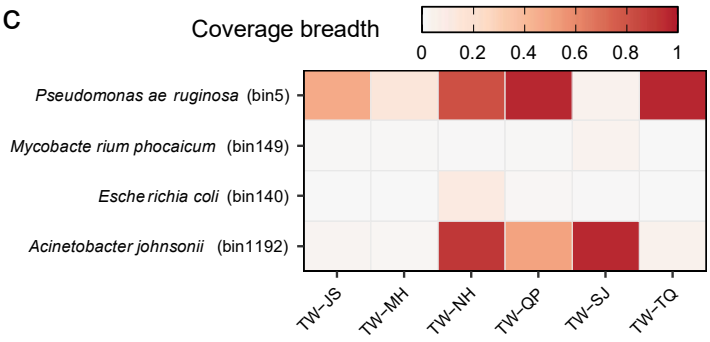
Following the reviewer's suggestion, we re-evaluated genome-wide recruitment patterns of MAGs (including pathogen-associated MAGs) in drinking water samples using inStrain (v1.10.0)[1]. All metagenomic reads were competitively aligned against a combined reference database consisting of all non-redundant MAGs recovered in this study, minimizing cross-mapping artifacts and false-positive signals. Detailed procedures are provided in **Supplementary Method 8**. On this basis, we further assessed coverage breadth, the number of scaffolds with mapped reads, population ANI (popANI), and read identity to systematically evaluate recruitment robustness under low-abundance conditions.

Supplementary Method 8 | Competitive genome-wide read recruitment to MAGs from tap water metagenomes. "To assess the robustness of low-abundance pathogen signals in tap water samples, we performed competitive genome-wide read recruitment to MAGs using Bowtie2 alignment paired with downstream inStrain (v1.10.0)[19] profiling. All non-redundant MAGs recovered in this study were concatenated to generate a unified reference database (non-redundant MAG set), with standardized renaming and bin annotation of all contigs and scaffolds to ensure unique genome identifiers. Clean paired-end reads from drinking water samples were mapped to the combined MAGs dataset using Bowtie2[20] (v2.5.1, end-to-end mode), enabling competitive alignment of each read across all reference genomes and reducing single-reference mapping bias. SAM files were converted to sorted BAM format using SAMtools (v1.17)[21]. For key pathogen-associated MAGs, including *Pseudomonas aeruginosa* (bin5), *Mycobacterium phocaicum* (bin149), *Escherichia coli* (bin140), and *Acinetobacter johnsonii* (bin1192), mapped reads were extracted and analyzed using inStrain profile to calculate genome-wide coverage, coverage breadth, and nucleotide diversity."

Reference:
[1] Olm MR, Crits-Christoph A, Bouma-Gregson K, Firek BA, Morowitz MJ, Banfield JF. inStrain profiles population microdiversity from metagenomic data and sensitively detects shared microbial strains. *Nat Biotechnol* 39, 727-736 (2021).

2) Genome coverage breadth of pathogen-associated MAGs in tap water

Re-analysis of genome coverage breadth showed that for *Pseudomonas aeruginosa* (bin5) and *Acinetobacter johnsonii* (bin1192), mapped reads in tap water samples were broadly distributed across the genome rather than restricted to conserved regions (**Supplementary Fig. 18c**). Specifically, *P. aeruginosa* exhibited a genome breadth of ~96.5% across multiple samples with average read identity of 96.8%–98.8%. These results support that the mapped reads reflect a genome-wide recruitment pattern consistent with the presence of *P. aeruginosa*, rather than stochastic or localized misalignment. In contrast, for low-abundance pathogens such as *E. coli*, the detected signals are near the detection limit of shotgun metagenomics in low-biomass systems and should therefore be interpreted with caution.



