nature portfolio

Peer Review File

Citywide metagenomics reveals microbial and resistome transmission in urban wastewater

Corresponding Author: Professor Tianmu Chen

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)
The manuscript entitled “Undiscovered crosstalk network of microbiome and resistome in an urban wastewater system” employs shotgun metagenomic sequencing to investigate ESKAPE pathogens and mobile genetic elements (MGEs) in an urban wastewater system. The study presents a short time series, with samples collected between February and April of a single year, and applies a wide range of bioinformatics and statistical tools to support its analyses.

I have two major comments regarding the study design and interpretation of the results.

Major comments

Although approximately 300 samples were collected across multiple locations, several key conclusions are based on a limited temporal window. Seasonal variability is well documented as a major driver of microbial community composition and resistome dynamics in wastewater systems, yet this factor is not adequately considered or discussed. The short sampling period raises concerns about the generalizability of the findings, and the limitations imposed by the lack of seasonal coverage should be explicitly acknowledged and discussed.

Line 178 refers to the upregulation and downregulation of antibiotic resistance genes (ARGs), which implies that transcriptomic analyses were performed. However, no transcriptomic study is described in the Methods or Results sections. The basis for these statements is therefore unclear. Although Supplemental Figure 14 presents a volcano plot, there is no accompanying methodological description explaining how gene expression or differential abundance analyses were conducted. This analysis appears to be missing or insufficiently described in the manuscript and requires clarification.

Minor comments

Line 112: The current wording, together with the cited reference, implies that the authors conducted the work described in reference 18 (“Federal Funding for the Study of Antimicrobial Resistance in Nosocomial Pathogens: No ESKAPE”). This should be clarified to avoid misattribution.

The year of sample collection is not stated when first describing the sampling period (February–April). The reader only learns later in the manuscript that samples were collected in 2023. This information should be clearly stated upfront.

Line 478: Composite samplers were used; however, the sampling duration is not specified. Please clarify whether composites were collected over 12 hours, 24 hours, or another interval.

While this cannot be addressed retroactively, it is worth noting that short-read sequencing technologies are increasingly being complemented or replaced by long-read platforms such as Oxford Nanopore and PacBio. Long-read sequencing can provide more informative assemblies and improved resolution for network and MGE analyses. A brief discussion acknowledging this limitation would strengthen the manuscript.

(Remarks on code availability)
I did not try the code, but all codes related to this study are placed in the github link.

Reviewer #2

(Remarks to the Author)
What are the noteworthy results?
It shows factors that can likely predict patterns of resistance and sources. Suggests that local transmission is impacting the microbial communities and their associated resistance markers. Makes a strong case for local/granular surveillance, rather than a uniform approach. This is in line with what is known from other surveillance studies. One size will not fit all, and environmental factors could be altering/impacting the signal, along with selection pressure. Some papers have mentioned mutagens that could be impacting and selecting for specific strains/lineages - this paper calls for analysis of the water milieu with respect to antimicrobials and disinfectants, along with mutagens. Lines 183-197 give a nice description of the major findings.
I really liked Figure 6. I'd use it in my classes. I'd like to understand the role of pH, as well as TSS. I think it would be interesting to look at an even more, fine-scale, analysis of particular hospitals/plants as they seem to share specific connections - I'd also like to better understand the nature of connections to the plants (residential homes, dental offices, clinics), more diverse populations/larger populations. I wish the authors had done more temporal/long-term sampling. It would be really interesting to see if air travel introduced novel genes to the environment that became more prevalent over time, but that might be too difficult to monitor.

Will the work be of significance to the field and related fields? How does it compare to the established literature?
There are a few papers that describe air travel contributions to a local wastewater resistome. Most have been on a much grander scale. This paper shows the potential for monitoring community spread and the increase of particular genetic elements associated with resistance. It also shows the potential for long-term monitoring of household and hospital resistance markers in the context of air travel and the introduction of new elements. I think I would have liked more of an analysis of particular plants and sites, as there seem to be more patterns where some of the sources are sharing more of the genes, and it may be due to other factors.

If the work is not original, please provide relevant references.
NA

Does the work support the conclusions and claims, or is additional evidence needed?
I think so. I don't have major concerns. There are plenty of suppositions and conclusions that may change over time with more sampling and more temporal samples. It may be that strain-specific monitoring could be done long-term here. Lines 346-350 speak to some of the challenges. I think more needs to be done to truly conclude what the authors are saying, but this would be another paper.

Are there any flaws in the data analysis, interpretation, and conclusions? Do these prohibit publication or require revision?
None that I can detect. It is a specific example of an approach in a particular geographical region at a fine level of scale. I think the approaches here that identify particular mobile elements are of great use.

Is the methodology sound? Does the work meet the expected standards in your field?
Yes

Is there enough detail provided in the methods for the work to be reproduced?
Yes = the methods were clear and detailed

Typo on line 330 ^39
line 50 "A" high spatial

(Remarks on code availability)
Code is not my area of expertise

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The authors have addressed the majority of my previous comments. I only have one observation. The raw data seems not be available as stated here:

"Raw sequencing data have been deposited in the China National GeneBank (https://db.cngb.org/) under accession number PRJCA039605"

Raw Sequencing data cannot be found in that website.

I even went here and no success:

https://ngdc.cncb.ac.cn

I did search for the Bioproject and Biosample in Supplementary Table 12 and they can be found in NCBI.

(Remarks on code availability)
I checked the alpha and beta resistome and ESKAPE code in R and they worked.

Reviewer #2

(Remarks to the Author)
Thank you for your thorough and thoughtful responses to my comments. I think everything was addressed, and I appreciate the additional analyses and figures. The manuscript is much improved.

(Remarks on code availability)
I have no more comments

RESPONSE TO REVIEWERS' COMMENTS for the paper "Undiscovered crosstalk network of microbiome and resistome in an urban wastewater system**"**

Dear Reviewers:

Thank you for your thorough evaluation and constructive feedback concerning our manuscript entitled “Undiscovered crosstalk network of microbiome and resistome in an urban wastewater system” (*NCOMMS-25-87105*). We have thoroughly revised our manuscript, addressed all the comments, and additionally conducted new analyses to further strengthen the robustness of our work. We appreciate the positive comments highlighting the contribution of our work. Specific revisions and responses to each comment are provided in detail below.

