Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Class 1 Integrons and Carbapenemase Genes Distinguish Hospital Effluent Across Urban Wastewater Sources

Ishaq Khan
University of Swat
Ihtisham Naeem
University of Swat
Syed Shujait Ali
University of Tabuk
Murad Ali Rahat
University of Swat
Arshad Iqbal
University of Swat
Shafee Ur Rehman
University of South Asia
Muhammad Shafiq
drshafiq@stu.edu.cn

Shantou University Medical College

Article

Keywords: antimicrobial resistance, wastewater metagenomics, class 1 integrons, carbapenemase genes, hospital effluent, mobilome, One Health

Posted Date: November 16th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11388375/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

In low- and middle-income countries, untreated wastewater from hospital, slaughterhouse, and community sources discharges into open drainage channels. However, how antibiotic resistance genes (ARGs) and mobile genetic elements (MGEs, collectively the mobilome) partition across these distinct urban sources remains poorly resolved.

Methods

We performed shotgun metagenomic sequencing of six wastewater profiles from hospital, slaughterhouse, and community drainage channels in Peshawar, Pakistan across two sampling intervals 30 days apart. ARGs and mobilome elements were quantified using taxonomically stratified bioinformatic workflows.

Results

We identified 814 ARGs across 25 drug classes, dominated everywhere by beta-lactams (74.4%). Total ARG richness did not distinguish the three sources, but mobilome architecture differed markedly. Class 1 integrons (*intI1*, *qacEΔ1*, *tniA*, *tniB*) accounted for 25.6% of mobilome abundance in hospital effluent versus 15.0–16.5% elsewhere (4.4- to 9.7-fold higher copies per cell; permutation $p = 0.067$). Critical carbapenemase (*blaNDM*, *blaVIM*) and glycopeptide determinants (*vanH*, *vanR*) occurred exclusively in hospital effluent. Across the dataset, 44 bacterial genera were associated with both ARGs and MGEs. Slaughterhouse wastewater carried 12.6-fold higher *Acinetobacter* abundance and highest *blaOXA* burden. Total ARG and mobilome burdens showed no correlation (Spearman $\rho = 0.429$, $p = 0.419$). An episodic *Vibrio* bloom showed high resistance with minimal mobilome capacity.

Conclusions

Resistance burden and mobilisation potential were uncoupled across urban wastewater sources. Hospital effluent concentrates critical carbapenemase genes and class 1 integron machinery in this urban drainage network.

Introduction

In 2019, drug-resistant bacterial infections caused an estimated 1.27 million deaths worldwide [1]. Resistance was associated with a further 4.95 million deaths. Six pathogens caused most fatal cases: *Escherichia coli*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Streptococcus pneumoniae*, and *Pseudomonas aeruginosa*. Mortality is projected to reach 1.91 million annual deaths by 2050 [2]. South Asia faces the heaviest burden. In this pathogen group, carbapenem-resistant Gram-negative bacteria are listed in the critical category of the WHO Bacterial Priority Pathogens List [3].

Untreated urban wastewater brings together resistance determinants from human and animal sources. A metagenomic analysis of sewage from 79 sites in 60 countries found that resistomes separated clearly between high- and low-income regions [4]. Slaughterhouse wastewater is a separate route. Veterinary antibiotics select for resistant Enterobacteriaceae in livestock. Abattoir drains in Pakistan yield *blaNDM*-positive isolates that reach meat distribution networks [5]. Community drains receive hospital and slaughterhouse discharge together with ambient environmental bacteria.

Several conditions increase AMR pressure in Pakistan. *Klebsiella pneumoniae* isolates carrying *blaNDM-1* occur in Pakistani animal facilities [5]. NDM carbapenemases remain endemic in South Asia [6]. In regional tertiary hospitals, prescription audits record extensive use of Watch and Reserve antibiotics [7, 8]. Yet metagenomic sewage data from Khyber Pakhtunkhwa are limited. Previous work in Peshawar relied on culture plates and PCR assays [9]. A livestock manure survey in the province sequenced metagenomes but omitted quantitative resistome analyses [10], and our pilot in Mardan examined community drains alone [11]. No study has compared resistome and mobilome profiles across connected hospital, slaughterhouse, and community drains in Peshawar.