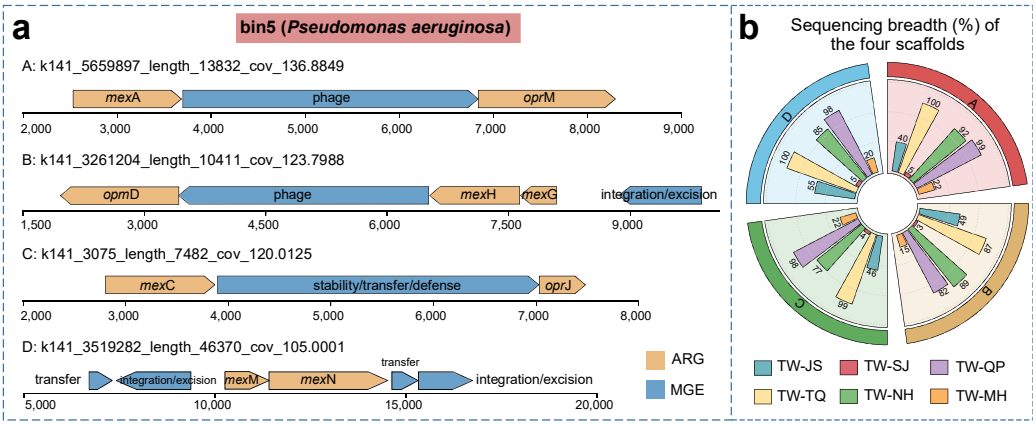
Supplementary Fig. 18c Heatmap showing the coverage breadth of pathogen-associated MAGs across six tap water samples, determined by competitive genome-wide mapping using inStrain.

3) Evidence for the presence of ARGs and MGEs in drinking water

We thank the reviewer for this important suggestion. To directly address the reviewer's request — to provide evidence that ARGs and MGEs associated with the *P. aeruginosa* lineage (bin5) are present in the drinking water compartment — we performed two complementary analyses focusing on tap water metagenomic reads across six samples, as shown in the newly added **Supplementary Fig. 13**.

(1) Assembly-level: Multiple bin5-associated scaffolds assembled from tap water metagenomes carried colocalized ARGs and MGEs. Specifically, these included components of multidrug efflux gene clusters (such as *mex*A-*opr*M, *mex*C-*opr*J, and *opm*D-*mex*H-*mex*G) colocalized with integrative elements, transposases, and phage-related sequences. **(2) Read recruitment to ARG/MGE-carrying scaffolds:** Metagenomic reads from six drinking water samples mapped back to these ARG/MGE-carrying scaffolds, with coverage breadth reaching 85%–100% in multiple samples (e.g., TW-TQ, TW-NH, TW-QP), confirming that the genomic context is present. In addition, independent read-level annotation using the SARG-OAP pipeline corroborated the assembly-level ARG profiles identified in bin5-associated scaffolds within the same drinking water samples, providing assembly-independent validation for the authentic presence of these resistance sequences in drinking water metagenomes.

Together, these results provide direct metagenomic evidence that ARGs and MGEs associated with the *P. aeruginosa* lineage (bin5) are indeed present in the drinking water compartment, directly addressing the reviewer's concern. Corresponding revisions have been incorporated into the main text (**Lines 286–287**).



Supplementary Fig. 13 | ARG–MGE co-localization and scaffold recruitment patterns of *Pseudomonas aeruginosa* (bin5) across tap water samples. a, Genomic context and environmental distribution of ARG-MGE co-localization in *Pseudomonas aeruginosa* (bin5). Schematic representation of four representative scaffolds (A-D) showing the physical linkage between ARGs (light orange) and MGEs (blue). **b**, Polar bar plot illustrating the sequencing breadth (%) of these scaffolds across six distinct sampling sites.

Lines 286–287: "Genomic context analysis revealed co-localization of ARGs and MGEs on pathogen-associated scaffolds (e.g., bin5), indicating their genetic linkage (Supplementary Fig. 13)."

4) Clarification of the impact of low absolute abundances

We agree that the low absolute abundances of ARGs and pathogens in drinking water should be explicitly considered when interpreting the ecological and public health implications. Accordingly, we have revised the ***Abstract*** to better contextualize the conclusions and avoid overinterpretation of the drinking water risk signal (**Lines 29–31**).

Lines 29–31: "wastewater discharge promotes AMR accumulation in natural water bodies, facilitating its persistence in finished drinking water, albeit with order-of-magnitude reductions in absolute abundance after disinfection."