Reviewer #1:

General comment: The manuscript entitled “Undiscovered crosstalk network of microbiome and resistome in an urban wastewater system” employs shotgun metagenomic sequencing to investigate ESKAPE pathogens and mobile genetic elements (MGEs) in an urban wastewater system. The study presents a short time series, with samples collected between February and April of a single year, and applies a wide range of bioinformatics and statistical tools to support its analyses.

Response: We sincerely thank you for a very careful and detailed review of our manuscript and for your overall positive impression of our work at scale. We also thank you for providing particularly constructive comments.

R1Q1: I have two major comments regarding the study design and interpretation of the results. Major comments: Although approximately 300 samples were collected across multiple locations, several key conclusions are based on a limited temporal window. Seasonal variability is well documented as a major driver of microbial community composition and resistome dynamics in wastewater systems, yet this factor is not adequately considered or discussed.

Response: We thank you for highlighting this critical point. We fully agree that seasonal variability is a well-documented driver of microbial community composition and resistome dynamics in wastewater systems[1-3]. While we acknowledge the study covers a limited temporal window, the sampling period (February to April) was strategically selected to ensure meteorological stability. This period in Xiamen represents early spring, characterized by moderate temperatures and minimal precipitation. By strictly avoiding the typhoon season

and heavy rainfall periods (typically May to October), we aimed to minimize the confounding effects of stormwater dilution and surface runoff. Therefore, our study design prioritized a high-resolution spatial analysis under stable environmental conditions. This allows us to attribute observed variations more confidently to facility-specific activities. Importantly, we would like to clarify that long-term temporal monitoring is already underway to build upon these findings. This work serves as an essential baseline for future longitudinal programs and provides a high-granularity snapshot of the intensity and interconnectedness of ARGs transmission within urban wastewater networks. We have updated the Discussion to explicitly clarify the rationale behind this temporal selection and its advantages for establishing a robust urban baseline, to read:

L457-L470: "First, we utilized intensive sampling to establish a high-resolution spatial baseline of the urban microbiome and resistome. While our sampling timeframe (February–April) was strategically designed to capture a period of meteorological stability and avoid the confounding effects of heavy seasonal rainfall, we acknowledge that this cross-sectional approach cannot fully resolve seasonal dynamics or temporal directionality in community change. Seasonal variability is a well-documented driver of microbial community composition and resistome dynamics in wastewater systems[52,53], yet this factor was not fully captured within our study period. Similarly, capturing the long-term ecological integration of novel ARGs introduced via international air travel remains inherently challenging within a restricted timeframe. Although our three-month window revealed a compelling pattern of high-frequency occurrence and broad spatial distribution for key externally introduced determinants, suggesting their capacity for rapid local stabilization, their extended evolutionary trajectories are yet to be fully elucidated."

R1Q2: The short sampling period raises concerns about the generalizability of the findings, and the limitations imposed by the lack of seasonal coverage should be explicitly acknowledged and discussed.

Response: We thank you for highlighting the potential limitations associated with the short sampling period and the lack of seasonal coverage. We have revised the Discussion to explicitly acknowledge this constraint and to clarify both the rationale underlying our study design and the future directions for investigating long-term transmission dynamics across seasons. While we fully recognize that longitudinal studies are essential for capturing the full ecological trajectory and seasonal variability of aquatic microbial communities[4], the primary objective of this study was to establish a high-resolution spatial baseline of the urban microbiome and resistome. To achieve this, we conducted an intensive sampling campaign within a meteorologically stable window, which helped minimize the confounding influence of short-term precipitation events and temperature fluctuations. This approach allowed us to more clearly resolve spatial heterogeneity and potential transmission linkages across urban

environmental compartments. Consequently, the resulting spatial "snapshot" provides important public health insights by identifying potential transmission nodes and anthropogenic reservoirs within the city. Second, although seasonal processes may influence the magnitude or frequency of microbial dissemination, the existence of transmission linkages themselves is unlikely to be entirely mediated by seasonal variation. In other words, seasonal dynamics and spatial transmission processes are not mutually exclusive. Therefore, identifying these spatial relationships within a controlled sampling window represents an important step toward understanding the structure of urban microbial transmission networks. Finally, following your suggestion, we have clarified in the revised manuscript that this study serves as a foundational framework for our ongoing longitudinal monitoring efforts. These future investigations aim to integrate the high-density spatial observations presented here with long-term temporal data in order to better characterize seasonal dynamics and their influence on urban microbiome and resistome transmission.

L457-L470: "First, we utilized intensive sampling to establish a high-resolution spatial baseline of the urban microbiome and resistome. While our sampling timeframe (February–April) was strategically designed to capture a period of meteorological stability and avoid the confounding effects of heavy seasonal rainfall, we acknowledge that this cross-sectional approach cannot fully resolve seasonal dynamics or temporal directionality in community change. Seasonal variability is a well-documented driver of microbial community composition and resistome dynamics in wastewater systems[52,53], yet this factor was not fully captured within our study period. Similarly, capturing the long-term ecological integration of novel ARGs introduced via international air travel remains inherently challenging within a restricted timeframe. Although our three-month window revealed a compelling pattern of high-frequency occurrence and broad spatial distribution for key externally introduced determinants, suggesting their capacity for rapid local stabilization, their extended evolutionary trajectories are yet to be fully elucidated."

R1Q3: Line 178 refers to the upregulation and downregulation of antibiotic resistance genes (ARGs), which implies that transcriptomic analyses were performed. However, no transcriptomic study is described in the Methods or Results sections. The basis for these statements is therefore unclear. Although Supplemental Figure 14 presents a volcano plot, there is no accompanying methodological description explaining how gene expression or differential abundance analyses were conducted. This analysis appears to be missing or insufficiently described in the manuscript and requires clarification.

Response: We sincerely appreciate your valuable feedback. We acknowledge that the use of terms such as "upregulation" and "downregulation" in Line 178 could lead to confusion, as these terms are typically associated with transcriptomic analyses. We apologize for the incorrect terminology used in this context, as no transcriptomic study was performed

in our research. Instead, the terms were intended to describe changes in the relative abundance of these genes across different environments. We have revised the manuscript to reflect this more accurate description:

L181-L184: "Compared with residential wastewater, FLI showed a substantial increase in the abundance of 435 ARGs, with tetracycline, erythromycin, and macrolides determinants being the most prominent, whereas only 29 genes showed a significant decrease in abundance."