Here, we profiled the resistome and mobilome of hospital, slaughterhouse, and community wastewater in Peshawar across two sampling rounds 30 days apart. Sequencing yielded 21.2 to 29.2 million read pairs per sample. Using CARD v3.2.7 and a curated mobile element catalogue, we quantified ARG richness, resistance mechanisms, drug classes, and mobile genetic element families. We also tracked bacterial genera co-carrying resistance genes and mobile platforms. The results provide a metagenomic baseline for resistance burden and mobilisation capacity in an untreated urban drainage network.

Materials and Methods

Study Area and Site Selection

Field sampling targeted three drainage junctions within the Peshawar urban catchment in Khyber Pakhtunkhwa, Pakistan (reference coordinates 34.0150° N, 71.5249° E), encompassing one medical, one abattoir, and one residential municipal site. Hospital effluent (HW1, HW2) was collected directly from the primary outfall of a major tertiary-care hospital. Slaughterhouse (SLW1, SLW2) and community (CW1, CW2) samples came from open municipal drainage canals in the same drainage network. No wastewater treatment infrastructure operated at any location. Raw effluent flowed directly into surface channels. Collection was repeated 30 days later (T1, August 2024; T2, September 2024). The paired collections give two temporal observations at each outfall. They are not spatial replicates.

Sample Collection and Biomass Concentration

Grab samples (500 mL) were taken at each outfall in sterile polypropylene bottles, kept on ice and brought to the laboratory within 6 h. Samples were stored at 4°C before filtration. Biomass was collected by vacuum filtration through 0.22 µm polyethersulfone membranes (Millipore, USA). The retained biomass was weighed and stored at − 80°C until extraction (Supplementary Material).

DNA Extraction and Quality Assurance

DNA was extracted from the filters with the PureLink Microbiome DNA Purification Kit (Invitrogen, Thermo Fisher Scientific, USA). Cells were lysed by mechanical bead-beating combined with enzymatic digestion, and DNA was purified on silica columns. Yield and purity were checked by NanoDrop spectrophotometry. A260/A280 ratios were 1.8–2.0 and A260/A230 ratios were 2.0–2.2. Integrity was verified on 1% agarose gels, and only intact preparations (> 10 kb) were sequenced.

Metagenomic Library Preparation and Sequencing

Libraries were prepared with the Illumina TruSeq DNA PCR-Free Library Preparation Kit, which avoids amplification-induced coverage bias. DNA was sonicated to a target insert length of ~ 350 bp (Covaris adaptive focused acoustics). Blunt-end repair, 3' adenylation and indexed adapter ligation followed in sequence. Libraries were sequenced on an Illumina NovaSeq 6000 instrument using 150 bp paired-end chemistry, producing 21.2 to 29.2 million read pairs per metagenome (Supplementary Material: Table S1).

Bioinformatics Pipeline and Host Decontamination

Computational processing was conducted through the Wekemo BioinCloud analytical environment [12]. Complete bioinformatics software tools, versions, and reference databases used throughout this pipeline are catalogued in Supplementary Material: Table S2. Initial sequence quality was audited using FastQC v0.11.9. Read trimming and quality filtering used Trimmomatic v0.39 (ILLUMINACLIP:adapters:2:30:10, SLIDINGWINDOW:4:20, MINLEN:50), preserving reads with average Phred scores > Q20. Potential human contamination was filtered out with KneadData v0.10.0 by aligning against the human reference genome GRCh38 using Bowtie2 v2.4.5 in very-sensitive mode. Only unmapped, host-decontaminated reads were passed to downstream taxonomic and functional workflows.