We thank the reviewer again for encouraging us to critically re-evaluate these methodological and interpretive issues, which has significantly strengthened the manuscript.

Reviewer #2:

The revised manuscript has improved substantially and addresses many of the concerns raised in the previous round, which this reviewer appreciates. While most points have been clarified, two issues still require attention because they are central to the manuscript's main claims and interpretation.

Response: We sincerely thank the reviewer for the encouraging comments and for acknowledging the substantial improvements in our revised manuscript. We are gratified that most of the previous concerns have been clarified to the reviewer's satisfaction. We also appreciate the reviewer identifying the two remaining issues that are central to the manuscript's main claims and interpretation. We fully agree that addressing these points is critical for the scientific rigor and clarity of our study. We have carefully considered these comments and have revised the manuscript accordingly.

1. *First, the claim that the different water systems are hydrologically connected, despite lacking physical connectivity, needs clearer justification. To support this, the authors conducted SourceTracker analyses and examined shared ARG subtypes, reporting*

substantial resistome overlap. However, this evidence alone does not necessarily demonstrate hydrologic connectivity or transfer between systems. For example, river water and reservoir source water are not physically connected, so shared ARGs could reflect widespread background occurrence, common upstream sources, or similar selective pressures rather than direct exchange. Because this point is fundamental to the study, the authors should strengthen and clarify the inference and explicitly distinguish resistome similarity from physical transfer.

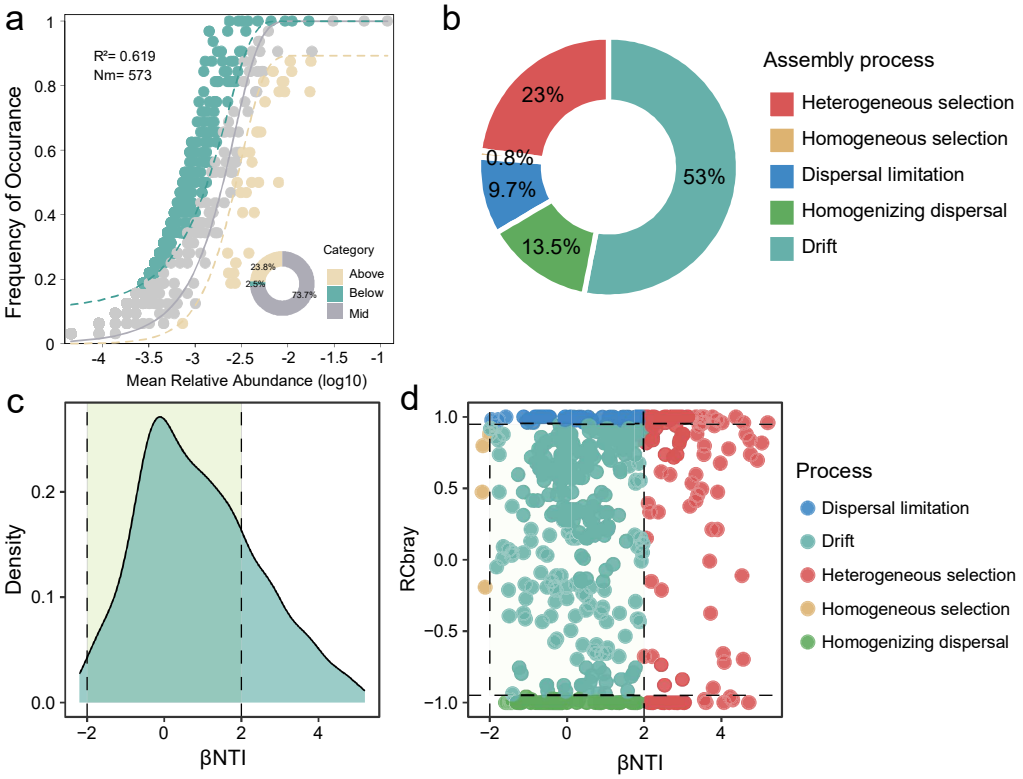
Response: We thank the reviewer for this important comment regarding the interpretation of connectivity among different water systems.