L186-L188: "In contrast, HOS exhibited an increase in the abundance of 368 ARGs and a decrease in the abundance of only 8 **(Supplementary Fig. 15B)**."

Additionally, we have now added a more detailed methodological description in the Methods section, which outlines how the ARG abundance data were processed, analyzed, and classified according to the risk assessment framework[5], to read:

L603-L609: "To evaluate the potential health risks associated with the identified ARGs, we applied a risk classification approach in which each ARG was matched with the risk level of that specific gene as proposed in established frameworks for assessing antimicrobial resistance[28]. This allowed us to assign risk categories to each gene based on its abundance, mobility, and potential for horizontal gene transfer. We further used RES samples as a baseline to benchmark differences in risk levels between external inputs and locally high-risk areas."

R1Q4: Line 112: The current wording, together with the cited reference, implies that the authors conducted the work described in reference 18 ("Federal Funding for the Study of Antimicrobial Resistance in Nosocomial Pathogens: No ESKAPE"). This should be clarified to avoid misattribution.

Response: We sincerely appreciate your careful observation and important suggestion to avoid misattribution. We acknowledge that the original wording at this line may have created an unintended implication that our study conducted the work described in that reference, which was not our intention. We have accordingly revised the relevant description:

L112: "We specifically surveyed the presence of ESKAPE pathogens in wastewater."

R1Q5: The year of sample collection is not stated when first describing the sampling period (February–April). The reader only learns later in the manuscript that samples were collected in 2023. This information should be clearly stated upfront.

Response: We thank you for your careful observation and valuable suggestion regarding the sampling year information. We acknowledge the oversight that the sampling year (2023) was not specified when the sampling period was first described. The issue has been fully

rectified in the revised manuscript:

L73-78: "Between February and April 2023, over 300 samples were collected from 50 sites across all six administrative districts: Tong'an (TA), Jimei (JM), Haicang (HC), Xiang'an (XA), Huli (HL), and Siming (SM), comprising 27 residential communities (RES), 7 hospitals (HOS), and 16 WTP sampling sites."

R1Q6: Line 478: Composite samplers were used; however, the sampling duration is not specified. Please clarify whether composites were collected over 12 hours, 24 hours, or another interval.

Response: We thank you for pointing out this oversight regarding the unspecified composite sampling duration in Line 478. This detail has been explicitly clarified in the revised manuscript at this position, where we state that the composite wastewater samples were collected over a 24-hour interval using sterile composite samplers. The corrected sentences now read:

L528-L530: "For each sampling event, 24-hour composite sampling was conducted, and four 500mL wastewater samples were collected using sterile composite samplers and immediately transported to the laboratory under 4°C storage conditions."

R1Q7: While this cannot be addressed retroactively, it is worth noting that short-read sequencing technologies are increasingly being complemented or replaced by long-read platforms such as Oxford Nanopore and PacBio. Long-read sequencing can provide more informative assemblies and improved resolution for network and MGE analyses. A brief discussion acknowledging this limitation would strengthen the manuscript.

Response: We appreciate your insightful comment regarding the use of short-read sequencing technologies. While our study utilized short-read sequencing, which is widely employed in metagenomic analyses, we recognize the inherent limitations of this approach, particularly in resolving complex genomic structures. We agree with your observation that long-read sequencing platforms, such as Oxford Nanopore and PacBio, offer more accurate genome assemblies, facilitate better strain-level resolution, and provide a clearer understanding of the dynamics of horizontal gene transfer[6,7]. We have now included a brief discussion of this limitation in the revised manuscript, to read:

L477-L482: "Third, while short-read sequencing provided valuable insights, it has inherent limitations in resolving complex genomic structures. Future studies using long-read technologies such as Oxford Nanopore or PacBio could offer improved assembly resolution for recovering complete microbial genomes, identifying strain-level information, and facilitating better characterization of ARGs and MGEs."

Reviewer #2:

General comment: What are the noteworthy results?

It shows factors that can likely predict patterns of resistance and sources. Suggests that local transmission is impacting the microbial communities and their associated resistance markers. Makes a strong case for local/granular surveillance, rather than a uniform approach. This is in line with what is known from other surveillance studies. One size will not fit all, and environmental factors could be altering/impacting the signal, along with selection pressure. Some papers have mentioned mutagens that could be impacting and selecting for specific strains/lineages - this paper calls for analysis of the water milieu with respect to antimicrobials and disinfectants, along with mutagens. Lines 183-197 give a nice description of the major findings.

Response: We sincerely appreciate your insightful and comprehensive summary of our study's noteworthy results, as well as the constructive suggestions for further analysis of the aquatic environment.

R2Q1: I really liked Figure 6. I'd use it in my classes. I'd like to understand the role of pH, as well as TSS.

Response: We are gratified that you found Figure 6 to be a valuable pedagogical tool. In response to your inquiry regarding the roles of pH and TSS, we have expanded our discussion to further elucidate their respective influences and contributions.

L433-L439: "Furthermore, our results highlight TSS as a critical structural determinant. The observed inhibitory effect of TSS beyond the 400 mg/L threshold suggests that excessive particulate matter may restrict niche breadth through mass-transfer limitations or the competitive dominance of a few generalist taxa. Such particulate-induced ecological filtering could potentially fragment the horizontal transmission pathways of ARGs by sequestering free DNA or limiting the physical contact required for conjugation."

