Explaining the resistomes in a megacity's water supply catchment: Roles of microbial assembly-dominant taxa, niched environments and pathogenic bacteria

Dong Wu[1,2,3], Jue Zhao[4], Yinglong Su[1,3], Mengjie Yang[1], Jan Dolfing[5], David W. Graham[6*], Kai Yang[1,3*], Bing Xie[1,3*]

[1] Shanghai Engineering Research Center of Biotransformation of Organic Solid Waste, Shanghai Key Lab for Urban Ecological Processes and Eco-Restoration, School of Ecological and Environmental Sciences, East China Normal University, Shanghai 200241, PR China

[2] Key Laboratory of Environmental Pollution Monitoring and Disease Control, Ministry of Education, Guizhou Medical University, Guizhou 550001, PR China

[3] Shanghai Institute of Pollution Control and Ecological Security, Shanghai 200092, PR China

[4] Department of Civil and Environmental Engineering and Research Institute for Sustainable Urban Development, The Hong Kong Polytechnic University, Kowloon, Hong Kong

[5] Faculty Energy and Environment, Northumbria University, Newcastle upon Tyne, NE1 8QH, UK

[6] School of Engineering, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK

*Corresponding authors:
Prof. Bing Xie; East China Normal University; Shanghai 200241, P.R. China. Email: bxie@des.ecnu.edu.cn; Phone: (+86) 021-54341276.

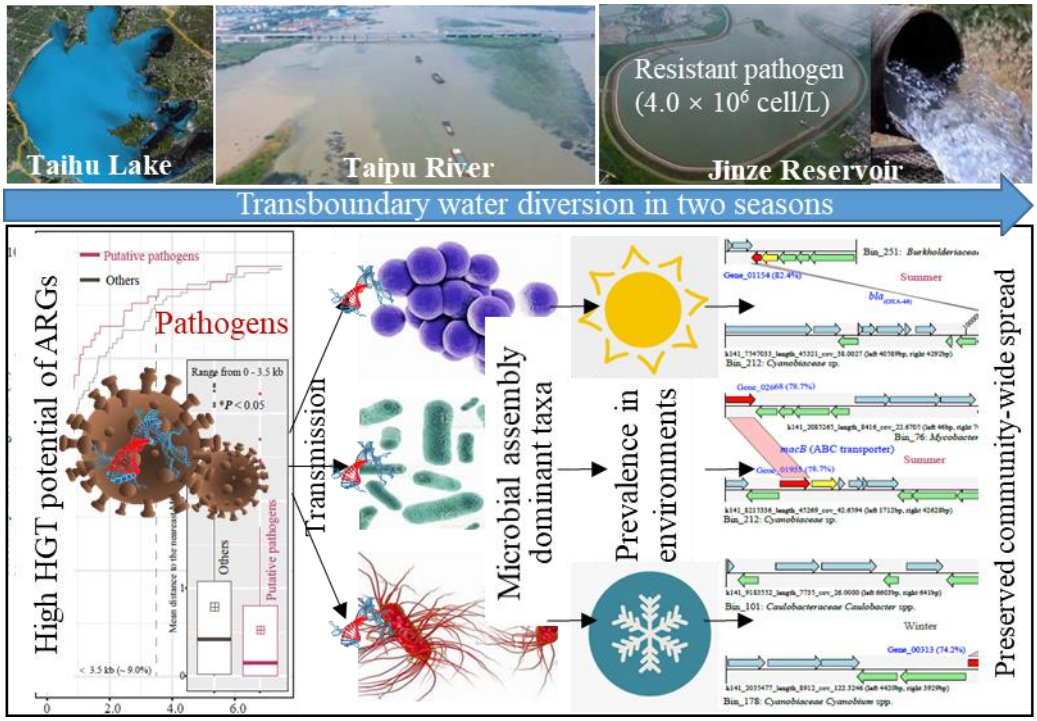
Prof. Kai Yang; East China Normal University; Shanghai 200241, P.R. China. Email: kyang@re.ecnu.edu.cn; Phone: (+86) 021-54341069

Prof. David W Graham, Newcastle University; Newcastle upon Tyne, United Kingdom. Email: david.graham@ncl.ac.uk; Phone: (+44) 0191-208-7925

Highlights

- Resistome size in the reservoir was comparable to that in sewers
- Microbial assembly process significantly influenced dynamics of resistome
- Putative resistant bacteria were at the same abundance across the catchment
- Putative resistant pathogens greatly contributed to the transmission of ARGs
- Intragenomic ARGs frequently transferred to microbial assembly-dominant taxa

Graphic Abstract



Abstract

Antibiotic resistance genes (ARGs) in drinking water sources suggest the possible presence of resistant microorganisms that jeopardize human health. However, explanations for the presence of specific ARGs in situ are largely unknown, especially how their prevalence is affected by local microbial ecology, taxa assembly and community-wide gene transfer. Here, we characterized resistomes and bacterial communities in the Taipu River catchment, which feeds a key drinking water reservoir to a global megacity, Shanghai. Overall, ARG abundances decreased significantly as the river flowed downstream towards the reservoir ($P < 0.01$), whereas the waterborne bacteria assembled deterministically ($|\beta\text{NRI}| > 2.0$) as a function of temperature and dissolved oxygen conditions with the assembly-dominant taxa (*e.g.* *Ilumatobacteraceae* and *Cyanobiaceae*) defining local resistomes ($P < 0.01$, Cohen's $D = 4.22$). Bacterial hosts of intragenomic ARGs stayed at the same level across the catchment (60 ~ 70 genome copies per million reads). Among them, the putative resistant pathogens (*e.g.* *Burkholderiaceae*) carried mixtures of ARGs that exhibited high transmission probability (transfer counts = 126, $P < 0.001$), especially with the microbial assembly-dominant taxa. These putative resistant pathogens had densities ranging form 3.0 to 4.0×10^6 cell/L, which was more pronouncedly affected by resistome and microbial assembly structures than environmental factors (SEM, std-coeff $\beta = 0.62$ *vs*. 0.12). This work shows that microbial assembly and resistant pathogens play predominant roles in prevelance and dissemination of resistomes in receiving water, which deserves greater attention in devisng control strategies for reducing in-situ ARGs and resistant strains in a catchment.

Keywords: *antibiotic resistance, community-wide gene transfer, putative resistant pathogen, metagenome binning, microbial assembly, water supply catchment*

Abbreviations

- Antibiotic resistance (AR)
- Aminoglycosides (Agy)
- Antibiotic resistance genes (ARGs)
- Average nucleotide identity (ANI)
- Bacitracin (Bac)
- Beta(β)-lactams (Bla)
- Chloramphenicol (Cp)
- Horizontal gene transfer (HGT)
- Fecal coliform (FC)
- Fluoroquinolones (Fq)
- Macrolides (Mls)
- Metagenome-assembled genomes (MAGs)
- Mobile genetic elements (MGEs)
- Open reading frames (ORFs)
- Peptides (Ptd)
- Polymerase chain reaction (PCR)
- Polymyxins (Pmx)
- Multidrugs (MDR)
- Rifamycin (Rmc)
- Sulfonamides (Sa)
- Tetracyclines (Tc)

1. Introduction

Antibiotic resistance (AR) is natural and ancient (D'Costa *et al.* 2011). As opposed to dogma that views the presence of antibiotic resistance genes (ARGs) in a place as the consequence of selective pressures, growing literature suggests that the microbial ecological context and niches may be more important to local ARG prevalence in natural and anthropogenic settings (Mahnert *et al.* 2019, Pehrsson *et al.* 2016). However, the niche-influenced assembly depend on various factors, such as microbial phylogenetic and taxonomic structures in human gut (Smillie *et al.* 2011), salinity in soil (Tan *et al.* 2019), microplastics in rivers (Wang *et al.* 2020), and partiles in the air (Xie *et al.* 2019). Strong evidence further suggests that the spread of ARGs is affected by microbial and environmental variables related to mobile genetic elements (MGEs) (Quintela-Baluja *et al.* 2021, Su *et al.* 2015, Wu *et al.* 2019).

As such, different studies have broadly applied this 'microbe-environment-MGE' triad co-driven theory to explain the dynamics of resistomes in riverine, atmospheric, soil and wastewater environments (Gibson *et al.* 2015, Quintela-Baluja *et al.* 2019, Zhu *et al.* 2021, Zhu *et al.* 2017), and even in pristine Arctic areas (McCann *et al.* 2019). However, the specific roles of microbial assembly featured key taxa (Munck *et al.* 2015), pertinent environmental conditions (Sun *et al.* 2020), and MGE-mediated gene transfer (Quintela-Baluja *et al.* 2021) playing in local resistomes have yet to be holistically analyzed (Larsson *et al.* 2018, Winter *et al.* 2021).

A flowing water supply catchment contains a series of distinctive microbial structures (Chen *et al.* 2020) and spatial ecosystems (Maavara *et al.* 2017), and is a useful platform for tracking ARG distribution and analyzing 'microbe-environment-MGE' co-driven mechanisms (Amos *et al.* 2018). A recent study showed that representative taxa groups in different compartments of a wastewater network imposed different selection effects on resistomes across a 'sewer to

river' continuum (Quintela-Baluja *et al.* 2019). But notably, to detect and analyze the dynamics of these niche compartment-specific bacterial taxonomic assemblages, few studies have included ecological assembly-based models (Ning *et al.* 2020), nor considered the in-situ microbial ecology driven environment-filtering effects (Yan *et al.* 2016).

MGEs can be useful biomarkers for AR in water catchments (Quintela-Baluja *et al.* 2021). MGE-laden ARGs, frequently associated with human allochthonous pathogens (Forsberg *et al.* 2012), often relate greater horizontal gene transfer (HGT) potential (D'Costa *et al.* 2006). Therefore, a generally greater prevalence of waterborne ARGs and MGEs exist in places with greater anthropogenic impact (Elder *et al.* 2021, Liang *et al.* 2019). However, clinically important ARGs (*e.g* $bla_{\text{CTX-M}}$ and *mcr*-1) hosted by potential pathogens, like *Enterobacteria, Aeromonads, and V. cholera*, have also been observed in places where human impacts are less apparent (Chen *et al.* 2019, Dang *et al.* 2020, Rysz *et al.* 2013, Walsh *et al.* 2011). This prompts the speculation of a continuous ARG-transfer between the microbial assembly-dominant taxa, which constantly exist in the ecosystem at high abundances, and human-associated microbiomes (Forsberg *et al.* 2014, Hassan *et al.* 2021, Ning *et al.* 2020). However, between human pathogens and assembly-dominant taxa, the intragenomic transfer of 'ARG - MGE' matching pairs have not been evidenced at the community-wide scale yet, especially in the water supply catchment.

Here, we study the main drinking water supply catchment of Shanghai, the largest city in China, to better understand how microbial dynamics and environmental drivers impact in-situ resistomes across a catchment. Using high throughput metagenomic sequencing, microbial community assembly modeling, genome binning related bioinformatics and cell quantification techniques, three goals were achieved: i) detecting and analyzing the hypothesized dominance of microbial assembly over the dynamics of resistome; ii) elucidating the hypothetical ARG

transfer between the microbial assembly-dominant taxa and putative human allochthonous pathogens via genome-warranted approaches; and iii) quantifying density of putative pathogenic ARG-hosts and estimating potential hazards exposed to the catchment users across seasons and geographic gradients.