Addressing the distinction between similarity and physical transfer. As noted by the reviewer, our previous three-level analyses (SourceTracker, shared ARG subtypes, and sequence homology) reflected ecological connectivity/similarity rather than direct physical hydrological connectivity. Therefore, to improve conceptual precision, we have uniformly revised "hydrological connectivity" to "ecological connectivity" throughout the manuscript (**Lines 239**, **423**). These multi-level analyses provided robust evidence for compositional similarity across water compartments (Supplementary Fig. 7), but they did not directly address the reviewer's alternative explanation that the observed similarity could arise from shared environmental selective pressures.

Testing the alternative explanation of shared selective pressures. To further evaluate whether shared selective pressures could explain the observed similarity, we applied an ecological community assembly framework, including the neutral community model (NCM)[1] and null model analyses based on βNTI and RCbray[2] (Supplementary Fig. 19). The rationale is that if shared environmental selection were the dominant force, community assembly patterns would be expected to be primarily governed by deterministic processes. Conversely, dominance of stochastic processes (e.g., ecological drift and dispersal-related dynamics) would suggest that the observed similarity cannot be explained solely by shared environmental conditions[3,4]. Our results showed that stochastic processes dominated ARG host community assembly (76.2%), with ecological drift as the major contributor (53.0%), while deterministic selection played a minor role (23.8%). In addition, the NCM showed a relatively high goodness-of-fit for the resistome ($R^2 = 0.6185$), indicating that neutral processes substantially contribute to ARG distribution patterns across compartments.

These ecological assembly results indicate that the observed compositional similarity across water compartments is not primarily driven by shared environmental selection, but rather reflects the influence of stochastic ecological

processes, providing a potential ecological explanation for the observed similarity. Detailed methodological descriptions have been added in **Supplementary Method 9**, and the corresponding results have been incorporated into **Supplementary Fig. 19**.



Supplementary Fig. 19 | Ecological assembly mechanisms of ARGs and their host communities. a, Neutral community model (NCM) illustrating the relationship between mean relative abundance and detection frequency of ARG subtypes. Dashed lines indicate the predicted occurrence and 95% confidence intervals; points are colored based on their fit to the model (Above, Below, or Mid). The inset donut chart summarizes the proportions of these three categories. **b**, Relative contributions of deterministic and stochastic ecological processes inferred from null model analysis. The donut chart quantifies the influence of heterogeneous selection, homogeneous selection, dispersal limitation, homogenizing dispersal, and drift. **c**, Density distribution of βNTI values for ARG-hosting microbial communities. Dashed lines indicate the βNTI thresholds (−2 and 2) used to distinguish deterministic from stochastic assembly processes. **d**, Ecological assembly processes inferred from the integrated βNTI and RCbray framework. Colors denote different assembly mechanisms, including heterogeneous selection, homogeneous selection, dispersal limitation, homogenizing dispersal, and drift.

Supplementary Method 9 | Ecological assembly processes of ARGs and ARG host communities

"To characterize the ecological assembly mechanisms of ARG communities and their host microbial genomes, we integrated the Neutral Community Model (NCM)[22] and a phylogeny- and abundance-based null model framework[23]. First, the NCM was used to estimate the contribution of stochastic processes (e.g., dispersal and ecological drift) to ARG community assembly, with the goodness-of-fit (R^2) representing the relative importance of neutral

processes. Second, for ARG host communities, the beta nearest taxon index (βNTI) was calculated to assess phylogenetic turnover and infer deterministic selection, where $|\beta NTI| > 2$ indicates heterogeneous or homogeneous selection. For comparisons with $|\beta NTI| \leq 2$, the Bray–Curtis-based Raup–Crick metric (RCbray) was further applied to partition stochastic processes, including dispersal limitation, homogenizing dispersal, and ecological drift. Together, this framework enabled quantitative estimation of the relative contributions of deterministic and stochastic processes to community assembly."

Reference:

[1] Sloan WT, Lunn M, Woodcock S, Head IM, Nee S, Curtis TP. Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environ Microbiol* 8, 732-740 (2006).