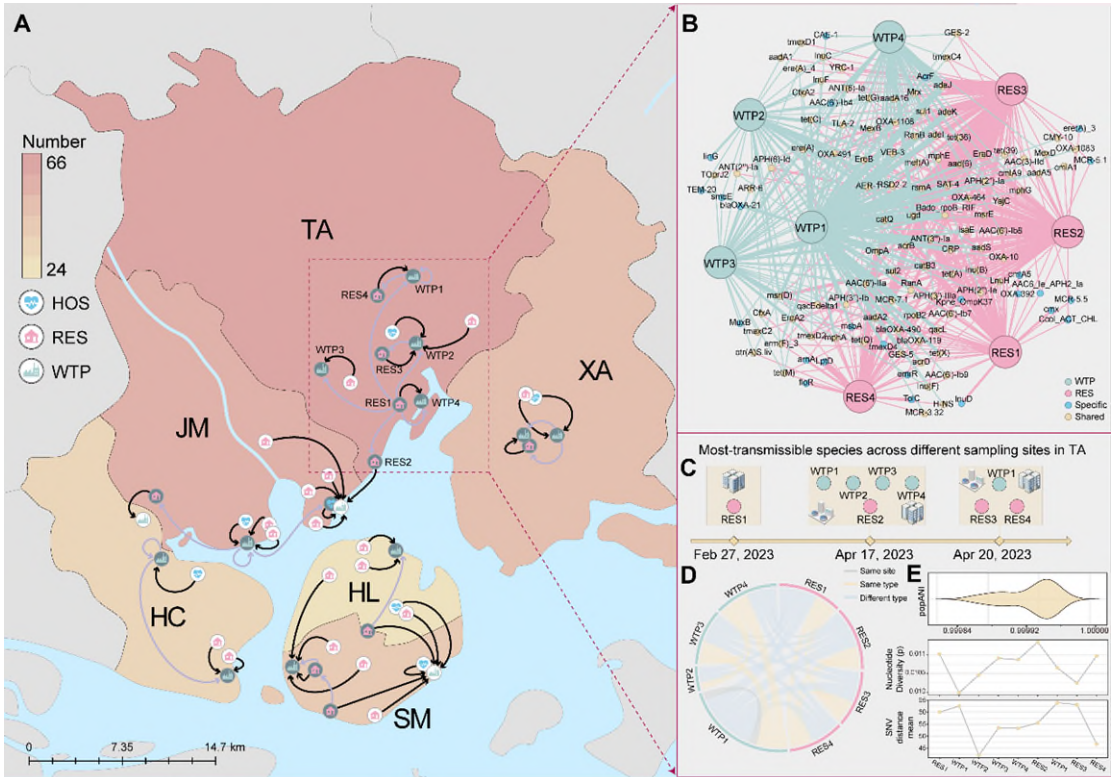
R2Q2: I think it would be interesting to look at an even more, fine-scale, analysis of particular hospitals/plants as they seem to share specific connections.

Response: We sincerely appreciate your suggestion to further explore fine-scale connections among specific facilities. In response, we conducted a more detailed investigation of site-level transmission dynamics, focusing on the most prevalent and highly transmissible species identified in our dataset, *Aliarcobacter cryaerophilus*, within the TA district. From a resistomic perspective, we analyzed the exchange of ARGs across all sampling sites in the TA district, categorizing them by their occurrence in different types. The resulting co-occurrence network reveals that a vast majority of ARGs are extensively shared across the two types, providing robust evidence of persistent genetic connectivity within the urban water infrastructure. Furthermore, our microbial-sharing analysis focused on the spatiotemporal

dynamics of the most prevalent species, *A. cryaerophilus*. Temporal tracking identified RES1 as the initial detection site in February 2023, followed by frequent transmission events across the district. The genomic integrity of these transmission events is confirmed by popANI exceeding 99.984%, demonstrating that these transmission events happened across genomic near-identical clones. We further explored the intra-population genetic signatures of this clonal expansion. We observed a reduction in nucleotide diversity at WTP1, marking the lowest diversity point in the district. This may suggest a potent selective sweep as the population transitioned from residential sources into the treatment plant, where intense environmental filters purged the standing genetic variation. Together, these fine-scale analyses provide a more detailed picture of how specific facilities may function as connected nodes within the urban water network. Rather than acting solely as passive conduits, elements of the infrastructure may actively shape microbial evolution by facilitating clonal dissemination and imposing strong selective pressures.

L324-L344: "To elucidate the finer dynamics underlying this cross-environment purifying selection, we performed a fine-scale, intra-district investigation centered on *Aliarcobacter cryaerophilus*, the most prevalent and highly transmissible species in our dataset. Focusing specifically on the interconnected infrastructure within the TA district, we mapped the resistome and microbiome trajectories across particular RES and WTP nodes **(Supplementary Fig. 17A)**. Resistomic profiling revealed a highly integrated co-occurrence network, with a vast majority of ARGs extensively shared across RES and WTP sites, underscoring persistent genetic connectivity within the local urban water infrastructure **(Supplementary Fig. 17B)**. Furthermore, spatiotemporal tracking of *A. cryaerophilus* captured a discrete clonal expansion event. Following initial detection at a residential source in February 2023, near-identical clones (popANI > 99.984%) rapidly disseminated to downstream specific sites across the district. Crucially, examining the intra-population genetic signatures during this transmission revealed a precipitous drop in nucleotide diversity at WTP1, marking the lowest diversity point within the local network **(Supplementary Fig. 17C-E)**. This reduction highlights how localized infrastructural filters can act as strict evolutionary bottlenecks. Together, these fine-scale findings corroborate our broader macro-ecological observations, demonstrating that the urban water system does not merely transport pathogens, but may actively reshape their evolutionary template through discrete, high-pressure selective events at specific infrastructural nodes."

L447-L449: "Our fine-scale tracking of specific intra-district transmissions provides direct mechanistic validation of this dynamic, suggesting that localized infrastructural filters likely impose strong genetic bottlenecks."



Supplementary Figure 17. (For review) Spatiotemporal resistome and microbiome connectivity within the TA district. A. Map illustrating sampling locations with color gradients corresponding to site-specific sampling intensity. All sites where *Aliarcobacter cryaerophilus* was detected are shaded in dark green. Black arrows symbolize directional connectivity between sampling nodes, while purple arrows represent the temporal sequence of transmission events of Arcobacter cryaerophilus within the same district. **B.** The co-occurrence network illustrates the exchange of ARGs across different sewage matrices, with color coding distinguishing shared and type-specific ARGs. **C.** Temporal timeline identifies the chronological order of detection at all involved sites. **D.** The chord diagram quantifies the frequency of transmission events occurring among these specific locations. **E.** The genomic characteristics of the prevalent strains and the intra-population genetic variation.

R2Q3: I'd also like to better understand the nature of connections to the plants (residential homes, dental offices, clinics), more diverse populations/larger populations.