2. Materials and methods

2.1 Study areas, sampling, and measurement of physicochemical parameters

There were six sampling sites in the studied watershed (**Fig. 1a**), including Taihu Lake pumping station (TP) as the origin; Pingwang Station (PW), Lili Station (LL), and Luxu Station (LX) in Taipu River; and the inlet (JRi) and outlet of Jinze Reservoir (JRe), which provides drinking water to 24 million Shanghai citizens. The sampling campaign was commenced in parallel with Taihu Basin Management Authority's transboundary water diversion from Taihu Lake in Jiangsu Province to the downstream Taipu River towards Shanghai, from June to September 2019 (summer) and November 2019 to February 2020 (winter). A total of 180 individual samples (5 replica×36) were collected along the water catchment. At each sampling site, five replicas (200mL) within a 3 - 5 km sampling region (0.5 to 1.0m in depth) were collected by using a Schindler sampler and then were combined to a sterile PE bottle (1L) before being transferred in the lab for vacuum filtration. Specific information, including sample pretreatment, sampling date, and exact locations are provided in Supplementary Information (**SI-1**).

The pretreated water samples (after settling) were used for the measurement of COD, BOD_5, total nitrogen, total phosphorus, and nitrate using an automated discrete analyzer (SmartChem-200, Italy). Handheld instruments were used to detect temperature, pH (PH-Scan, Shanghai), and dissolved oxygen (DO) (HI9147, Hanna, Italy) on site. Fecal coliform (FC) levels were measured for pretreated water samples on the sampling day, according to previously published

culturing methods (Lamba *et al.* 2017).

2.2 DNA extraction, 16S rRNA gene and metagenomic sequencing

DNA was extracted from vacuum-filtered biosolids collected using sterile 0.22-mm membrane disc filters (Supor® Membrane, Pall Co., USA). Extraction was performed using Power Soil DNA extraction kits (MOBio, USA) and a Ribolyzer (FastPre-24, MP, US), according to manufacturers' suggested protocols. The quality checked samples were sent to Personal Biotech Company Ltd. (Shanghai, China) for library construction (size = 250 bp) and then were applied for 16S rRNA gene (V4 region) amplicon and metagenomic sequencing on the platform of Illumina-NovaSeq (Illumina, USA). The raw sequences were submitted to The National Center for Biotechnology Information (NCBI). The sequence read archive (SRA) can be retrieved via the accession numbers PRJNA730720 (metagenomics sequences) and PRJNA767498 (16S rRNA sequences). Details of primers, library construction and metagenomic sequencing were provided in Supplementary Information (**SI-2**)

2.3 Profile of resistome, dynamics of bacterial community and predicted function

Raw data were initially processed using FASTP to remove low-quality (>= Q20, length > 150bp) sequencing reads, which were analyzed by ARG-oap (v2.3) pipelines to map the mosaic of resistome (id = 80%, threshold value = 10-e10). The relative abundances of detected ARGs were normalized to the copy numbers of (per) identified 16S rRNA genes showing the prevalence of genes (Yang *et al.* 2016). All samples' 16S rRNA amplicon sequences were analyzed by Qiime2-DADA2 (v2020.11) to generate sequence variant (ASV) tables, the taxonomy of which were assigned by using the closed reference approach, against the trained SILVA 138 database. The rarefied BIOM table was further processed to remove the ASVs occurring in less than 10% of the samples (n = 4) with frequencies no less than one. To analyze

the microbial community assembly structures relative to environmental filtering effects, phylogenetic bin-based null model analysis (iCAMP) was utilized (Ning *et al.* 2020); the bin size and ses.cut ($\beta NRI/NTI$) thresholds of each season group were set as 24 and 1.96, respectively. The taxonomic groups mostly contributing to the iCAMP-identified process were regarded as the assembly-dominant taxa (significant phylogenetic signal < 0.05). The metabolic functions of the microbial community were predicted by using FAPROTAX pipelines that were constructed with a database specific to prokaryotic taxa (Louca *et al.* 2016). The details of feature table filtering and rarefication and the parameters invoked in iCAMP and metabolic function prediction pipelines were provided in Supplementary Information **SI-3**.

2.4 Construction of non-redundant metagenome-assembled genomes (MAGs)

The clean paired-end metagenomics reads (~20 Gb/sample) were grouped by sampling periods (season-resolved, **Table S2**) and co-assembled using MEGAHIT v1.13 with default parameters that merge complex bubbles of length <= l*kmer_size (20) and similarity = 0.95. These co-assembled contigs were clustered to recover draft genomes (v1.3.2, length >= 1500 bp) using MaxBin and metaBAT (iterations = 50) as suggested by MetaWRAP (Uritskiy *et al.* 2018). The generated draft genomes were refined to produce high-quality genomic bins using the built-in refining module of MetaWRAP (> 55% completeness & < 5% contamination). All bins were further processed with rapid pairwise genome comparisons at a 95% average nucleotide identity (ANI) threshold (Olm *et al.* 2017) to dereplicate the redundant MAGs (alignment fraction > 10%, greedy algorithms). The obtained non-redundant metagenome-assembled genomes (MAGs) were annotated for taxonomic classification (alignment > 70%) using Genome Taxonomy Database (GTDB, v1.4.0) as previously described (Wu *et al.* 2022).

2.5 Identification of intragenomic ARGs and putative pathogenic hosts

The co-assembled scaffolds of all MAGs were initially processed by Prodigal (v2.6.3; -c -p meta mode) to predict open reading frames (ORFs), which was processed by CD-HIT (v4.6, id = 90%, cov = 90%). ABRicate pipelines (v0.9.9; -id 70%, -qcov 75%) were used to detect the human virulent factors and pathogens (vfdb_setB_nt.fas and PATRIC v3.6.12) and ARG marker genes (SARG v2.2) (Arango-Argoty *et al.* 2018, Liu *et al.* 2019a, Wattam *et al.* 2014, Yin *et al.* 2018). An assembled MAG with queried scaffolds showing the coexistence of intragenomic ARGs and virulent/pathogenic marker genes was regarded as the putative resistant pathogen (Liang *et al.* 2019).

2.6 Estimation of HGT potential and community-wide ARGs transfer

The differences in HGT potentials of intragenomic ARGs between putative resistant pathogens and other MAGs were analyzed according to previously published methods with a few modifications (Li *et al.* 2017). In brief, locations of ARGs and MGEs in scaffolds of MAGs were retrieved from the DIAMOND blastx-mode output (alignment =1, -id 70%, -qcov 75%), with respective to the referencing databases of SARG v2.2 and MEGs90 downloaded from the deepARG-DB (Arango-Argoty *et al.* 2020, Buchfink *et al.* 2015). The minimum genetic distance (minDis) pairs of 'ARGs – MGEs' in the target genomes were calculated. After that, we randomly selected the bacterial MAGs in three groups including putative pathogenic, non-pathogenic, and all MAGs. This random sampling procedure was repeated 100,000 times. The resulting permutation table of minimum distance pairs with a stepwise increase of genetic distance intervals (increment = 50 bp) was used to calculate the probability of the selected ARGs encountering MGEs (Forsberg *et al.* 2014).

All the generated non-redundant ARG-carrying MAGs were applied to MetaCHIP (Song *et al.*

2019), which invokes Prodigal generated- ORFs and BLASTx to align and annotate ARGs (alignment to SARG v2.2) and MGEs as suggested (length > 200bp, -id 90%, -qcov 75%, threshold value = 10-e7). The HGT of 'ARG – MGE' matching pairs disseminated across the candidate MAGs were predicted via the best-match approach (-BP mode) with taxonomic hierarchy ranks from genus to phylum (-r pcofg). The details of HGT estimation methods and calculating pipelines were provided in Supplementary Information (**SI - 3**).

2.7 Quantification of MAGs and potential hazards

The Quant_bin module (default parameters) in the MetaWRAP was used to calculate the relative abundances of constructed MAGs, which were presented in genome copies per million clean sequencing reads (GPMR). The absolute cell counts in the water samples were determined by using a flow cytometer (CytoFLEX, Beckman Coulter, USA), according to the optimized protocols (Nescerecka *et al.* 2016). As suggested by the previous study (Liang *et al.* 2019), the cell density of the MAGs (copy/mL-water) in each sample was quantified according to **Eq. 1**, where *Map.r* referred to the percentage of reads in a sample successfully mapped onto the non-redundant MAG reference (BAM files) using Bowtie2 (Langmead and Salzberg 2012); The portion (and number) of reads in each sample's sequence files generally ranged from 60% to 75% (**Table S2**); *Ab.i* refers to the relative abundance of a target MAG (GPMR); and *i* represented the total number of MAGs.

$$MAG.density = \frac{\text{Absolute cell counts} \times \text{Map.r} \times Ab.i}{\sum_{1}^{n=i} Ab.i} \qquad \textbf{Eq. 1}$$

The relative abundance (coverages in the host MAGs) of intragenomic ARGs (ARG-cov/ppm) were calculated according to **Eq. 2**, where $Ab._{mag}$ referred to the relative abundance of the MAG quantified by Quant_bin module (GPMR) and Cov.g denoted the mean coverage of target ARG located on the identified MAGs' scaffolds; and the mean coverage values of the ARG-

associated scaffolds were calculated via built-in pipelines (pileup.sh) of BBmap (v38.87).

$$Intra.ARG = \text{Ab.mag} \times \text{Cov.g} \quad \textbf{Eq. 2}$$

The comparison of overall potential AR risks (a relative-risk index generated) of the resistome among the samples was initially assessed using MetaCompare (Oh *et al.* 2018). The FASTP-filtered clean metagenomic sequences were processed to calculate the abundance and mobility of environmental resistomes and their hosts' pathogenicity with default parameters (10e−10, id >60%, length >150 bp). The details of calculation of MAGs' relative abundances, reads mapping and waterborne cell quantification were provided in Supplementary Information (**SI – 2** & **SI – 3**).

2.8 Data processing and statistics

The data were log-transformed or scaled to improve sample normality or fitness to specific methods, such as the pairwise t-test and ANOVA test. For datasets that did not fit normal distributions, non-parametric methods were used like Wilcoxon rank-sum test and Kruskal-Wallis rank-sum test. Descriptive analyses of the collected data were performed in Excel 2010 (Microsoft Corp. USA), while the advanced statistical analyses (e.g. Mantel test, Procrustes analysis, partitioning of variations, iCAMP-modeling) were performed using *R* 4.0.2 (https://cran.r-project.org/). The statistical significance was defined at a 95% confidence interval, with a *P*-value of < 0.05 (two-tailed), unless stated otherwise. The details of used statistical methods and installed packages in *R* were explained in detail in Supplementary Information (**SI - 3**).

3. Results and discussion

3.1 Profile and prevalence of the resistomes across the catchment

The Taipu River originates from Taihu Lake in Jiangsu Province and flows to the Jinze Reservoir in Shanghai (**Fig. 1a**). Along this path, normalized ARG levels in bacterial communities decrease significantly (**Fig. 1b**) from -1.09 ± 0.17 to -1.41 ± 0.08 ($\log_{10}$(ARGs/16S rRNA gene) based on all samples collected over the study (Two-way ANOVA, F = 12.1, P_{site} < 0.01; P_{season} = 0.12). This decreasing trend from upstream to downstream is probably related to the reservoir resettlement (Chen *et al.* 2019) and strict management of the neighboring watershed, where no sewage outfalls and substantially fewer factories are located (**Fig. 1a**). Nevertheless, the mean relative ARG abundances in the Jinze Reservior resistome was -1.30 ± 0.15 $\log_{10}$(ARGs/16S rRNA gene), which is in the same proportional range as municipal wastewater or activated sludge (**Fig. S1**). Futhermore, the prevalence of specific ARGs is particularly concerning (**Fig. S2**), which encoded resistances to medical antiseptics (*quaE*, -2.70 ± 0.18 $\log_{10}$(ARGs/16S rRNA gene)) and frontline antibiotics used in clinical settings (Buffet-Bataillon *et al.* 2012, Petrovich *et al.* 2020), such as carbapenem (bla_{OXA}; -3.05 ± 0.50 $\log_{10}$(ARGs/16S rRNA gene)), rifamycin (*Arr*; -3.55 ± 0.22 $\log_{10}$(ARGs/16S rRNA gene)), vancomycin (*vanR*/*S*; -3.49 ± 0.30 $\log_{10}$(ARGs/16S rRNA gene)), and polymyxin (*mcr*-1/5; -4.41 ± 0.38 $\log_{10}$(ARGs/16S rRNA gene)). These ARGs may stem from the Taipu River's upstream quayside aquaculture farms (**Fig. 1a**), which often misuse antimicrobials and discharge their wastewater along with the water diversion process (Hong et al. 2018, Jiang et al. 2013, Perry et al. 2019, Su and Chen 2020).