[2] Stegen JC, Lin X, Fredrickson JK, Chen X, Kennedy DW, Murray CJ, Rockhold ML, Konopka A. Quantifying community assembly processes and identifying features that impose them. The *ISME Journal* 7, 2069-2079 (2013).

[3] Zhang H, Wang F, Yu S, Mao D, Wang X, Luo Y. Holocene chronosequences reveal the temporal dynamics of antibiotic resistance genes during sediment development. *Environ Sci Technol* 59, 18808-18818 (2025).

[4] Freire-Zapata V, Holland-Moritz H, Cronin DR, Aroney S, Smith DA, Wilson RM, et al. Microbiome-metabolite linkages drive greenhouse gas dynamics over a permafrost thaw gradient. *Nat Microbiol* 9, 2892-2908 (2024).

2. *Second, additional clarification is needed regarding the fate of ARGs in open systems such as rivers and reservoirs. Although the authors considered multiple factors in the Mantel tests and the PLS-SEM model, key confounders related to loss processes appear to be missing. In addition to host proliferation, it is important to consider bacterial mortality in the water column and decay of extracellular DNA or ARG signals due to natural processes (e.g., dilution, sunlight exposure, temperature, predation, settling). These processes should be incorporated where feasible or, at minimum, discussed explicitly as limitations. Without addressing them, the interpretation of ARG persistence and transport in open systems remains incomplete.*

Response: We thank the reviewer for raising this important point. We fully agree that natural attenuation processes (e.g., dilution, sunlight exposure, temperature, predation, settling) are critical for understanding ARG fate in open systems.

Clarification of already-included factors. We note that ***temperature*** — one of the factors mentioned by the reviewer — was already included in our Mantel tests and PLS-SEM model as part of the "Quality parameters" variable, which also included pH, DO, COD, TP, and NH_3-N. The model showed high explanatory power for ARG variation (Fig. 6).

Additional analysis to address the factor "*dilution*". To directly address the reviewer's concern about "*dilution*", we performed two complementary analyses

using sequencing-based microbial cell-equivalent concentration (MCEC; cell eq./mL) as a proxy for total microbial biomass along the spatial gradient from river to reservoir. **First**, we examined the variation trends of ARG abundance and MCEC across these sites. We observed lower levels of both ARG abundance and MCEC in source waters (SW) compared with river samples (RS/RD). Specifically, the average MCEC decreased from 2.11×10^5 cell eq./mL in river samples to 4.79×10^4 cell eq./mL in source waters, representing an approximately 4.4-fold reduction (**Supplementary Table 5**). **Second**, we further incorporated MCEC into the Mantel tests as a proxy for biomass attenuation (**Fig. 6a**). The results showed that both ARG composition and ARG host composition were significantly associated with MCEC ($p < 0.05$), suggesting that reductions in microbial biomass are significantly coupled with resistome attenuation. In addition, SEM analysis (**Fig. 6c**) showed that MCEC had a significant total effect (direct + indirect) on the resistome, with a standardized coefficient of 0.29 ($p < 0.05$). Detailed methodological descriptions of MCEC calculation and trend analysis have been provided in the ***Methods*** (**Lines 618–621**), and the relevant results have been integrated into the ***Discussion*** (**Lines 483–487**).

Supplemented acknowledgment of other loss processes and future perspectives. We acknowledge that other processes (UV degradation, predation, and settling) cannot be quantitatively partitioned with snapshot sampling data. We have explicitly discussed this limitation in the ***Discussion*** (**Lines 553–558**). Future studies integrating hydrodynamic modeling, UV photolysis kinetics, and microbial predation dynamics would help further resolve the contributions of these processes to ARG attenuation in open aquatic systems.