Response: We thank you for pointing out that the scale of population in different connections can influence the frequency of microbial sharing events related to wastewater treatment plants. To address this, we calculated the "Scaled population size" for each connection, defined as the log-transformed product of the populations of the two connected nodes using the formula:

$$Scaled\ population\ size = \log_{10}(Population1 \times Population2)$$

Our analysis demonstrates that population scale was not a significant predictor of genome

sharing events across any of the connection types, as evidenced by the lack of statistically significant correlations for HOSWTP, RESWTP, and WTPWTP (**Figure R1**). These findings indicate that the number of sharing events is not driven by the volume of the contributing populations, suggesting that the intrinsic characteristics or specific environmental pressures of each connection type may play a more critical role in shaping these microbial dynamics.

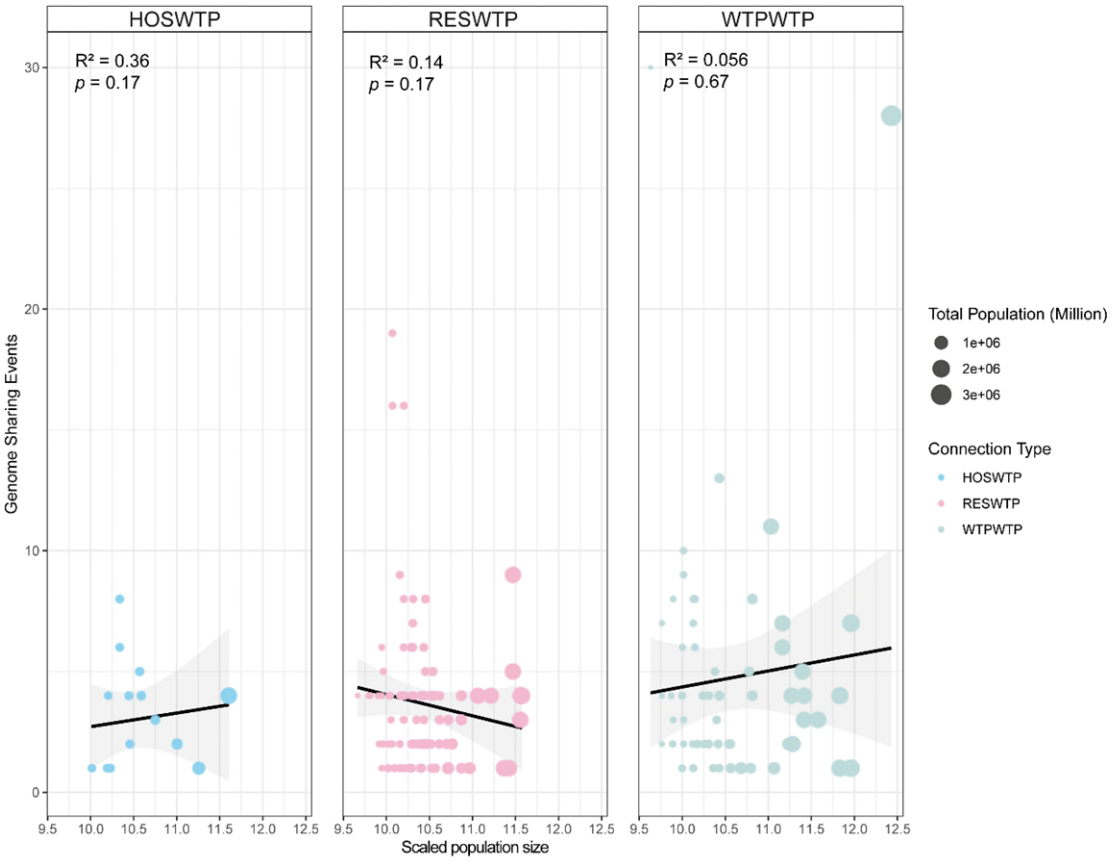


Figure R1. (For review) Correlation between the scaled population size and the frequency of genome sharing events. Scatter plots illustrate the relationship across three connection types: hospital-to-WTP (HOSWTP), residential-to-WTP (RESWTP), and WTP-to-WTP (WTPWTP). The x-axis represents the population potential, and bubble sizes indicate the total population involved in the respective connection. Solid black lines represent linear regression fits, with shaded regions denoting 95% confidence intervals. The R^2 and p-values are displayed for each panel.

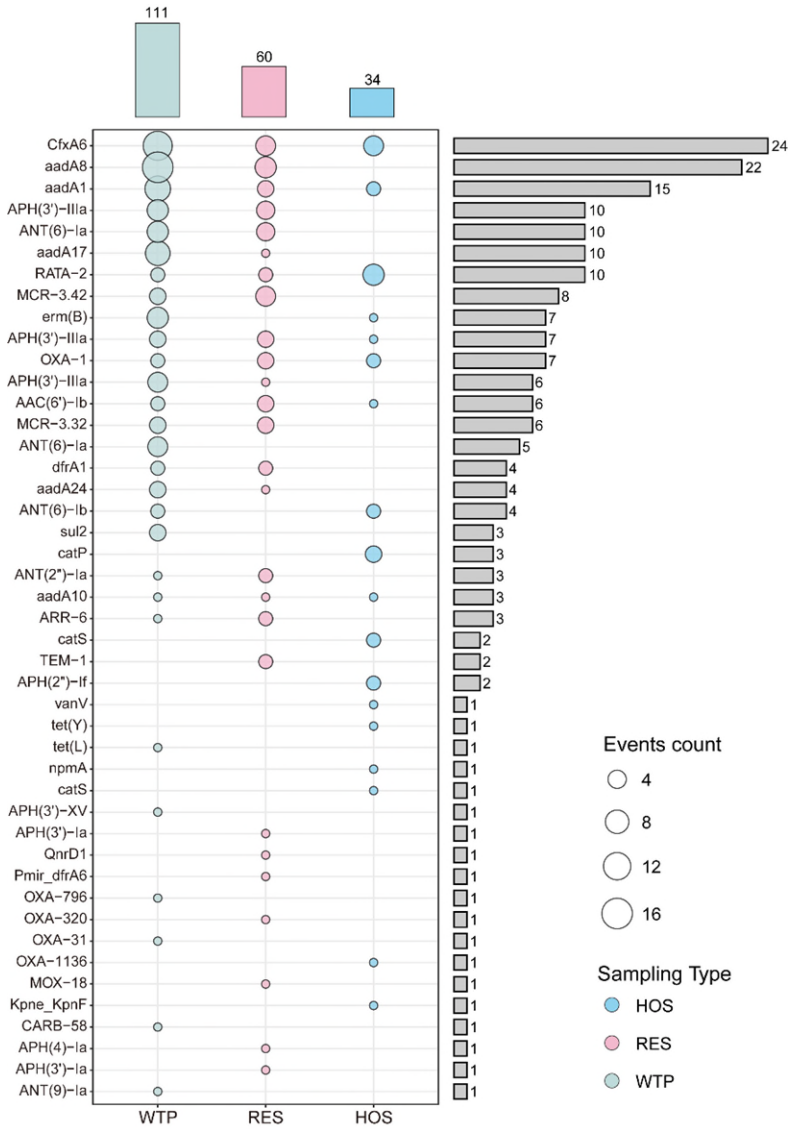
R2Q4: I wish the authors had done more temporal/long-term sampling. It would be really interesting to see if air travel introduced novel genes to the environment that became more prevalent over time, but that might be too difficult to monitor.