3.2 Resistomes determined by microbial assembly and environmental filtering

Resistomes across the catchment were comprised of 103 types of ARGs (**Fig. S2**), which were

relatively consistent at our sites during summer and winter smapling (Adonis2, $P > 0.05$, **Fig. S3**). The microbial community also exhibited few variations as the Taipu River flowed downstream (Adonis2, $P_{site} = 0.44$). The observed consistency of site-specific resistomes and microbiomes suggested a close association between ARGs and resident bacteria in the catchment (Procrustes tests, $M^2 = 0.78$, R = 0.46, $P = 0.002$, **Fig. S4**).

To investigate this relationship from a microbial ecological perspective (Ning *et al.* 2020), season-differentiated microbial assembly processes were simulated (Adonis2, $P_{season} = 0.01$). As shown in **Fig. 1c**, bacteria in the studied catchment appeared to assemble in a homogeneous selection-dominated approach (HoS, 48% ~ 53%) in both summer and winter samples (ses.cut = 1.96, iCAMP), irrespective of variations in geographic locations (One-way ANOVA, $P = 0.17$). Among the identified microbial assemblages (**Table S3** & **Table S4**), *Microcystis* (26.4% in abundance), *Cyanobiaceae* (10.6%), *Ilumatobacteraceae* (8.9%), and *Burkholderiales* sp. (0.75%) mostly contributed to the HoS process in summer (~ 85%n, βNRI = -2.5) and winter (~ 90%, βNRI = -2.2, **Fig. 1c**). Notably, these assembly-dominant taxa not only exhibited significant correlations to the resistome (Bray-Curtis dissimilarity indices distance, $P < 0.001$, **Fig. 2a**), but also imposed greater impact on ARGs compared with whole microbial taxa (Cohen's D = 4.22 *vs.*3.39, $P < 0.001$, **Fig.2a**).

The observed high importance (~ 50%) of HoS in the bacterial assembly process ($|\beta$NRI$| > 2.0$, **Fig. 1c**) suggests substantial environmental filtering effects on microbial assemblages (Xu *et al.* 2020, Yan *et al.* 2016). Here, the forward-selected environmental factors (rda, AIC = 278.5, $P < 0.05$, **Table 1**), including temperature and DO, significantly related to the deterministic assembly of the microbial community ($P < 0.05$, **Fig.2b**). Decreases in temperature and DO levels (**Table 1**) mirroed microbial responses, especially photoautotrophic processes, such as the oxygenic and photosynthetic pathways of *Cyanobacteria* (LEfSe, LDA = 3.0, $P < 0.01$, **Fig.**

306 **2c** & **Table S6**). Negative effects of low temperature and DO have been shown for the
307 planktobacteria, including *Microcystis and Ilumatobacteraceae* (Mo *et al.* 2018, Zakharova *et*
308 *al.* 2021), which also appear to be resistome-determining taxa in our systems (Cohen's D =
309 4.22, $P < 0.01$, **Fig. 2a**). Environmental filtering and related microbial metabolic responses had
310 a comparable infulence to that of assembly-dominant taxa on ARG demongraphics in the
311 studied catchment (VPA, 12 - 15%, **Table S5**).

312 *3.3 Seasonally distributed putative resistant pathogens and intragenomic ARGs*

313 There were 266 and 289 non-redundant MAGs obtained from summer and winter samples,
314 respectively, around 20% of which were identified as ARG-hosting bacteria (**Fig. 3**). Overall,
315 these ARG hosts were significantly more abundant across the catchment in summer than in
316 winter (76.8 ± 9.08 vs. 59.0 ± 5.69 GPMR, Pair-wise t-test, $P < 0.01$) and generally belonged
317 to the phyla of *Proteobacteria*, *Bacteroidota* and *Actinobacteria*. Among them, most putative
318 pathogens were *Mycobacteriaceae* (6.83 ± 0.54 GPMR) and *Burkholderiaceae* (19.7 ± 6.51
319 GPMR), of which the distribution and concentrations did not significantly vary across sampling
320 sites (One-way ANOVA, $P > 0.05$; ADONIS, $P = 0.1$, **Fig. S5**). The predominant subtaxon
321 species, such as *M. mycobacterium* spp., and *B. ramibacter* spp. and *Limnohabitans* spp. (**Fig.**
322 **3a**), are frequently associated with human AR infections (Fang *et al.* 2015, Fang *et al.* 2019).
323 In winter samples (**Fig. 3b**), MAGs of putative resistant pathogens were more dispersed across
324 *Actinobacteria* phylum, including *Aeromicrobium*, *Nocardioides*, and *Phycicoccus*. These taxa
325 are known to have ill effects on human dermal and intestine health (Singleton *et al.* 2022, Zhang
326 *et al.* 2022). However, the predominant putative resistant pathogens still primarily belonged to
327 *Burkholderiaceae* (12.9 ± 3.83 GPMR) and were evenly distributed across the river (One-way
328 ANOVA, $P > 0.05$).

329 Intragenomic ARGs of putative pathogens were detected with significantly higher diversity

and abundance than their counterparts located in other MAGs during the whole period of study (Kruskal-Wallis test, $\chi^2 = 10.6$, $P < 0.01$). In summer samples (**Fig. 3a**), *Burkholderiaceae* and *Mycobacteriaceae* MAGs hosted more than eight types of ARGs, encoding resistance to aminoglycosides, rifamycin, extended-spectrum β-lactams and macrolides (1.5 – 2.5 $\log_{10}$(ARG-cov/ppm)). In contrast, in winter samples, the predominant constituents of intragenomic resistome were bacitracin (~ 2.5 $\log_{10}$(ARG-cov/ppm)), rifamycin (~ 1.5 $\log_{10}$(ARG-cov/ppm)), and aminoglycoside (~ 1.0 $\log_{10}$(ARG-cov/ppm)) ARGs (**Fig 3b**), most of which were associated with *Burkholderiaceae* spp., such as *B. limnohabitans* and *Rhizobacter* spp.

3.4 Importance of putative pathogens to community-wide HGT of ARG

The observed larger and diverse intragenomic ARGs in putative pathogenic bacteria may result from frequent HGT of ARGs (Zhu *et al.* 2013). As shown in **Fig. 4a**, the normalized intragenomic MGE levels, often an approximation of HGT potential (Lamba *et al.* 2017), were significantly correlated with intragenomic ARGs across all samples (Pearson, $P < 0.05$). Importantly, MGEs with apparent links to putative pathogenic MAGs exhibited stronger linear correlations with intragenomic ARGs, suggesting that HGT of ARGs within in situ river microbial communities might be more likely to proceed via resistant pathogenic bacteria (Cohen's D effect size = 0.84 *vs.* 0.41, $P < 0.001$, **Fig. 4a**). This high HGT potential of ARGs in putative pathogens was seen in both summer and winter samples (Cohen's D effect size > 0.6, $P < 0.001$, **Fig. 4a**). Moreover, as shown in **Fig. 4b**, ARGs associated with pathogens had relatively short genetic distancesto MGEs on their genetic context contigs (e.g. < 3.5 kb), which is consistent with the previous work that showed greater HGT potential of ARGs in human pathogens compared with overall MAG assemblages (Forsberg *et al.* 2014).

The importance of pathogens to dissemination of ARGs was further suggested by a community-

wide HGT analysis of all ARG hosting MAGs using the MetaCHIP (Song *et al.* 2019). **Fig 5a** shows that HGT of ARGs more frequently associated with the putative resistant pathogens, irrespective of their roles (i.e. donor *vs.* recipients) in the HGT processes (Kruskal-Wallis test, $\chi^2 = 17.5$, $P_{path} < 0.001$, $P_{role} = 0.31$). Notably, HGT-involved MAGs primarily belonged to the subtaxa of *Proteobacteria* and *Actinobacteria* (**Fig 5a**), including *Burkholderiaceae* (count = 126.8 ± 78.4), *Ilumatobacteraceae* (count = 12.4 ± 6.1) and *Microbacteriaceae* (count = 7.25 ± 5.6). These MAGs not only exhibited more frequent ARG transmission (Kruskal-Wallis test, $\chi^2 = 96.3$, $P < 0.001$), but also were identified as the microbial assembly-dominant taxa that shaped the resistomes (**Fig. 2a** & **Table S4**).

Although *Cyanobacteria* dominated microbial assembly and resistome variations, evidence suggests they were not closely invovled in the community-wide ARG HGT (count < 3.0, **Fig 5a**) and their acquired intragenomic ARGs all originated from putative pathogenic MAGs of *Burkholderiaceae* sp. (**Fig. S7** & **Fig. S8**) with high HGT potentials (**Fig. 5a**). From the microbial synergistic perspective, *Burkholderiaceae* can utilize algal excremets (e.g. microcystin) as major metabolic substrates (Salter et al. 2021), therefore the cyanobacterial aggragates tend to closely structure with *Burkholderiaceae* in lake-river ecosystems (Eiler and Bertilsson 2004) . As reported by Guo et al. (2018), waterborne ARGs and their hosting bacteria were abundantly observed in the bacterioplanktonic aggregates. In the homogeneous selection dominated community (**Fig. 1**), *Cyanobacteria* exhibit high resilliance to environmental variations (Liu et al. 2019b, Mo et al. 2018), which could be beneficial to preserve their associated ARGs (**Fig. 5**). The predominant intragenomic ARGs, including β-lactam resistant (e.g. bla_{OXA}) and MLS resistant genes (*mac*B), appeared to be disseminated between microbial assembly-dominant taxa and putative resistant pathogens (**Fig. 5b**), which explains the high prevalence and the spread of ARGs across the catchment during the period of water diversion.

3.5 Implications of potential AR hazards and influencing factors

The prevalence of ARGs with potential clinical importance (**Fig. 1c** & **Fig. S2**) and putative resistant pathogens (**Fig. 3**) may suggest higher AR risks (Martinez *et al.* 2015), especially if they are found in drinking water sources (Dang *et al.* 2020, Walsh *et al.* 2011). Overall, summer samples showed higher potential AR hazards using MetaCompare (index values = 19.3 *vs.* 19.0). The potential exposure risks were significantly lower in the reservoir *vs*. upstream (Two-way ANOVA, $P < 0.05$, **Fig. S8**). This trend was mirrored by lower density of putative resistant pathogens in the reservoir water samples, *i.e.* $3.79 \pm 2.5 \times 10^6$ cell/L on average (One-way ANOVA, $F = 31.4$, $P < 0.001$, **Fig. 6a**). This value corresponds to a total amount of $13.3 \pm 0.89 \times 10^{15}$ cells of putative resistant pathogens that are transferred daily to Shanghai's water supply system during the water diversion process (3.5 million m^3/day). However, it should be noted that this study was confined to just one catchment. For better generalization of at-tap drinking water source-specific AR, a combination of metagenomics and culture-based studies that comprehensively encompass water sources, distribution networks and household tap water at a national or continental scale is needed.