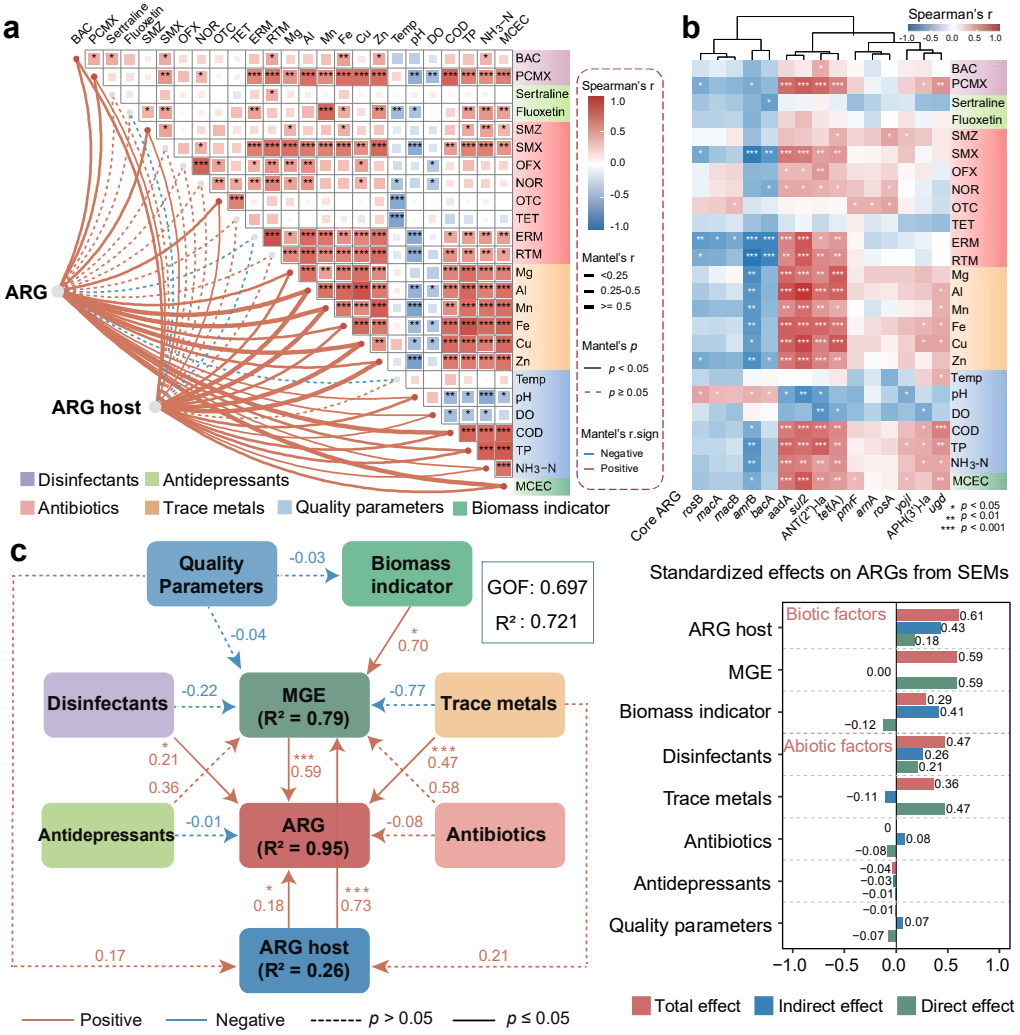


Fig. 6 | Key drivers influencing the dynamics of resistome within the urban water system. **a**, Mantel test diagrams illustrating relationships between various abiotic factors, ARGs, and bacterial hosts. Spearman correlation coefficients are represented by heatmap colors, with asterisks indicating statistical significance. Edge widths are proportional to Mantel's r statistic for the respective distance correlations, and edge color denotes statistical significance. *SMX, sulfamethoxazole; SMZ, sulfamethazine; ERM, erythromycin; RTM, roxithromycin; NOR, norfloxacin; OFX, ofloxacin; TET, tetracycline; OTC, oxytetracycline; BAC, benzalkonium chloride; PCMX, parachlorometaxylenol. MCEC, microbial cell-equivalent concentration.* **b**, Heatmap of Spearman correlations between core ARGs and abiotic factors (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$). **c**, Partial least squares structural equation modeling (PLS-SEM) quantifying direct and indirect drivers of ARG dynamics. Path colors indicate effect direction (yellow: positive; blue: negative), with R^2 values representing the proportion of variance explained for each dependent variable. Bar plots show standardized total, indirect, and direct effects of biotic/abiotic factors.