Metagenomic Assembly and Gene Catalogue Construction

De novo contig assembly from clean reads was performed using MEGAHIT v1.2.9 across multiple k-mer lengths (--k-list 21,29,39,59,79,99,119,141), setting a minimum contig length cutoff of 500 bp (--min-contig-len 500). Assembly quality and fragmentation statistics were tracked via QUAST v5.2.0. Protein-coding genes were called on assembled contigs using Prodigal v2.6.3 under metagenomic settings (-p meta). To remove redundancy across samples, predicted open reading frames were pooled and clustered with CD-HIT v4.8.1 at thresholds of 95% identity and 90% alignment coverage (-G 1 -c 0.95). Read coverage and relative abundance were quantified by mapping clean reads back to the non-redundant catalogue using Salmon v1.9.0 (--validateMappings --meta), generating abundance in Transcripts Per Million (TPM). Nucleotide sequences were translated using EMBOSS transeq. Taxonomic lineages for catalogued genes were determined by DIAMOND v2.0.15 searches against NCBI NR (--evalue 1e-5 --max-target-seqs 50 --id 70), followed by lowest common ancestor consensus parsing in BASTA (--alen 100 --identity 80 --evalue 1e-5 --minimum 3 --best_hit 1 --maj_perc 60).

Taxonomic Profiling

Community taxonomic profiling was performed directly from clean reads using Kraken2 v2.1.2 against an integrated reference database containing complete bacterial genomes from NCBI NT and RefSeq, as well as fungal, archaeal, and viral sequences, using a confidence parameter of 0.2 (--confidence 0.2). Bayesian reallocation of k-mer matches to resolve species-level proportions was executed with Bracken v2.6 (150 bp read length, threshold 10). Taxonomic read counts were evaluated unrarefied to avoid data loss. Ecological diversity metrics, including Shannon and Simpson indices and Bray-Curtis distances, were computed using the R package vegan. Inter-sector compositional shifts were evaluated by PERMANOVA with 999 permutations, complemented by linear discriminant analysis effect size (LEfSe; LDA > 2.0) and post-hoc Bonferroni-corrected Dunn tests.

Resistome Profiling

Antibiotic resistance determinants were identified by aligning translated gene sequences against the Comprehensive Antibiotic Resistance Database (CARD v3.2.7) with DIAMOND v2.0.15. Hits required sequence identity ≥ 80%, query coverage ≥ 70% and an E-value ≤ 1×10^{-5}. Annotations were structured through the Antibiotic Resistance Ontology (ARO) to map genes to biochemical resistance mechanisms and drug categories. Because multidrug determinants in CARD are assigned across several classes simultaneously, the overall drug-class sum (11.98 CPC) is greater than the mechanism total (5.56 CPC).

Mobile Genetic Element Profiling

Mobile elements were catalogued by searching translated proteins against the curated BioinCloud/Majorbio mobile element repository [14] using DIAMOND v2.0.15 with an E-value threshold ≤ 0.001. Detected features were categorised into two hierarchical tiers: broad functional categories at Level 1 (integrases, transposases, insertion sequences, plasmid-associated elements) and specific element families at Level 2 (such as *intI1*, IS26, IS91, *tniA*, and *tniB*). Abundances of mobile elements were normalised to copies per cell (CPC) via the 16S rDNA procedure described above.

Integrated Resistome–Mobilome Analysis

To determine which bacterial taxa were associated with resistance determinants and mobilisation platforms, functional annotations were taxonomically stratified on the Wekemo/Majorbio platform. Reads mapped to CARD (v3.2.7) and the curated MGE database were simultaneously linked to their Kraken2/Bracken taxonomic lineage, generating stratified profile matrices where each record represents a specific gene–taxon combination (e.g., Gene_ID|Taxonomy_Lineage). Genus-level abundances were expressed in Transcripts Per Million (TPM) by summing all reads assigned to a given genus–gene pair. To evaluate resistome–mobilome co-occurrence across bacterial hosts, genus-stratified ARG and MGE matrices were cross-referenced; genera with non-zero abundance for both at least one ARG and at least one MGE were identified across the six samples, defining 44 genera with co-detected resistance and mobilisation determinants (Fig. 5C; Supplementary Material: Table S6).