A more in-depth analysis using structural equation modelling (SEM) revealed direct effects concerning the distribution of resistome (std-coeff $\beta = 0.69$, $P < 0.05$) and microbial assembly-dominant bacteria (std-coeff $\beta = 0.15$, $P < 0.05$) on the absolute concentration of waterborne putative resistant pathogens across the studied catchment (**Fig. 6b**). Nevertheless, the environmental filtering factors including temperature and DO values (SEM, std-coeff $|\beta| >= 0.8$, $P < 0.05$), only showed indirect effects via their influence on the microbial assembly-dominant bacterial taxa (SEM, std-coeff $\beta = -0.45$, $P < 0.05$). This suggests that improvement of water quality may have a limited effect on reducing in-situ resistome transmission and resistant pathogens ($P > 0.05$, **Fig. 6b**), whereas the holistic management of wastewater inputs

containing pathogens (Jiang *et al.* 2013, Quintela-Baluja *et al.* 2019) and control of microbial assembly-dominant bacterial taxa, such as curbing algal (*e.g. Cyanobacteria*) blooms are critically needed (Guo *et al.* 2018, Mo *et al.* 2018, Xue *et al.* 2018).

4. Conclusions

Here, we show that microbial deterministic assembly process is critical to the resulting environmental antibiotic resistomes in the Taipu River catchment. The seasonal variations of ARGs, including the ones of particular importance to human health, were frequently hosted by putative pathogenic bacteria, which greatly contributed to the community-wide HGT of ARGs, especially with the assembly-dominant taxa.

The pipelines constructed to estimate ARGs' HGT potentials can be applied more generally in a wide range of environmental settings, for quantitative comparison and analysis of resistome mobility and potential AR risks. Importantly, in the annually operated water diversion process in the Taihu Basin, the waterborne ARGs released from upstream sources are 'well-preserved' in the niched bacterial community and transferred to the downstream drinking water reservoir, which presumably contains $3.0 \sim 4.0 \times 10^6$ cell/L of the putative resistant pathogens. Although our data are from samples prior to water treatment, there is an implicit concern about AR spread through the local water supply, possibly impacting 24 million Shanghai citizens. However, the work also shows importance of environmental factors and microbial spatial ecology on ARG fate, which provides a useful starting point for developing strategies to curb the AR transmission in the Shanghai or other water supply catchments.

CRediT Authorship Contribution Statement

K.Y. and B.X. designed and obtained funding for this study; D.W., Y-L.S., and J.Z. collected

samples and raw sequencing data and conducted the experiments; D.W., J.Z., and M-J. Y. undertook the bioinformatics analyses and contributed to the data visualization. D.W., D.W.G. Y-L.S., J.D., and B.X. wrote the paper. All of the authors read and approved the final manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was financially supported by the National Key Research and Development Program of China (2018YFC1901000), Major Science and Technology Program for Water Pollution Control and Treatment (2017ZX07207003-01), National Natural Science Foundation of China (42107457 and 22276059), Key Laboratory of Environmental Pollution Monitoring and Disease Control, Ministry of Education, Guizhou Medical University (KY2022383). Guangdong-Hong Kong-Macao Joint Laboratory for Contaminants Exposure and Health, Guangdong University of Technology (GHMJLCEH-12). The authors thank Editor, Reviewers, and Prof. Tong Zhang and his team from the University of Hong Kong for their suggestions and comments to improve the quality of the manuscript.

Supplementary Materials

The Supplementary Information (SI) is available free of charge on the Elsevier Publication Water Research website (https://www.journals.elsevier.com/water-research). The provided SI was comprised of four sections including sampling and pretreatment information (**SI-1**), details of utilized microbial molecular techniques (**SI-2**), bioinformatics methods and related

447 calculations (**SI-3**), and supplementary analyzing results (**SI-4**).

References

Amos, G.C.A., Ploumakis, S., Zhang, L., Hawkey, P.M., Gaze, W.H. and Wellington, E.M.H. (2018) The Widespread Dissemination of Integrons Throughout Bacterial Communities in a Riverine System. ISME J 12(3), 681-691. doi:10.1038/s41396-017-0030-8

Arango-Argoty, G., Garner, E., Pruden, A., Heath, L.S., Vikesland, P. and Zhang, L. (2018) Deeparg: A Deep Learning Approach for Predicting Antibiotic Resistance Genes from Metagenomic Data. Microbiome 6(1), 23. doi:10.1186/s40168-018-0401-z

Arango-Argoty, G.A., Guron, G.K.P., Garner, E., Riquelme, M.V., Heath, L.S., Pruden, A., Vikesland, P.J. and Zhang, L. (2020) Argminer: A Web Platform for the Crowdsourcing-Based Curation of Antibiotic Resistance Genes. Bioinformatics 36(9), 2966-2973. doi:10.1093/bioinformatics/btaa095

Buchfink, B., Xie, C. and Huson, D.H. (2015) Fast and Sensitive Protein Alignment Using Diamond. Nature Methods 12(1), 59-60. doi:DOI 10.1038/nmeth.3176

Buffet-Bataillon, S., Tattevin, P., Bonnaure-Mallet, M. and Jolivet-Gougeon, A. (2012) Emergence of Resistance to Antibacterial Agents: The Role of Quaternary Ammonium Compounds--a Critical Review. Int J Antimicrob Agents 39(5), 381-389. doi:10.1016/j.ijantimicag.2012.01.011

Chen, H., Li, Y., Sun, W., Song, L., Zuo, R. and Teng, Y. (2020) Characterization and Source Identification of Antibiotic Resistance Genes in the Sediments of an Interconnected River-Lake System. Environment International 137, 105538. doi:10.1016/j.envint.2020.105538

Chen, Y., Su, J.Q., Zhang, J., Li, P., Chen, H., Zhang, B., Gin, K.Y. and He, Y. (2019) High-Throughput Profiling of Antibiotic Resistance Gene Dynamic in a Drinking Water River-Reservoir System. Water Res 149, 179-189. doi:10.1016/j.watres.2018.11.007

D'Costa, V.M., King, C.E., Kalan, L., Morar, M., Sung, W.W., Schwarz, C., Froese, D., Zazula, G., Calmels, F., Debruyne, R., Golding, G.B., Poinar, H.N. and Wright, G.D. (2011) Antibiotic Resistance Is Ancient. Nature 477(7365), 457-461. doi:10.1038/nature10388

D'Costa, V.M., McGrann, K.M., Hughes, D.W. and Wright, G.D. (2006) Sampling the Antibiotic Resistome. Science 311(5759), 374-377. doi:10.1126/science.1120800

Dang, C., Xia, Y., Zheng, M., Liu, T., Liu, W., Chen, Q. and Ni, J. (2020) Metagenomic Insights into the Profile of Antibiotic Resistomes in a Large Drinking Water Reservoir. Environment International 136, 105449. doi:10.1016/j.envint.2019.105449

Eiler, A. and Bertilsson, S. (2004) Composition of Freshwater Bacterial Communities Associated with Cyanobacterial Blooms in Four Swedish Lakes. Environ Microbiol 6(12), 1228-1243. doi:10.1111/j.1462-2920.2004.00657.x

Elder, F.C.T., Proctor, K., Barden, R., Gaze, W.H., Snape, J., Feil, E.J. and Kasprzyk-Hordern, B. (2021) Spatiotemporal Profiling of Antibiotics and Resistance Genes in a River Catchment: Human Population as the Main Driver of Antibiotic and Antibiotic Resistance Gene Presence

485 in the Environment. Water Res 203, 117533. doi:10.1016/j.watres.2021.117533

486 Fang, H., Wang, H., Cai, L. and Yu, Y. (2015) Prevalence of Antibiotic Resistance Genes and
487 Bacterial Pathogens in Long-Term Manured Greenhouse Soils as Revealed by Metagenomic
488 Survey. Environ Sci Technol 49(2), 1095-1104. doi:10.1021/es504157v

489 Fang, P., Peng, F., Gao, X., Xiao, P. and Yang, J. (2019) Decoupling the Dynamics of Bacterial
490 Taxonomy and Antibiotic Resistance Function in a Subtropical Urban Reservoir as Revealed
491 by High-Frequency Sampling. Front Microbiol 10, 1448. doi:10.3389/fmicb.2019.01448

492 Forsberg, K.J., Patel, S., Gibson, M.K., Lauber, C.L., Knight, R., Fierer, N. and Dantas, G.
493 (2014) Bacterial Phylogeny Structures Soil Resistomes across Habitats. Nature 509(7502),
494 612-616. doi:10.1038/nature13377

495 Forsberg, K.J., Reyes, A., Wang, B., Selleck, E.M., Sommer, M.O. and Dantas, G. (2012) The
496 Shared Antibiotic Resistome of Soil Bacteria and Human Pathogens. Science 337(6098), 1107-
497 1111. doi:10.1126/science.1220761

498 Gibson, M.K., Forsberg, K.J. and Dantas, G. (2015) Improved Annotation of Antibiotic
499 Resistance Determinants Reveals Microbial Resistomes Cluster by Ecology. ISME J 9(1), 207-
500 216. doi:10.1038/ismej.2014.106

501 Guo, Y., Liu, M., Liu, L., Liu, X., Chen, H. and Yang, J. (2018) The Antibiotic Resistome of
502 Free-Living and Particle-Attached Bacteria under a Reservoir Cyanobacterial Bloom. Environ
503 Int 117, 107-115. doi:10.1016/j.envint.2018.04.045

504 Hassan, B., Ijaz, M., Khan, A., Sands, K., Serfas, G.I., Clayfield, L., El-Bouseary, M.M., Lai,
505 G., Portal, E., Khan, A., Watkins, W.J., Parkhill, J. and Walsh, T.R. (2021) A Role for
506 Arthropods as Vectors of Multidrug-Resistant Enterobacterales in Surgical Site Infections from
507 South Asia. Nat Microbiol 6(10), 1259-1270. doi:10.1038/s41564-021-00965-1

508 Jiang, L., Hu, X., Xu, T., Zhang, H., Sheng, D. and Yin, D. (2013) Prevalence of Antibiotic
509 Resistance Genes and Their Relationship with Antibiotics in the Huangpu River and the
510 Drinking Water Sources, Shanghai, China. Sci Total Environ 458-460, 267-272.
511 doi:10.1016/j.scitotenv.2013.04.038

512 Lamba, M., Graham, D.W. and Ahammad, S.Z. (2017) Hospital Wastewater Releases of
513 Carbapenem-Resistance Pathogens and Genes in Urban India. Environ Sci Technol 51(23),
514 13906-13912. doi:10.1021/acs.est.7b03380

515 Langmead, B. and Salzberg, S.L. (2012) Fast Gapped-Read Alignment with Bowtie 2. Nature
516 Methods 9(4), 357-359. doi:10.1038/nmeth.1923

517 Larsson, D.G.J., Andremont, A., Bengtsson-Palme, J., Brandt, K.K., de Roda Husman, A.M.,
518 Fagerstedt, P., Fick, J., Flach, C.F., Gaze, W.H., Kuroda, M., Kvint, K., Laxminarayan, R.,
519 Manaia, C.M., Nielsen, K.M., Plant, L., Ploy, M.C., Segovia, C., Simonet, P., Smalla, K., Snape,
520 J., Topp, E., van Hengel, A.J., Verner-Jeffreys, D.W., Virta, M.P.J., Wellington, E.M. and
521 Wernersson, A.S. (2018) Critical Knowledge Gaps and Research Needs Related to the
522 Environmental Dimensions of Antibiotic Resistance. Environment International 117, 132-138.
523 doi:10.1016/j.envint.2018.04.041

524 Li, L.G., Xia, Y. and Zhang, T. (2017) Co-Occurrence of Antibiotic and Metal Resistance
525 Genes Revealed in Complete Genome Collection. ISME J 11(3), 651–662.
526 doi:10.1038/ismej.2016.155

