

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | | |
|-------------------------------------|--|
| n/a | Confirmed |
| ☐ | ☒ The exact sample size (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| ☐ | ☒ A description of all covariates tested |
| ☐ | ☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☐ | ☒ For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted
<i>Give P values as exact values whenever suitable.</i> |
| ☒ | ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☐ | ☒ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☐ | ☒ Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i>), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	Paired-end sequencing (PE150) was performed on a DNBSEQ-T7 sequencer (BGI, Shenzhen, China) using the DNBSEQ-G50RS High-throughput Sequencing Set (BGI, Shenzhen, China).
Data analysis	All R scripts and python scripts used in this project are available at the following Github repository: https://github.com/XZYING147/Wastewater_project . Software used in the analysis of this study include fastp (v0.23.4), Bowtie2 (v2.4.5), KneadData (v0.10.0), MEGAHIT (v1.2.9), MetaBAT (v2.12.1), MaxBin2 (v2.2.6), CONCOCT (v1.0.0), MetaWRAP (v1.3.1), CheckM (v1.0.12), dRep (v3.4.2), CoverM (0.7.0), GTDB-Tk (v2.3.2), iTOL (v6.6), MetaPhlAn4 (v4.1.1), HUMAnN3 (v3.7), ARGs-OAP (v3.2.4), ResFinder (v4.6.0), inStrain (v1.9.0), and BLAST (v2.16.0).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

All shotgun sequencing data are all available under Bioproject ID PRJCA039605 at the platform of National Genomics Data Center (NGDC). The raw environmental parameters are available in Supplementary Table 12. Databases used in the analysis of this study included GRCh38.p13 (NCBI accession number GCA_000001405.28), GTDB genome database (database Release207), Chocophlan nucleotide database, Uniref90 protein database, SARG reference database (v2.2), CARD database (v3.3.0), ISfinder database, VFDB database.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender

Information was not used in the analysis.

Reporting on race, ethnicity, or other socially relevant groupings

Information was not used in the analysis.

Population characteristics

Information was not used in the analysis.

Recruitment

Information was not used in the analysis.

Ethics oversight

Information was not used in the analysis.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description

Here we conducted a city-wide comprehensive metagenomic snapshot (n=252) across a rapidly developing coastal metropolis, capturing the complete wastewater microbiota journey from hospitals, residential zones, and international transportation hubs to treatment facilities.

Research sample

We collected over 300 samples from 50 sampling sites spanning six administrative districts (i.e., TA, JM, HC, XA, HL, and SM). Our investigation focused on the major components of urban wastewater systems, comprising 27 residential communities (RES), 7 hospitals (HOS), and 16 wastewater treatment plants (WTP) sampling sites. Additionally, to compare the compositional differences in microbial communities and resistome between imported wastewater from external sources versus local wastewater, as well as to assess the impact of imported wastewater on the urban wastewater system's microbiome and resistome, we collected wastewater samples from international flights (FLI, N=22) arriving in Xiamen. No statistical methods were used to predetermine sample size. The sample size was determined to ensure a comprehensive spatiotemporal representation of Xiamen's wastewater system. We analyzed 16 wastewater treatment plants, 34 source-specific nodes (7 hospitals, 27 residential areas), and 22 international aircraft wastewater samples. This combination provides a representative and high-resolution capture of the city-wide and import-related microbial and antibiotic resistance gene landscapes.

Sampling strategy

Our sampling effort encompassed all wastewater treatment facilities in Xiamen to comprehensively capture a representative snapshot of the city's microbial community profile.

Data collection

HOS wastewater is systematically sampled at designated effluent discharge nodes within healthcare facilities, undergoing multistage disinfection protocols compliant with China's Medical Institution Water Pollutant Discharge Standards (GB 18466-2005) prior to integration into municipal sewage networks. RES wastewater is collected at primary pumping stations following community sewage aggregation, entering municipal networks untreated. WTP samples are obtained from untreated water in inlet collection channels (influent). FLI wastewater is collected through two protocols: (1) immediate sampling from international aircraft vacuum wastewater retainers post-landing, and (2) composite sampling of 2-3 flights consolidated in mobile decontamination units at airport aprons. Environmental parameters were automatically recorded by the integrated on-site monitoring sensors at each wastewater treatment plant. The historical data were directly exported from the facility's digital management systems by the investigators.

Timing and spatial scale	Detailed sampling time are provided in Supplementary Table 1. All samples were acquired in a three-month field campaign across 50 sampling sites spanning six administrative districts, including residential communities (RES, N=27), hospitals (HOS, N=7), and wastewater treatment plants (WTP, N=16). To evaluate microbial and antibiotic burden from external sources, we also collected wastewater from international flights arriving in Xiamen (FLI, N=22).
Data exclusions	A total of 72 samples yielded insufficient DNA concentrations (<0.001 ng/μL in 50μL sterile water), fell below Qubit detection thresholds, and were consequently excluded from sequencing library construction.
Reproducibility	All attempts to repeat the experiment were successful. Technical replicates were performed for a subset of samples during library preparation and sequencing, yielding highly consistent taxonomic and functional profiles. All samples were collected in at least duplicate to ensure the robustness of the observed microbial dynamics.
Randomization	Randomization was not applicable as samples were allocated into groups based on their predetermined functional source.
Blinding	Blinding was not relevant to this study because the investigation focused on an objective characterization of the wastewater microbiome and resistome across known geographical and functional categories (RES, HOS, WTP, FLI). There were no subjective treatments or clinical interventions that required blinding of investigators during sample collection or data analysis.
Did the study involve field work?	☒ Yes ☐ No

Field work, collection and transport

Field conditions	All environmental parameters are available in Supplementary Table 11.
Location	Field work was conducted across 50 sites in Xiamen, China (spanning approximately 24°23'N to 24°54'N and 117°53'E to 118°26'E). Detailed GPS coordinates and site descriptions for all residential, hospital, and wastewater treatment plants sampling nodes are provided in Supplementary Table 1.
Access & import/export	All samples and environmental parameters were collected, and the detailed information was available in Supplementary Table 1 and Supplementary Table 11.
Disturbance	No disturbance were caused by this study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks	N/A
Novel plant genotypes	N/A
Authentication	N/A