Lines 483–487 (Discussion): "In addition, microbial biomass decreased along the spatial gradient from river to source water and was consistently associated with both ARG abundance and composition, suggesting that biomass reduction, potentially driven by dilution, may act in concert with other biotic and abiotic processes to influence resistome dynamics in UWS."

Lines 553–558 (Discussion): "Although the PLS-SEM model in this study exhibited high explanatory power for ARG variation, several natural attenuation processes in open aquatic systems (e.g., sunlight exposure, predation, and settling) were not quantitatively characterized due to the snapshot sampling design. Future studies integrating UV photolysis kinetics, microbial predation dynamics, and particle-associated transport would help further resolve the contributions of these processes to ARG fate in open aquatic systems."

Lines 618–621 (Methods): "In addition, sequencing-derived microbial cell-equivalent concentration (MCEC) was estimated based on ARGs-OAP single-copy marker gene normalization, DNA extraction yield, and filtered water volume, and expressed as microbial cell equivalents per milliliter (cell eq./mL)."

3. **Reviewer #2's comments regarding the previous comments of Reviewer #3***:*

1) Overall, the authors have largely addressed Reviewer #3's comments and revised the manuscript accordingly.
2) With respect to Comment 14 (filtration method), I believe the authors should provide additional clarification and, where possible, further justification. Reviewer #3's central concern relates to whether the sampled microbial community is representative of the water column (i.e., source water) or whether the approach may bias the sample toward organisms associated with filters.
Although I agree with the authors that substantive microbial growth on the filters during sampling is unlikely under typical conditions, the reported 24-hour sampling period still raises two related questions:
a) Potential contribution from distribution system biofilms/pipe-associated communities: Over a prolonged sampling duration, the collected material may reflect not only microorganisms suspended in the water source but also organisms intermittently shed from pipes or other infrastructure, which could affect representativeness.
b) Comparability across sampling locations and study inferences about connectivity: If other matrices in the study (e.g., wastewater treatment plant effluent, river water, source water) were collected as grab samples or over much shorter time windows, the markedly different temporal integration (24-hour composite vs. grab/short-duration sampling) could introduce a mismatch in sampling timescales. Given that the manuscript emphasizes ecological connectivity across compartments, the authors should more explicitly discuss how differences in sampling duration may influence comparability and interpretation of connectivity signals.

In my view, the response and the revised manuscript would be strengthened by a clearer explanation of: (i) the rationale for the 24-hour filtration duration; (ii) the extent to which this approach is expected to capture water-column communities versus potential distribution-system contributions; and (iii) how differences in temporal integration across sample types are addressed analytically or, at minimum, explicitly acknowledged

as a limitation.

Response: We thank the reviewer for the constructive comments and for acknowledging that we have substantially addressed ***Reviewer #3***'s previous concerns. We agree that, in low-biomass drinking water systems, the use of long-duration large-volume filtration warrants additional clarification regarding sample representativeness and comparability across different environmental matrices. In response, we have expanded the relevant discussion in the revised manuscript and explicitly clarified the potential limitations of this sampling strategy.

(1) Rationale for the 24-hour filtration duration. The primary purpose of the 24-hour continuous large-volume filtration was to obtain sufficient microbial biomass from the extremely low-biomass drinking water environment to meet the DNA input requirements for downstream metagenomic sequencing and genome-resolved analyses. Conventional short-duration or small-volume sampling often fails to yield sufficient DNA for robust metagenomic analyses, particularly for the detection of low-abundance microorganisms and ARGs. Previous studies have suggested that large-volume filtration is often necessary to improve microbial recovery efficiency and metagenomic detection sensitivity in low-biomass drinking water systems[1, 2]. We have clarified this rationale in the revised ***Methods*** section (**Lines 606–610**):

> **Lines 606–610**: "Tap water samples (TW, n = 6) were collected from six administrative districts across the city, whose microorganisms were collected by filtering over 2,000 L using a domestic reverse osmosis (R.O.) purification device (A.O. Smith, USA) operated at a constant flow rate of 1.5 L·min^{-1} for approximately 24 h to ensure sufficient biomass for metagenomic and genome-resolved analyses."