527 Liang, J., Mao, G., Yin, X., Ma, L., Liu, L., Bai, Y., Zhang, T. and Qu, J. (2019) Identification
528 and Quantification of Bacterial Genomes Carrying Antibiotic Resistance Genes and Virulence
529 Factor Genes for Aquatic Microbiological Risk Assessment. Water Res 168, 115160.
530 doi:10.1016/j.watres.2019.115160

531 Liu, B., Zheng, D., Jin, Q., Chen, L. and Yang, J. (2019a) Vfdb 2019: A Comparative
532 Pathogenomic Platform with an Interactive Web Interface. Nucleic Acids Res 47(D1), D687-
533 D692. doi:10.1093/nar/gky1080

534 Liu, L., Chen, H., Liu, M., Yang, J.R., Xiao, P., Wilkinson, D.M. and Yang, J. (2019b) Response
535 of the Eukaryotic Plankton Community to the Cyanobacterial Biomass Cycle over 6 Years in
536 Two Subtropical Reservoirs. ISME J 13(9), 2196-2208. doi:10.1038/s41396-019-0417-9

537 Louca, S., Parfrey, L.W. and Doebeli, M. (2016) Decoupling Function and Taxonomy in the
538 Global Ocean Microbiome. Science 353(6305), 1272-1277. doi:10.1126/science.aaf4507

539 Maavara, T., Lauerwald, R., Regnier, P. and Van Cappellen, P. (2017) Global Perturbation of
540 Organic Carbon Cycling by River Damming. Nat Commun 8, 15347.
541 doi:10.1038/ncomms15347

542 Mahnert, A., Moissl-Eichinger, C., Zojer, M., Bogumil, D., Mizrahi, I., Rattei, T., Martinez,
543 J.L. and Berg, G. (2019) Man-Made Microbial Resistances in Built Environments. Nat
544 Commun 10(1), 968. doi:10.1038/s41467-019-08864-0

545 Martinez, J.L., Coque, T.M. and Baquero, F. (2015) What Is a Resistance Gene? Ranking Risk
546 in Resistomes. Nat Rev Microbiol 13(2), 116-123. doi:10.1038/nrmicro3399

547 McCann, C.M., Christgen, B., Roberts, J.A., Su, J.Q., Arnold, K.E., Gray, N.D., Zhu, Y.G. and
548 Graham, D.W. (2019) Understanding Drivers of Antibiotic Resistance Genes in High Arctic
549 Soil Ecosystems. Environment International 125, 497-504. doi:10.1016/j.envint.2019.01.034

550 Mo, Y., Zhang, W., Yang, J., Lin, Y., Yu, Z. and Lin, S. (2018) Biogeographic Patterns of
551 Abundant and Rare Bacterioplankton in Three Subtropical Bays Resulting from Selective and
552 Neutral Processes. ISME J 12(9), 2198-2210. doi:10.1038/s41396-018-0153-6

553 Munck, C., Albertsen, M., Telke, A., Ellabaan, M., Nielsen, P.H. and Sommer, M.O. (2015)
554 Limited Dissemination of the Wastewater Treatment Plant Core Resistome. Nat Commun 6,
555 8452. doi:10.1038/ncomms9452

556 Nescerecka, A., Hammes, F. and Juhna, T. (2016) A Pipeline for Developing and Testing
557 Staining Protocols for Flow Cytometry, Demonstrated with Sybr Green I and Propidium Iodide
558 Viability Staining. Journal of Microbiological Methods 131, 172-180.
559 doi:10.1016/j.mimet.2016.10.022

560 Ning, D., Yuan, M., Wu, L., Zhang, Y., Guo, X., Zhou, X., Yang, Y., Arkin, A.P., Firestone,
561 M.K. and Zhou, J. (2020) A Quantitative Framework Reveals Ecological Drivers of Grassland

562 Microbial Community Assembly in Response to Warming. Nat Commun 11(1), 4717.
563 doi:10.1038/s41467-020-18560-z

564 Oh, M., Pruden, A., Chen, C., Heath, L.S., Xia, K. and Zhang, L. (2018) Metacompare: A
565 Computational Pipeline for Prioritizing Environmental Resistome Risk. FEMS Microbiol Ecol
566 94(7), fiy079. doi:10.1093/femsec/fiy079

567 Olm, M.R., Brown, C.T., Brooks, B. and Banfield, J.F. (2017) Drep: A Tool for Fast and
568 Accurate Genomic Comparisons That Enables Improved Genome Recovery from
569 Metagenomes through De-Replication. ISME J 11(12), 2864-2868.
570 doi:10.1038/ismej.2017.126

571 Pehrsson, E.C., Tsukayama, P., Patel, S., Mejia-Bautista, M., Sosa-Soto, G., Navarrete, K.M.,
572 Calderon, M., Cabrera, L., Hoyos-Arango, W., Bertoli, M.T., Berg, D.E., Gilman, R.H. and
573 Dantas, G. (2016) Interconnected Microbiomes and Resistomes in Low-Income Human
574 Habitats. Nature 533(7602), 212-216. doi:10.1038/nature17672

575 Petrovich, M.L., Zilberman, A., Kaplan, A., Eliraz, G.R., Wang, Y., Langenfeld, K., Duhaime,
576 M., Wigginton, K., Poretsky, R., Avisar, D. and Wells, G.F. (2020) Microbial and Viral
577 Communities and Their Antibiotic Resistance Genes Throughout a Hospital Wastewater
578 Treatment System. Front Microbiol 11, 153. doi:10.3389/fmicb.2020.00153

579 Quintela-Baluja, M., Abouelnaga, M., Romalde, J., Su, J.Q., Yu, Y., Gomez-Lopez, M., Smets,
580 B., Zhu, Y.G. and Graham, D.W. (2019) Spatial Ecology of a Wastewater Network Defines the
581 Antibiotic Resistance Genes in Downstream Receiving Waters. Water Res 162, 347-357.
582 doi:10.1016/j.watres.2019.06.075

583 Quintela-Baluja, M., Frigon, D., Abouelnaga, M., Jobling, K., Romalde, J.L., Gomez Lopez,
584 M. and Graham, D.W. (2021) Dynamics of Integron Structures across a Wastewater Network -
585 Implications to Resistance Gene Transfer. Water Res 206, 117720.
586 doi:10.1016/j.watres.2021.117720

587 Rysz, M., Mansfield, W.R., Fortner, J.D. and Alvarez, P.J. (2013) Tetracycline Resistance Gene
588 Maintenance under Varying Bacterial Growth Rate, Substrate and Oxygen Availability, and
589 Tetracycline Concentration. Environ Sci Technol 47(13), 6995-7001. doi:10.1021/es3035329

590 Salter, C., VanMensel, D., Reid, T., Birbeck, J., Westrick, J., Mundle, S.O.C. and Weisener,
591 C.G. (2021) Investigating the Microbial Dynamics of Microcystin-Lr Degradation in Lake Erie
592 Sand. Chemosphere 272, 129873. doi:10.1016/j.chemosphere.2021.129873

593 Singleton, C.M., Petriglieri, F., Wasmund, K., Nierychlo, M., Kondrotaite, Z., Petersen, J.F.,
594 Peces, M., Dueholm, M.S., Wagner, M. and Nielsen, P.H. (2022) The Novel Genus, 'Candidatus
595 Phosphoribacter', Previously Identified as Tetrasphaera, Is the Dominant Polyphosphate
596 Accumulating Lineage in Ebpr Wastewater Treatment Plants Worldwide. ISME J.
597 doi:10.1038/s41396-022-01212-z

598 Smillie, C.S., Smith, M.B., Friedman, J., Cordero, O.X., David, L.A. and Alm, E.J. (2011)
599 Ecology Drives a Global Network of Gene Exchange Connecting the Human Microbiome.
600 Nature 480(7376), 241-244. doi:10.1038/nature10571

Song, W., Wemheuer, B., Zhang, S., Steensen, K. and Thomas, T. (2019) Metachip: Community-Level Horizontal Gene Transfer Identification through the Combination of Best-Match and Phylogenetic Approaches. Microbiome 7(1), 36. doi:10.1186/s40168-019-0649-y

Su, J.Q., Wei, B., Ou-Yang, W.Y., Huang, F.Y., Zhao, Y., Xu, H.J. and Zhu, Y.G. (2015) Antibiotic Resistome and Its Association with Bacterial Communities During Sewage Sludge Composting. Environ Sci Technol 49(12), 7356-7363. doi:10.1021/acs.est.5b01012

Sun, J., Liao, X.P., D'Souza, A.W., Boolchandani, M., Li, S.H., Cheng, K., Luis Martinez, J., Li, L., Feng, Y.J., Fang, L.X., Huang, T., Xia, J., Yu, Y., Zhou, Y.F., Sun, Y.X., Deng, X.B., Zeng, Z.L., Jiang, H.X., Fang, B.H., Tang, Y.Z., Lian, X.L., Zhang, R.M., Fang, Z.W., Yan, Q.L., Dantas, G. and Liu, Y.H. (2020) Environmental Remodeling of Human Gut Microbiota and Antibiotic Resistome in Livestock Farms. Nat Commun 11(1), 1427. doi:10.1038/s41467-020-15222-y

Tan, L., Wang, F., Liang, M., Wang, X., Das, R., Mao, D. and Luo, Y. (2019) Antibiotic Resistance Genes Attenuated with Salt Accumulation in Saline Soil. J Hazard Mater 374, 35-42. doi:10.1016/j.jhazmat.2019.04.020

Uritskiy, G.V., DiRuggiero, J. and Taylor, J. (2018) Metawrap-a Flexible Pipeline for Genome-Resolved Metagenomic Data Analysis. Microbiome 6(1), 158. doi:10.1186/s40168-018-0541-1

Walsh, T.R., Weeks, J., Livermore, D.M. and Toleman, M.A. (2011) Dissemination of Ndm-1 Positive Bacteria in the New Delhi Environment and Its Implications for Human Health: An Environmental Point Prevalence Study. Lancet Infect Dis 11(5), 355-362. doi:10.1016/s1473-3099(11)70059-7

Wang, J., Qin, X., Guo, J., Jia, W., Wang, Q., Zhang, M. and Huang, Y. (2020) Evidence of Selective Enrichment of Bacterial Assemblages and Antibiotic Resistant Genes by Microplastics in Urban Rivers. Water Res 183, 116113. doi:10.1016/j.watres.2020.116113

Wattam, A.R., Abraham, D., Dalay, O., Disz, T.L., Driscoll, T., Gabbard, J.L., Gillespie, J.J., Gough, R., Hix, D., Kenyon, R., Machi, D., Mao, C., Nordberg, E.K., Olson, R., Overbeek, R., Pusch, G.D., Shukla, M., Schulman, J., Stevens, R.L., Sullivan, D.E., Vonstein, V., Warren, A., Will, R., Wilson, M.J., Yoo, H.S., Zhang, C., Zhang, Y. and Sobral, B.W. (2014) Patric, the Bacterial Bioinformatics Database and Analysis Resource. Nucleic Acids Res 42(Database issue), D581-591. doi:10.1093/nar/gkt1099

Winter, M., Buckling, A., Harms, K., Johnsen, P.J. and Vos, M. (2021) Antimicrobial Resistance Acquisition Via Natural Transformation: Context Is Everything. Current Opinion in Microbiology 64, 133-138. doi:10.1016/j.mib.2021.09.009

Wu, D., Jin, L., Xie, J., Liu, H., Zhao, J., Ye, D. and Li, X.-d. (2022) Inhalable Antibiotic Resistomes Emitted from Hospitals: Metagenomic Insights into Bacterial Hosts, Clinical Relevance, and Environmental Risks. Microbiome 10(1), in press. doi:10.1186/s40168-021-01197-5