Response: We sincerely appreciate your insightful suggestion regarding the importance of temporal monitoring to track the potential introduction of novel resistance genes via air travel. While we agree that such processes are inherently complex and, as you noted, can be

difficult to monitor, it is noteworthy that even within our study's three-month sampling window, we observed a compelling pattern of high-frequency occurrence and wide spatial distribution for several key resistance genes. As illustrated in the newly added bubble plot (**Supplementary Figure 10**), genes such as *CfxA6*, *aadA8*, and *aadA1* exhibited significant event counts in clinical, residential areas, and wastewater treatment plants. This high prevalence across different environmental compartments suggests that these genes may have already established a stable presence in the local environment following their introduction. However, we fully acknowledge that the robustness and long-term trajectory of this phenomenon require more extended observation to be definitively confirmed. To this end, we would like to clarify that a more comprehensive long-term monitoring project is already underway to further validate these preliminary findings and track the evolutionary dynamics of these genes over a multi-year period. We have also enriched the discussion in the revised manuscript to address these temporal considerations and the inherent challenges of environmental monitoring.

L149-L152: Notably, throughout the three-month sampling window, specific resistance determinants (including *CfxA6*, *aadA8*, and *aadA1*) exhibited high-frequency occurrence and broad spatial distribution across clinical, residential, and municipal wastewater compartments (**Supplementary Fig. 10**).

L457-L470: "First, we utilized intensive sampling to establish a high-resolution spatial baseline of the urban microbiome and resistome. While our sampling timeframe (February–April) was strategically designed to capture a period of meteorological stability and avoid the confounding effects of heavy seasonal rainfall, we acknowledge that this cross-sectional approach cannot resolve seasonal dynamics or temporal directionality in community change. Seasonal variability is a well-documented driver of microbial community composition and resistome dynamics in wastewater systems[52,53], yet this factor was not fully captured within our study period. Similarly, capturing the long-term ecological integration of novel ARGs introduced via international air travel remains inherently challenging within a restricted timeframe. Although our three-month window revealed a compelling pattern of high-frequency occurrence and broad spatial distribution for key externally introduced determinants, suggesting their capacity for rapid local stabilization, their extended evolutionary trajectories are yet to be fully elucidated."



Supplementary Figure 10. (For review) Distribution and prevalence of antibiotic resistance genes (ARGs) across different sampling environments. The bubble plot (left) illustrates the occurrence and frequency of specific ARGs detected across WTP, RES, and HOS. The size of each bubble corresponds to the event count, representing the frequency of a specific gene's detection within each sampling type. The corresponding bar chart (right) quantifies the total cumulative detection frequency for each ARG, with numerical values indicating the absolute count. Color coding distinguishes the sampling types.

R2Q5: Will the work be of significance to the field and related fields? How does it compare to the established literature?

There are a few papers that describe air travel contributions to a local wastewater resistome. Most have been on a much grander scale. This paper shows the potential for monitoring community spread and the increase of particular genetic elements associated with resistance. It also shows the potential for long-term monitoring of household and hospital resistance

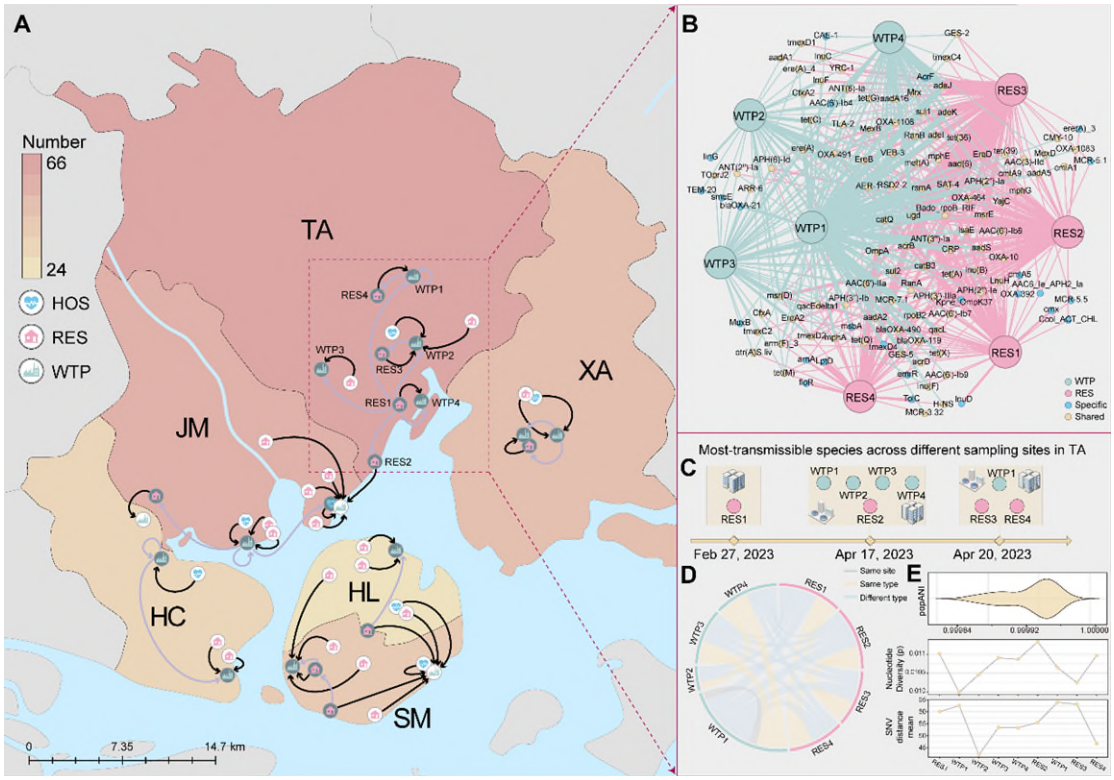
markers in the context of air travel and the introduction of new elements. I think I would have liked more of an analysis of particular plants and sites, as there seem to be more patterns where some of the sources are sharing more of the genes, and it may be due to other factors.