(2) Water-column communities versus potential distribution-system contributions. We agree that tap water samples integrate contributions from both bulk water communities and distribution-system biofilms, and we cannot definitively partition these sources. This integrated signal represents the overall microbial and ARG burden to which consumers are exposed at the point of use. The detection of upstream-derived ARG variants with high sequence identity across compartments (Supplementary Fig. 7c) provides independent evidence for ecological connectivity. Future studies using biofilm-specific tracers or controlled pipe-loop experiments could help quantitatively partition bulk water versus biofilm contributions in tap water resistome dynamics. As suggested, we have supplemented this concern and perspective in the revised ***Discussion*** (**Lines 456–461**):

> **Lines 456–461**: "Tap water samples may integrate contributions from both bulk water communities and biofilms, which represents the overall microbial and ARG burden to which

consumers are exposed at the point of use. The detection of upstream-derived ARG variants across compartments supports ecological connectivity, and future tracer or pipe-loop studies could help quantitatively resolve biofilm versus bulk water contributions."

(3) Comparability of temporal integration across sample types. We agree with the reviewer that differences in sampling time scales may affect direct comparability across environmental matrices. In this study, tap water samples were collected using a 24-hour time-integrated continuous filtration strategy, while samples from other compartments were collected in biological triplicates within a 24-hour window and subsequently pooled into composite samples (Lines 605–606). All samples were collected during a period of stable system operation without major rainfall, operational fluctuations, or treatment disturbances (Line 596). Therefore, all sample types share a comparable temporal integration window, supporting the validity of cross-comparison. Although differences in temporal integration are not expected to fundamentally alter the overall association patterns of microbial communities and resistomes across compartments, they may influence short-term community variability and the detection of some low-abundance microorganisms and ARGs. As suggested, we have added a clarifying statement at the end of the "Sample collection design" subsection in ***Methods*** (**Line 611–612**) and a statement in the ***Discussion*** noting this consideration (**Lines 551–553**).

Line 611–612 (Methods): "Thus, all samples were collected within a comparable 24-hour temporal window."

Lines 551–553 (Discussion): "Nevertheless, differences in temporal integration among sample types may influence short-term community variability and the detection of some low-abundance microorganisms and ARGs."

Reference:

[1] Pinto AJ, Schroeder J, Lunn M, Sloan W, Raskin L. Spatial-temporal survey and occupancy-abundance modeling to predict bacterial community dynamics in the drinking water microbiome. *Mbio* 5, e01135-01114 (2014).

[2] Hozalski RM, Zhao X, Kim T, LaPara TM. On-site filtration of large sample volumes improves the detection of opportunistic pathogens in drinking water distribution systems. *Appl Environ Microbiol* 90, e0165823 (2024).

We thank the reviewer for the thorough and constructive comments, which have significantly improved the clarity and rigor of our manuscript.

Responses to the Comments of Reviewers

Dear Reviewers
,
We are delighted and honored to receive your letter informing us that our manuscript (Manuscript Number: **NCOMMS-25-80871B**) has been accepted in principle for publication in *Nature Communications*. We sincerely thank the Reviewers for their constructive insights, which have greatly improved the quality and rigor of our st udy.

Response to Reviewers' Comments

Reviewer #1:

I thank the authors for their exemplary response to my comments and feel they have addressed my concerns sufficiently.

> **Response:** We thank the reviewer for the positive evaluation of our revised manuscript and for recognizing our efforts in addressing the previous comments. We greatly appreciate the reviewer's constructive feedback, which has helped improve the quality and clarity of the manuscript.

Reviewer #2:

I appreciate the authors' efforts in revising the manuscript to address the comments raised.

> **Response:** We thank the reviewer for the positive assessment and for acknowledging the improvements made in the revised manuscript. We are grateful for the valuable comments, which have strengthened the overall rigor of our work.