Wu, D., Su, Y., Xi, H., Chen, X. and Xie, B. (2019) Urban and Agriculturally Influenced Water Contribute Differently to the Spread of Antibiotic Resistance Genes in a Mega-City River

641 Network. Water Res 158, 11-21. doi:10.1016/j.watres.2019.03.010

642 Xie, J., Jin, L., He, T., Chen, B., Luo, X., Feng, B., Huang, W., Li, J., Fu, P. and Li, X. (2019)
643 Bacteria and Antibiotic Resistance Genes (Args) in Pm2.5 from China: Implications for Human
644 Exposure. Environ Sci Technol 53(2), 963-972. doi:10.1021/acs.est.8b04630

645 Xu, X., Wang, N., Lipson, D., Sinsabaugh, R., Schimel, J., He, L., Soudzilovskaia, N.A.,
646 Tedersoo, L. and Algar, A.C. (2020) Microbial Macroecology: In Search of Mechanisms
647 Governing Microbial Biogeographic Patterns. Global Ecology and Biogeography 29(11),
648 1870-1886. doi:10.1111/geb.13162

649 Xue, Y., Chen, H., Yang, J.R., Liu, M., Huang, B. and Yang, J. (2018) Distinct Patterns and
650 Processes of Abundant and Rare Eukaryotic Plankton Communities Following a Reservoir
651 Cyanobacterial Bloom. ISME J 12(9), 2263-2277. doi:10.1038/s41396-018-0159-0

652 Yan, Q., Li, J., Yu, Y., Wang, J., He, Z., Van Nostrand, J.D., Kempher, M.L., Wu, L., Wang, Y.,
653 Liao, L., Li, X., Wu, S., Ni, J., Wang, C. and Zhou, J. (2016) Environmental Filtering Decreases
654 with Fish Development for the Assembly of Gut Microbiota. Environmental Microbiology
655 18(12), 4739-4754. doi:10.1111/1462-2920.13365

656 Yang, Y., Jiang, X.T., Chai, B.L., Ma, L.P., Li, B., Zhang, A.N., Cole, J.R., Tiedje, J.M. and
657 Zhang, T. (2016) Args-Oap: Online Analysis Pipeline for Antibiotic Resistance Genes
658 Detection from Metagenomic Data Using an Integrated Structured Arg-Database.
659 Bioinformatics 32(15), 2346-2351. doi:10.1093/bioinformatics/btw136

660 Yin, X., Jiang, X.T., Chai, B., Li, L., Yang, Y., Cole, J.R., Tiedje, J.M. and Zhang, T. (2018)
661 Args-Oap V2.0 with an Expanded Sarg Database and Hidden Markov Models for Enhancement
662 Characterization and Quantification of Antibiotic Resistance Genes in Environmental
663 Metagenomes. Bioinformatics 34(13), 2263-2270. doi:10.1093/bioinformatics/bty053

664 Zakharova, Y., Bashenkhaeva, M., Galachyants, Y., Petrova, D., Tomberg, I., Marchenkov, A.,
665 Kopyrina, L. and Likhoshway, Y. (2021) Variability of Microbial Communities in Two Long-
666 Term Ice-Covered Freshwater Lakes in the Subarctic Region of Yakutia, Russia. Microb Ecol.
667 doi:10.1007/s00248-021-01912-7

668 Zhang, Y., Chen, J., Chen, H., Liu, L., Liu, C. and Teng, Y. (2022) An Integrated
669 Multidisciplinary-Based Framework for Characterizing Environmental Risks of Heavy Metals
670 and Their Effects on Antibiotic Resistomes in Agricultural Soils. J Hazard Mater 426, 128113.
671 doi:10.1016/j.jhazmat.2021.128113

672 Zhu, G., Wang, X., Yang, T., Su, J., Qin, Y., Wang, S., Gillings, M., Wang, C., Ju, F., Lan, B.,
673 Liu, C., Li, H., Long, X.E., Wang, X., Jetten, M.S.M., Wang, Z. and Zhu, Y.G. (2021) Air
674 Pollution Could Drive Global Dissemination of Antibiotic Resistance Genes. ISME J 15(1),
675 270-281. doi:10.1038/s41396-020-00780-2

676 Zhu, Y.G., Johnson, T.A., Su, J.Q., Qiao, M., Guo, G.X., Stedtfeld, R.D., Hashsham, S.A. and
677 Tiedje, J.M. (2013) Diverse and Abundant Antibiotic Resistance Genes in Chinese Swine
678 Farms. Proc Natl Acad Sci U S A 110(9), 3435-3440. doi:10.1073/pnas.1222743110

679 Zhu, Y.G., Zhao, Y., Li, B., Huang, C.L., Zhang, S.Y., Yu, S., Chen, Y.S., Zhang, T., Gillings,

680 M.R. and Su, J.Q. (2017) Continental-Scale Pollution of Estuaries with Antibiotic Resistance
681 Genes. Nat Microbiol 2, 16270. doi:10.1038/nmicrobiol.2016.270

682

Declaration of interests

☐The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☒The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Dong Wu reports financial support was provided by National Natural Science Foundation of China. Dong Wu reports financial support was provided by Key Laboratory of Environmental Pollution Monitoring and Disease Control, Ministry of Education. Dong Wu reports financial support was provided by Guangdong-Hong Kong-Macao Joint Laboratory for Contaminants Exposure and Health. Bing Xie, Kai Yang reports financial support was provided by National Key Research and Development Program of China. Bing Xie reports financial support was provided by National Natural Science Foundation of China.

Table 1 Measurement and forward-selection of physiochemical parameters of water samples in Taipu Catchment

#Forward-selected factors	TP[1] (Summer|Winter)		PW[2] (Summer|Winter)		LL[2] (Summer|Winter)		LX[2] (Summer|Winter)		JZi[3] (Summer|Winter)		JZe[3] (Summer|Winter)	
TN (sd)	1.56	1.55	1.81	1.39	1.66	1.37	1.71	1.37	1.17	0.80	1.01	1.36
0.1> *P* > 0.05	(0.03)	(0.20)	(1.01)	(0.15)	(0.21)	(0.29)	(0.16)	(0.3)	(0.28)	(0.17)	(0.16)	(0.11)
Nitrate-N (sd)	0.67	0.63	0.73	0.62	0.57	0.50	0.75	0.62	0.58	0.22	0.25	0.12
0.1 > *P* > 0.05	(0.11)	(0.06)	(0.60)	(0.08)	(0.09)	(0.14)	(0.19)	(0.11)	(0.27)	(0.15)	(0.15)	(0.08)
Ammonium-N (sd)	1.01	0.98	1.10	0.89	0.87	0.91	1.11	0.58	0.51	0.77	0.71	1.21
0.1> *P* > 0.05	(0.13)	(0.10)	(0.71)	(0.25)	(0.11)	(0.21)	(0.06)	(0.18)	(0.31)	(0.23)	(0.14)	(0.10)
DO (sd)	6.3	5.5	7.2	4.81	6.21	5.03	5.40	4.03	6.41	5.05	6.30	5.82
P = 0.01; AIC = 201.6	(0.90)	(0.70)	(1.82)	(1.59)	(2.03)	(1.06)	(0.55)	(1.85)	(1.83)	(0.44)	(1.09)	(1.17)
Temperature (sd)	22.5	14.6	21.4	13.3	21.8	12.6	23.3	11.6	20.7	10.8	21.7	11.7
P = 0.04; AIC = 201.7	(5.20)	(2.47)	(1.82)	(5.68)	(1.78)	(5.73)	(2.02)	(5.89)	(1.50)	(5.53)	(1.71)	(5.97)
*FCs (sd)	6.23	5.20	2.05	7.12	2.57	2.07	2.36	10.0	5.00	6.35	5.70	6.25
0.1> *P* > 0.05	(4.43)	(3.56)	(1.71)	(7.05)	(1.97)	(1.48)	(1.87)	(8.52)	(2.49)	(1.71)	(3.76)	(2.55)
Flowrate (sd)	112.85	81.80	154.28	149.06	236.20	230.33	289.73	280.58	123.97	148.36	216.92	158.61
P = 0.05; AIC = 204.5	(18.7)	(1.70)	(35.1)	(22.7)	(58.6)	(13.1)	(13.3)	(47.6)	(61.2)	(26.5)	(33.6)	(64.4)

[1] TP water was sampled from the effluents from Taihu Lake pumping stations; [2] PW, LL, and LX river were sampled from the middle of Taipu River;

[3] JZ water were sampled from Jinze Reservoir; *FCs were presented in the unit of colony forming unit (CFU)/mL-water.

†pH of Taipu River samples kept at 6.95 – 7.03 during the whole period of study.

#Parameters having a *P* value lower than 0.1 in the rda forward section to explain the variations of resistome were listed, and the AIC index were provided along the with significant variables