Response: We sincerely thank you for your positive assessment of our study's significance and for precisely recognizing its unique value in monitoring local community spread and introducing new resistance elements, especially when compared to the grander scale of existing literature. Encouraged by this validation of our research scope and objectives, we fully agree with your insightful suggestion to delve deeper into the analysis of particular wastewater treatment plants. In the revised manuscript, we have incorporated a more granular, fine-scale investigation centered on the most prevalent species identified in our dataset, *A. cryaerophilus*, focusing on the resistome and microbiome dynamics specifically within the TA district. From a resistomic perspective, we analyzed the exchange of ARGs across all sampling sites, categorizing them by their occurrence in different types. The resulting co-occurrence network reveals that a considerable number of ARGs are shared across these types, suggesting potential genetic connectivity within the urban water infrastructure. Furthermore, our microbiome analysis explored the spatiotemporal patterns of *A. cryaerophilus*. Temporal tracking identified RES1 as the initial detection site in February 2023, followed by observed transmission events across the district. The genomic similarity of these events is supported by popANI values exceeding 99.984%, indicating the likely spread of near-identical clones rather than completely distinct strains. We further explored the intra-population genetic signatures of this expansion and observed a reduction in nucleotide diversity at WTP1, marking the lowest diversity point in the district. This localized reduction could potentially reflect a selective sweep as the population transitioned from residential sources into the treatment plant, where environmental pressures might influence the standing genetic variation. Overall, these fine-scale observations highlight the possibility that specific nodes in the urban water system may play a role in shaping the evolutionary trajectories of such pathogens, though extended monitoring would be required to fully confirm these dynamics.

L324-L344: "To elucidate the finer dynamics underlying this cross-environment purifying selection, we performed a fine-scale, intra-district investigation centered on *Aliarcobacter cryaerophilus*, the most prevalent and highly transmissible species in our dataset. Focusing specifically on the interconnected infrastructure within the TA district, we mapped the resistome and microbiome trajectories across particular RES and WTP nodes **(Supplementary Fig. 17A)**. Resistomic profiling revealed a highly integrated co-occurrence network, with a vast majority of ARGs extensively shared across RES and WTP sites, underscoring persistent genetic connectivity within the local urban water infrastructure **(Supplementary Fig. 17B)**. Furthermore, spatiotemporal tracking of *A. cryaerophilus*

captured a discrete clonal expansion event. Following initial detection at a residential source in February 2023, near-identical clones (popANI > 99.984%) rapidly disseminated to downstream specific sites across the district. Crucially, examining the intra-population genetic signatures during this transmission revealed a precipitous drop in nucleotide diversity at WTP1, marking the lowest diversity point within the local network **(Supplementary Fig. 17C-E)**. This reduction highlights how localized infrastructural filters can act as strict evolutionary bottlenecks. Together, these fine-scale findings corroborate our broader macro-ecological observations, demonstrating that the urban water system does not merely transport pathogens, but may actively reshape their evolutionary template through discrete, high-pressure selective events at specific infrastructural nodes."

L447-L449: "Our fine-scale tracking of specific intra-district transmissions provides direct mechanistic validation of this dynamic, suggesting that localized infrastructural filters likely impose strong genetic bottlenecks."



Supplementary Figure 17. (For review) Spatiotemporal resistome and microbiome connectivity within the TA district. A. Map illustrating sampling locations with color gradients corresponding to site-specific sampling intensity. All sites where *Aliarcobacter cryaerophilus* was detected are shaded in dark green. Black arrows symbolize directional connectivity between sampling nodes, while purple arrows represent the temporal sequence of transmission events of Arcobacter cryaerophilus within the same district. **B.** The co-occurrence network illustrates the exchange of ARGs across different sewage matrices, with color coding distinguishing shared and type-specific ARGs. **C.** Temporal timeline identifies

the chronological order of detection at all involved sites. **D.** The chord diagram quantifies the frequency of transmission events occurring among these specific locations. **E.** The genomic characteristics of the prevalent strains and the intra-population genetic variation.

R2Q6: Does the work support the conclusions and claims, or is additional evidence needed?

I think so. I don't have major concerns. There are plenty of suppositions and conclusions that may change over time with more sampling and more temporal samples. It may be that strain-specific monitoring could be done long-term here. Lines 346-350 speak to some of the challenges. I think more needs to be done to truly conclude what the authors are saying, but this would be another paper.

Response: We sincerely appreciate your positive assessment that our work adequately supports the presented conclusions and claims, and for the constructive perspectives on future research directions. We are grateful for the recognition that no major concerns exist with the current evidence, and we fully agree with the insightful comments that additional long-term monitoring would further refine the suppositions and conclusions presented here, and these important follow-up directions align closely with our own research plans. Specifically, the present study utilized an intensive sampling strategy over a short time period, with the primary aim of investigating the transmission and evolutionary dynamics of the microbiome and resistome in wastewater during a period of relatively stable temperature and precipitation in this coastal city. This work is intended to lay a solid foundation for subsequent long-term research, while also providing empirical support for environmental management practices and public health safeguards.

R2Q7: Are there any flaws in the data analysis, interpretation, and conclusions? Do these prohibit publication or require revision?

None that I can detect. It is a specific example of an approach in a particular geographical region at a fine level of scale. I think the approaches here that identify particular mobile elements are of great use.