Pairwise associations between bacterial genera, ARGs, and MGEs were evaluated using Spearman rank correlation (ρ) on $\log_{10}$(TPM + 1)-transformed data. Because asymptotic tests are invalid at n = 6, two-sided exact permutation *p*-values were computed across all 720 possible sample permutations. Benjamini-Hochberg FDR correction was applied within each comparison family. Constrained redundancy analysis (RDA) was performed on Hellinger-transformed ARG and MGE matrices. Between-group variation was tested with PERMANOVA (999 permutations). Temporal variation within sites was calculated as 1 minus the Bray-Curtis distance between T1 and T2. The class 1 integron platform was evaluated as a composite index summing *intI1*, *qacEΔ1*, *tniA*, and *tniB* in CPC, and expressed as a percentage of total sample mobilome. The hospital sector was compared against non-hospital samples using a two-versus-four exact permutation test across all 15 possible sample assignments (attainable one-sided floor $p = 1/15 = 0.067$).

Results

Microbial Community Composition

Composition, Diversity, and Shared Taxa

Bacterial reads accounted for 95.87–99.62% of classified sequences across all six samples. Pseudomonadota was the dominant phylum in community wastewater, whereas Bacteroidota and Bacillota characterised the slaughterhouse profile (Fig. 1A; Supplementary Material: Table S7). Hospital effluent had high proportions of *Thauera* (mean 24.34%), with *Aliarcobacter* reaching 15.71% in HW1. In contrast, *Acinetobacter* occurred at 12.6-fold higher relative abundance in slaughterhouse than hospital effluent (8.08% vs. 0.64%; Dunn adjusted $p = 0.049$). *Bifidobacterium* showed the opposite pattern, with 3.9-fold higher abundance in hospital than slaughterhouse effluent (1.65% vs. 0.42%; Dunn adjusted $p = 0.049$; Supplementary Material: Table S8). Thirty-eight genera differed significantly between sectors after Bonferroni correction (10 community vs. hospital, 3 community vs. slaughterhouse, 25 hospital vs. slaughterhouse). An episodic *Vibrio* bloom reached 71.06% of CW2 reads compared with 0.03% in CW1, dropping Shannon diversity to 3.50 relative to 4.56 in slaughterhouse and 4.10 in hospital wastewater (Supplementary Material: Fig. S1, S4A). At the whole-community level, principal coordinates analysis (PCoA) of Bray-Curtis dissimilarities showed that bacterial composition did not separate significantly by drainage sector (Axis 1 = 42.1%, Axis 2 = 22.4%; PERMANOVA $R^2 = 0.374$, $F = 0.897$, $p = 0.667$; Fig. 1B). Hospital effluent contained 3,677 bacterial species, compared with 4,745 in slaughterhouse and 5,130 in community effluent. Across the dataset, 2,865 species (45.7%) occurred in all three sectors. Shared species partitioned unequally: exclusive overlap between community and slaughterhouse channels encompassed 1,069 species, 3.5-fold higher than the community–hospital overlap (302 species) and 6.0-fold higher than the hospital–slaughterhouse intersection (179 species; Fig. 1C; Supplementary Material: Table S13). LEfSe identified 324 genus-level biomarkers (Supplementary Material: Fig. S4B). Campylobacterota reached 17.11% in HW1, driven by *Aliarcobacter*, while *Campylobacter* itself remained below 0.02%.

Resistome Architecture

ARG Richness and Total Burden

Overall ARG richness showed no clear divergence among the three drainage sources. Alignment against CARD v3.2.7 catalogued 814 distinct ARGs representing 25 antibiotic classes and 6 resistance mechanisms (Supplementary Material: Fig. S6, Table S3). After multiple testing adjustment, Kruskal-Wallis tests revealed no individual gene with significant sector-specific divergence (peak H = 4.800, unadjusted p = 0.091; all FDR-corrected $p \geq 0.577$). Total ARG burden ranged from 0.371 CPC in SLW1 to 1.727 CPC in CW2. This 4.7-fold variation was driven primarily by sampling timepoint. Specifically, CW2 carried 2.3-fold higher ARG abundance than its paired sample, CW1 (0.751 CPC). HW1 contained 710 unique ARGs compared with 493 in CW1 and HW2. Across all six profiles, a core set of 20 genes made up 26.9% of total resistome abundance, headed by *qnrVC* variants, *sul1*, *msrE*, *aph(3")-Ib*, and *mphE* (Supplementary Material: Table S3).