Figure1

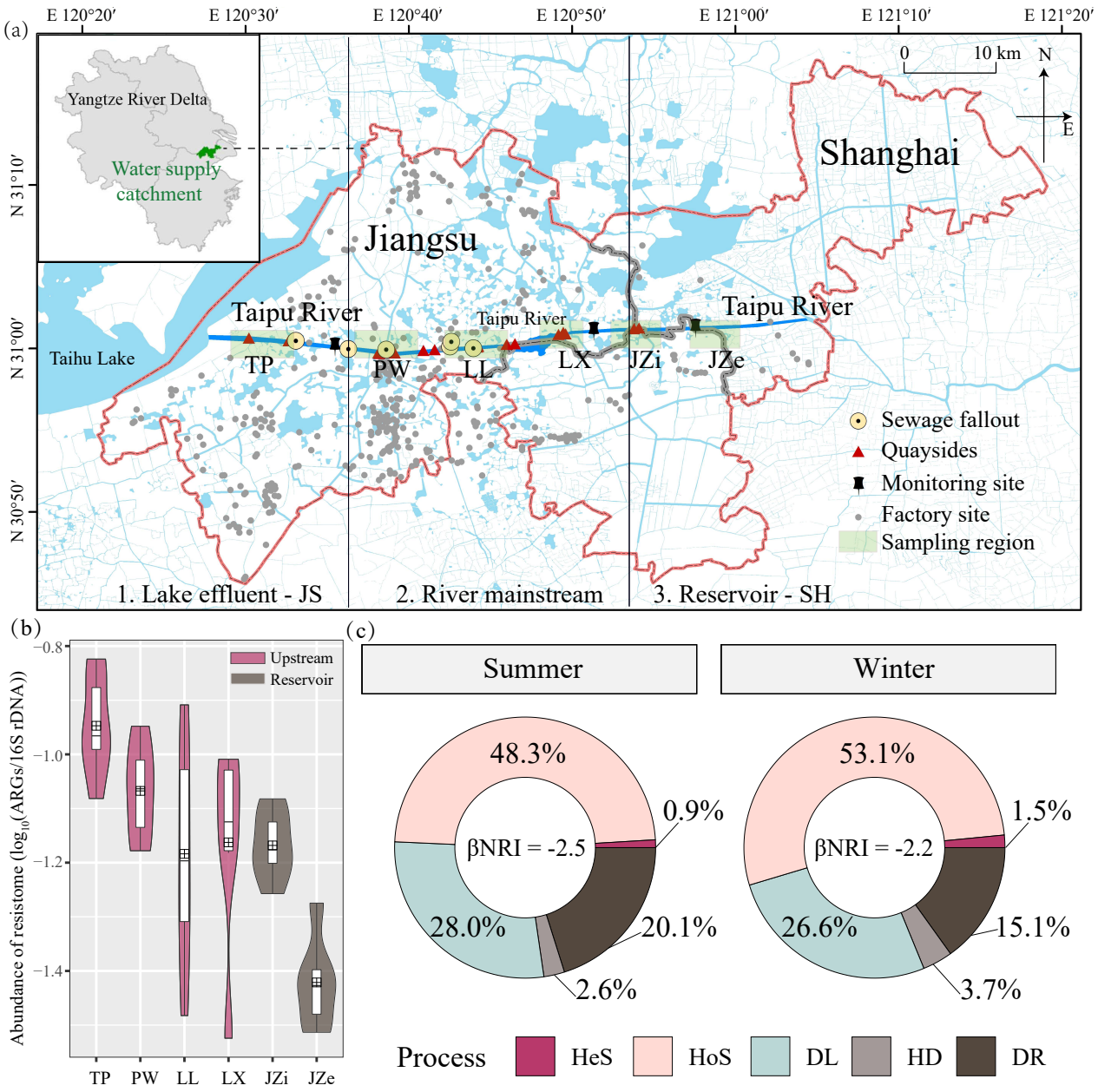


Fig. 1 (a) Taipu River functions as a canal that connected Taihu Lake in Jiangsu Province and Jinze Reservoir in Shanghai. The flow rates of the catchment were managed by the pumping stations on the Taipu River. **(b)** Samples were collected from six regions along the catchment (in green shades) including Taihu Lake pumping stations (TP), Pingwang Station (PW), Lili Station (LL), Luxu Station (LX), Jinze Reservoir influent (JRi), and Jinze Reservoir effluent (JRe). The relative abundance of resistome in each site was averaged during the whole period of study, and the boxplots denoted in the same color were statistically the same. **(c)** the iCAMP-estimated relative importance of different microbial ecological assembly processes and the homogeneous selection (HoS) had predominant influences over heterogeneous selection (HeS), dispersal limitation (DL), homogenizing dispersal (HD) and drifting (DR), which suggested strong environmental filtering effects on the deterministic bacterial assembly process (|βNRI| > 2.0).

Figure2

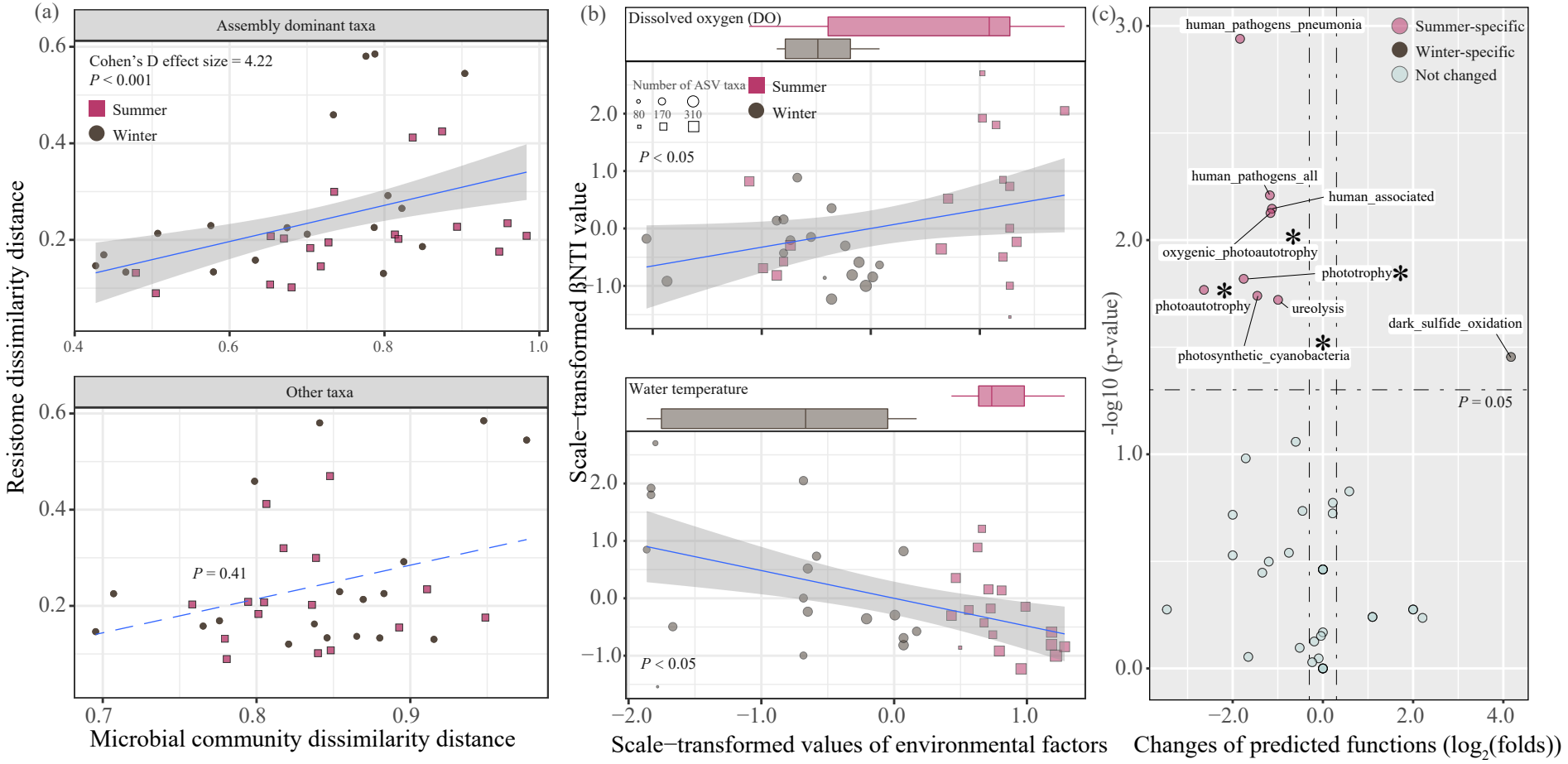


Fig. 2 (a) Correlations between bacterial community structures (assembly - dominant taxa vs. others) and compositions of antibiotic resistomes in each sampling site along the catchment (Bray-Curtis transformed distance). **(b)** Linear regression between the identified environmental factors (DO and temperature) and beta nearest-taxon-index (βNTI). (c) Microbial metabolic functions predicted by FAPROTAX. Pathways had significantly higher intensity in summer and winter were depicted in red and grey (Wilcoxon rank-sum test, $P < 0.05$), respectively. The pathways detected as seasonal variation representative ones using lefse analyses (LDA = 3.0) were annotated with an asterisk.

Figure3

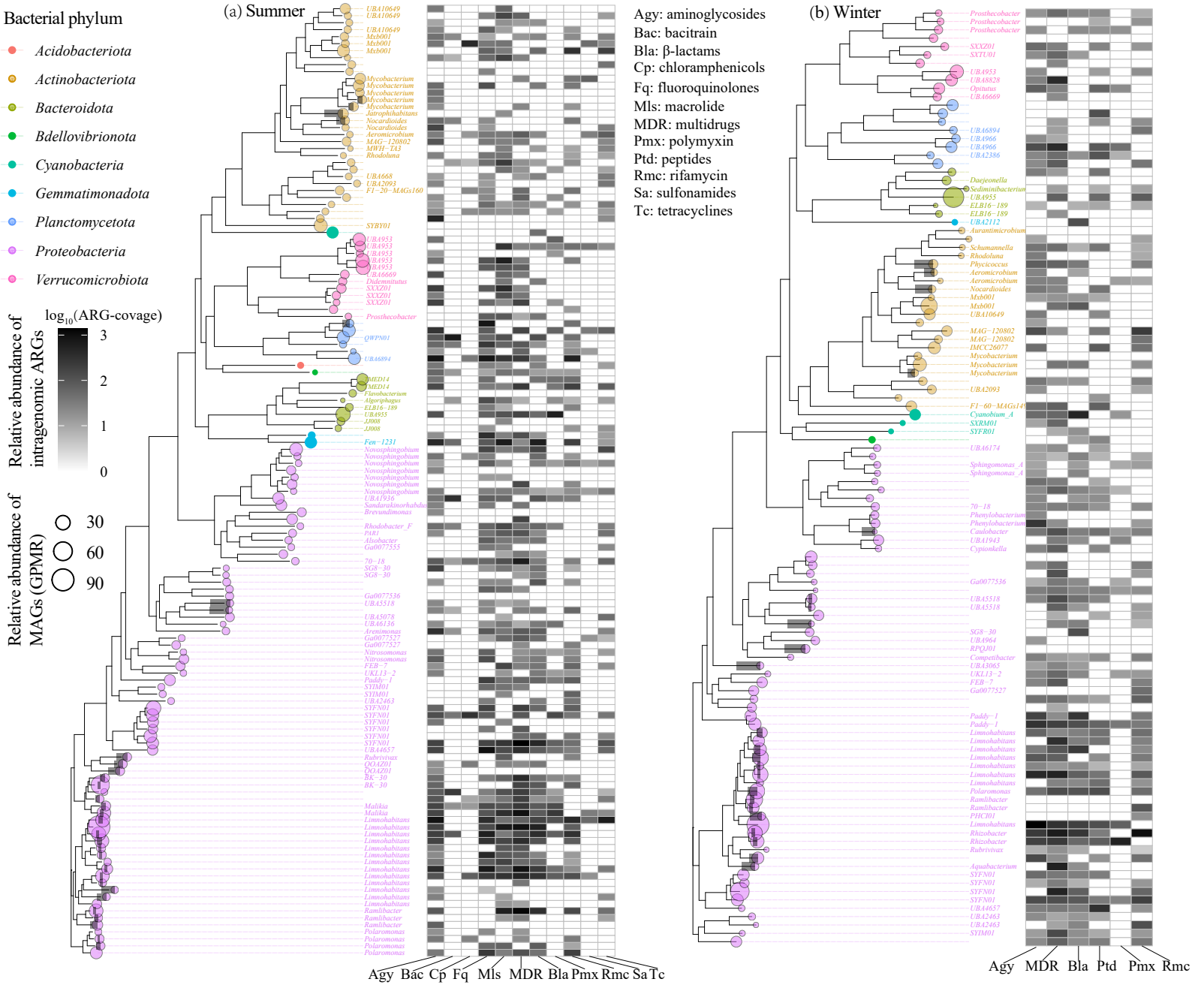


Fig. 3 Distribution and phylogenic trees of the assembled ARGs-hosting metagenome-assembled genomes (MAGs) in summer **(a)** and winter **(b)**. The MAGs of taxa belonging to the same phylum were denoted in the same color. Among them, genomes that were identified as putative resistant pathogens were depicted in grey shades. The relative abundances of the MAGs (GPMR) are shown in proportion to the size of the nodes in the phylogenic trees. The relative abundance of intra-genomic ARGs was shown in heatmaps.