Response: We sincerely appreciate your positive evaluation and insightful comments on our data analysis, interpretation, and conclusions, as well as the recognition of our work's strengths. We are grateful that no flaws were identified in these critical aspects of the study, and we are particularly encouraged by the your affirmation of the utility of our approaches for identifying specific mobile elements. Our study focuses on a fine-scale analysis of the approach in a specific geographical region, and we are pleased that this targeted research design and its findings are well-received. We will retain the core framework of our data analysis and conclusion derivation in the revised manuscript to ensure the consistency and rigor of the research.

R2Q8: Is the methodology sound? Does the work meet the expected standards in your field?

Yes

Response: We sincerely appreciate your positive assessment of the soundness of our methodology and the conformity of our work to the field's expected standards.

R2Q9: Is there enough detail provided in the methods for the work to be reproduced?

Yes = the methods were clear and detailed

Response: We appreciate your positive comment on the clarity and detail of our methods section.

R2Q10: Typo on line 330 ^39, line 50 "A" high spatial

Response: Thank you for your careful check and for identifying these typos. We have now revised the mentioned sentences. The corrected sentences now read:

L355-L357: "This design yielded a temporally narrow yet city-scale "snapshot" of microbiome and resistome features and enabled construction of an ecological network spanning aircraft, hospital, residential, and treatment compartments[43] **(Fig. 6)**."

L50-L52: "A high spatial resolution genomic-resolved approach is therefore essential to understand both ecological structures and evolutionary processes sustaining ARG persistence."

Reference

1 Fierer, N. *et al.* A Metagenomic Investigation of Spatial and Temporal Changes in Sewage Microbiomes across a University Campus. *mSystems* **7**, e0065122 (2022). https://doi.org/10.1128/msystems.00651-22

2 LaMartina, E. L., Mohaimani, A. A. & Newton, R. J. Urban wastewater bacterial communities assemble into seasonal steady states. *Microbiome* **9**, 116 (2021). https://doi.org/10.1186/s40168-021-01038-5

3 Becsei, A. *et al.* Time-series sewage metagenomics distinguishes seasonal, human-derived and environmental microbial communities potentially allowing source-attributed surveillance. *Nat Commun* **15**, 7551 (2024). https://doi.org/10.1038/s41467-024-51957-8

4 Zhou, Z. *et al.* Unravelling viral ecology and evolution over 20 years in a freshwater lake. *Nat Microbiol* **10**, 231–245 (2025). https://doi.org/10.1038/s41564-024-01876-7

5 Zhang, A. N. *et al.* An omics-based framework for assessing the health risk of antimicrobial resistance genes. *Nat Commun* **12**, 4765 (2021). https://doi.org/10.1038/s41467-021-25096-3

6 Eisenhofer, R. *et al.* A comparison of short-read, HiFi long-read, and hybrid strategies for genome-resolved metagenomics. *Microbiol Spectr* **12**, e0359023 (2024). https://doi.org/10.1128/spectrum.03590-23

7 Xia, Y. *et al.* Strategies and tools in illumina and nanopore-integrated metagenomic analysis of microbiome data. *Imeta* **2**, e72 (2023). https://doi.org/10.1002/imt2.72

RESPONSE TO REVIEWERS' COMMENTS for the paper "Citywide metagenomics reveals microbial and resistome transmission in urban wastewater."

Dear Reviewers:

We sincerely thank you for your continued evaluation of our manuscript entitled "Citywide metagenomics reveals microbial and resistome transmission in urban wastewater" (*NCOMMS-25-87105A*). We are deeply grateful for the positive final assessments and are delighted that the additional analyses and extensive revisions provided in the previous round have met the reviewers' expectations. For this current revision, we have corrected the minor discrepancy regarding the data repository link to ensure complete data accessibility. We would like to take this final opportunity to express our sincere gratitude to the reviewers for their time, expertise, and meticulous attention to detail. Your collective efforts have significantly strengthened the robustness of our study. Our specific response to the final observation is provided in detail below.

Reviewer #1:

R1Q1: The authors have addressed the majority of my previous comments. I only have one observation. The raw data seems not be available as stated here: "Raw sequencing data have been deposited in the China National GeneBank (https://db.cngb.org/) under accession number PRJCA039605". Raw Sequencing data cannot be found in that website. I even went here and no success: https://ngdc.cncb.ac.cn. I did search for the Bioproject and Biosample in Supplementary Table 12 and they can be found in NCBI. I checked the alpha and beta resistome and ESKAPE code in R and they worked.

Response: We sincerely thank you for your meticulous evaluation of our work, including the time and effort taken to verify the data repositories and successfully test our R scripts for the resistome and ESKAPE analyses. We apologize for the confusion regarding the raw sequencing data. We made an error in the original manuscript by incorrectly naming the repository as the China National GeneBank (CNGB). The raw sequencing data were actually deposited in the National Genomics Data Center (NGDC) at the China National Center for Bioinformation (CNCB). The dataset under accession number PRJCA039605 has now been fully released and is publicly accessible. The data can be directly accessed via the following link: [https://ngdc.cncb.ac.cn/bioproject/browse/PRJCA039605]. Furthermore, we have updated the "Data Availability" section in the revised manuscript to accurately reflect the correct database name and URL.

L750-L754: "Raw sequencing data have been deposited in the National Genomics Data

Center (https://ngdc.cncb.ac.cn/) under accession number PRJCA039605 [https://ngdc.cncb.ac.cn/bioproject/browse/PRJCA039605]. Details of the FASTQ files from the reference studies used for comparative analyses are provided in Supplementary Data 12. Source data are provided with this paper."

Finally, we would like to take this opportunity to express our deepest gratitude for your constructive and rigorous feedback throughout the entire peer review process. Your insightful comments and meticulous checks have significantly elevated the quality, transparency, and reproducibility of this study. We truly appreciate the time and expertise you have dedicated to evaluating our manuscript.

Reviewer #2:

General comment: Thank you for your thorough and thoughtful responses to my comments. I think everything was addressed, and I appreciate the additional analyses and figures. The manuscript is much improved. I have no more comments.

Response: We sincerely thank you for your positive feedback and for the time and effort dedicated to evaluating our manuscript. We are pleased that our revisions, along with the additional analyses and figures, have satisfactorily addressed your previous concerns. We deeply appreciate your constructive comments and guidance throughout the review process, which have significantly contributed to improving the overall quality, clarity, and rigor of our study.