Figure4

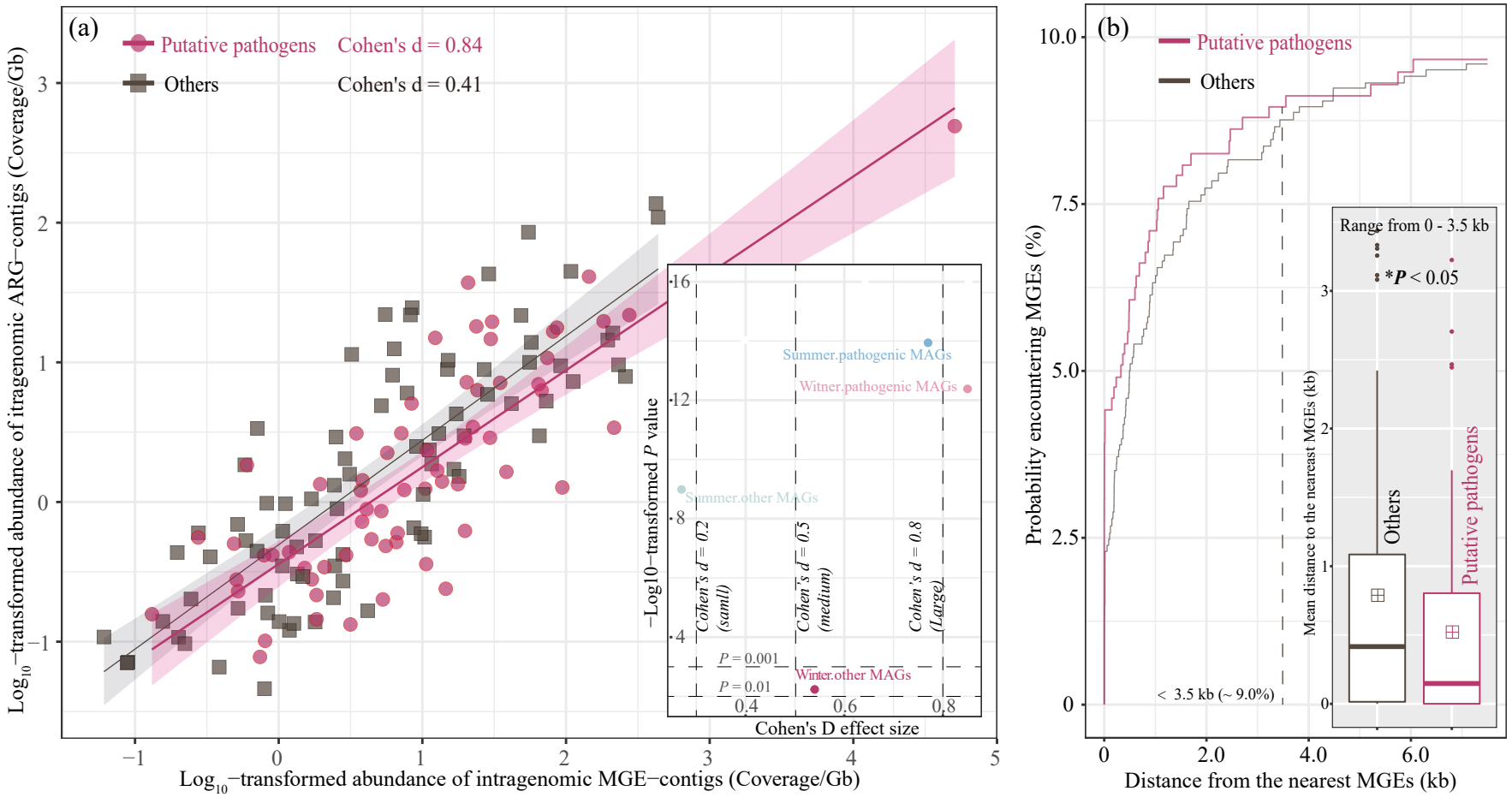


Fig. 4 Linear regression between the abundance of intragenomic MGEs and ARGs in summer and winter samples **(a)**; the Cohen's D effect sizes representing the closeness of regressions between subgroups intragenomic ARGs and MGEs were plotted in the subpanel. **(b)** The metagenome-assembled genomes (MAGs) representing human virulent putative pathogens and others were denoted in red and grey, respectively. The probability of the horizontal gene transfer (HGT) of ARGs across the MAG-assemblages was represented by the incidence of encountering the nearest MGEs. The MAGs of putative pathogens exhibited significantly greater potential HGT of ARGs relative to other genomes within the nearest (ARG - MGE) matching-pair distance of 3.5 kb.

Figure5

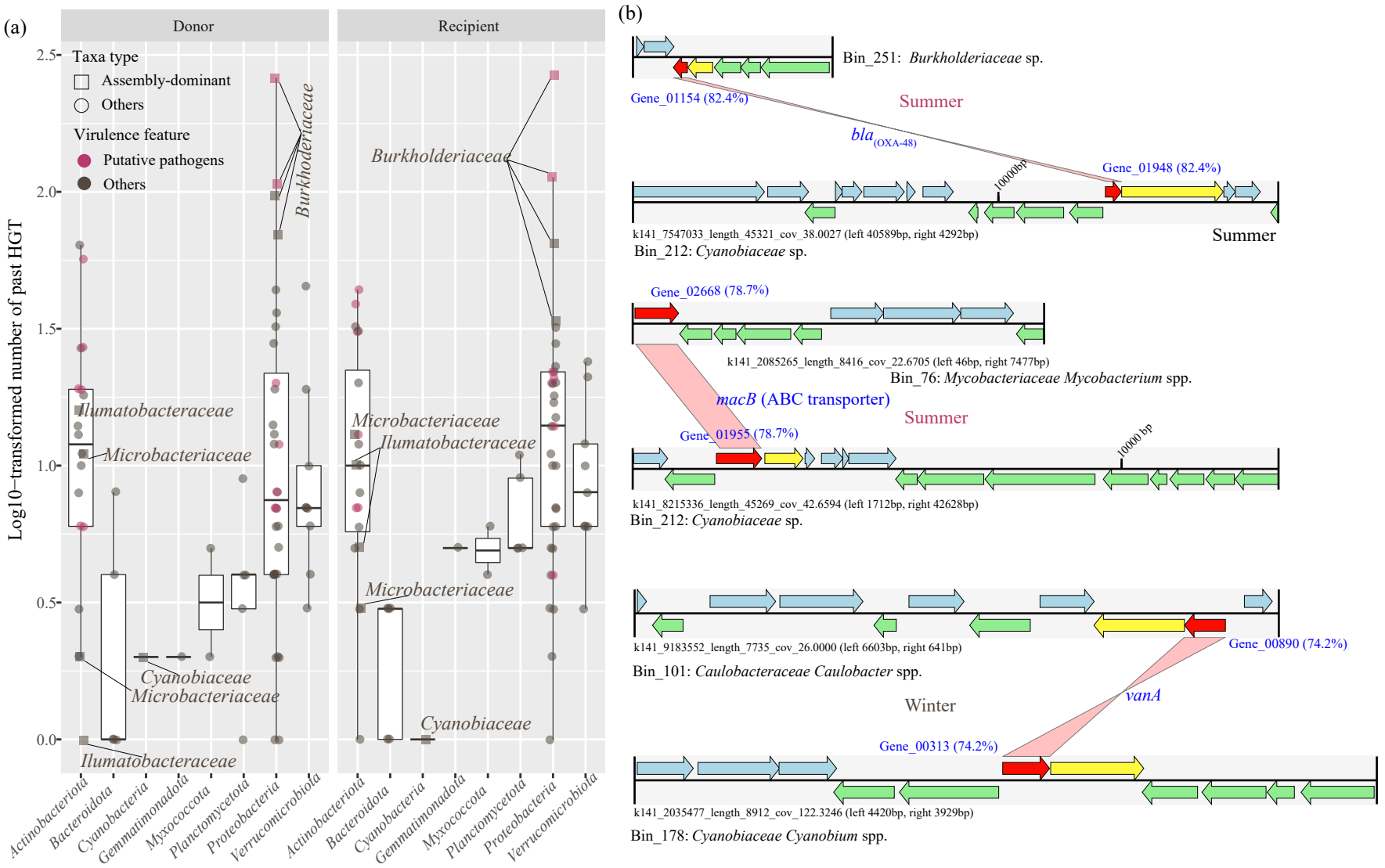


Fig. 5 (a) Community-wide horizontal gene transfer (HGT) of ARGs was analyzed using MetaCHIP. Metagenome-assembled genomes (MAGs) functioning as donors and recipients of the disseminated ARGs were identified and their corresponding count numbers of HGTs were provided. The putative pathogens were highlighted in red and the MAGs belonging to the identified microbial assembly-dominant taxa were depicted in squares. **(b)** The representative contexts of the genetic structure concerning the intragenomic 'ARG - MGE' matching pairs were observed in both summer and winter samples. The identity of matching fragments and gene ID were annotated. The genetic open reading frames (ORFs) segments identified as ARGs (red) and MGEs (yellow) were highlighted, while the other genes located in the forward and reverse strands were colored in blue and green, respectively.

Figure6

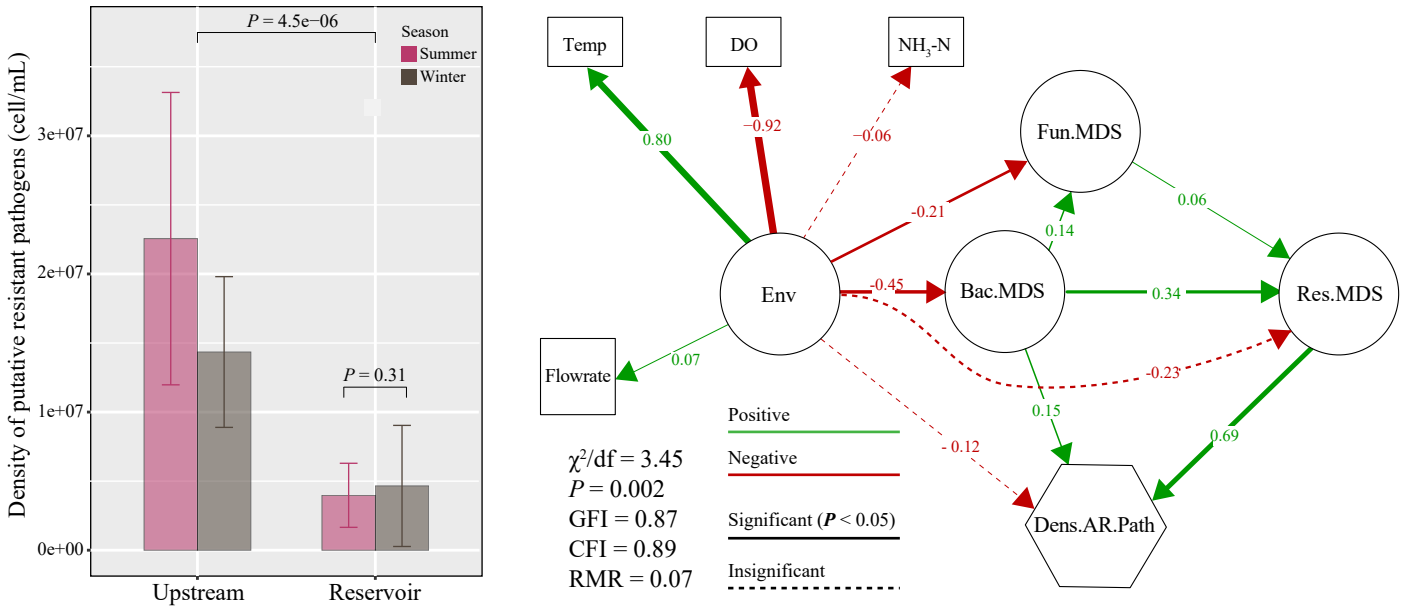


Fig. 6 (a) Mean of cell densities of putative resistant pathogens in the upstream and drinking water reservoir samples collected from the Taipu River Catchment. **(b)** Structural equation modeling (SEM) analysis of the impacts of forward-selected environmental factors (Env), structure (Bray-Curtis multidimensional scaling components) of microbial assembly - dominant taxa (Bac.MDS) and antibiotic resistome (Res.MDS) on the variations of cell densities of waterborne putative resistant pathogens (Dens.AR.Path) in catchment. In the constructed SEM, the co-variances among Env, predicted microbial metabolic functions (Fun.MDS), Bca.MDS and Res.MDS were considered. The relationships fitting a significant correlation ($P < 0.05$) either in positive (green) or negative (red) is depicted using a solid line. The width of lines are in proportion to the standard path coefficients (std-coeff β).