Drug Class Distribution and Resistance Mechanisms

Determinants conferring resistance to beta-lactam antibiotics represented the largest fraction of the resistome across all sites, making up 74.4% of drug-class-attributed abundance (Fig. 2A; Supplementary Material: Table S14). Cephalosporin, penam, and carbapenem genes contributed the largest shares. Hospital wastewater carried the heaviest burden across these top three classes. Because CARD assigns certain multidrug determinants to more than one class, the drug-class denominator (11.98 CPC across six samples) exceeds the mechanism-level total (5.56 CPC). Fluoroquinolone resistance increased 26-fold between the two community timepoints (CW1: 0.024 CPC; CW2: 0.621 CPC). Hospital wastewater carried higher macrolide and aminoglycoside resistance than slaughterhouse effluent. Sulfonamide resistance was evenly distributed, whereas glycopeptide genes occurred almost exclusively in hospital samples. Enzymatic antibiotic inactivation constituted the predominant resistance mechanism at 72.8% of aggregate ARG load, reaching its highest level in HW1 (1.156 CPC; Supplementary Material: Table S9). Target protection contributed 16.1% overall, concentrated in CW2 (0.633 CPC). Efflux (5.7%), target replacement (3.8%), target alteration (1.6%), and reduced permeability (trace) accounted for the remaining resistance burden.

AMR Gene Families and Site-Specific Patterns

Carbapenem and vancomycin resistance genes of major clinical concern occurred almost exclusively in hospital wastewater. Among 129 identified AMR gene families, beta-lactamases predominated, with CMY, TEM, and GES contributing the highest abundances (Fig. 2B; Supplementary Material: Table S4). High-consequence metallo-beta-lactamase genes (NDM and VIM) and vancomycin resistance operon genes (*vanH* and *vanR*) were completely absent from abattoir and community drains, apart from a minimal *vanR* detection in CW1. GES beta-lactamases also predominated in hospital effluent. Other beta-lactamases did not follow this hospital-skewed distribution. OXA determinants were distributed across all three sectors. TEM genes were higher in CW1 and SLW2 than in HW1, indicating independent carriage in community and livestock reservoirs. Conversely, sulfonamide-resistant *sul* genes and aminoglycoside-modifying phosphotransferases *aph(3")* and *aph(6)* occurred ubiquitously across all sampled channels. No individual gene met statistical criteria for sector enrichment under LEfSe (LDA > 2.0; Supplementary Material: Fig. S2)

or Dunn's tests after Bonferroni correction (minimum KW $p = 0.091$). High-LDA genes such as *qnrVC* (4.81) and *aph(3")-Ib* (4.55) reflected high abundance in CW2 rather than consistent sector differentiation.

WHO Priority Pathogen–Resistance Associations

Genus-attributed ARG abundance totalled 3,498.7 TPM across the dataset. *Vibrio* and Salmonella accounted for 85.4% of this total, reflecting the episodic CW2 bloom (Fig. 3A; Supplementary Material: Table S15). Across consistently detected taxa, *Pseudomonas* carried the highest ARG abundance in hospital effluent (HW mean: 42.3 TPM versus 26.6 in community and 19.9 in slaughterhouse). A similar hospital-skewed distribution was evident for *Klebsiella* (HW mean: 4.1 TPM, 7.7-fold higher than community). Abattoir wastewater carried higher pathogen-linked resistance carriage for *Acinetobacter* (SLW mean: 35.9 TPM, a 6.0-fold excess over hospital levels) and *Escherichia* (SLW mean: 61.4 TPM, a 1.8-fold increase over hospital effluent). While total *blaNDM* abundance was greatest in hospital discharge, *Escherichia*-linked *blaNDM* was highest in slaughterhouse wastewater (0.767 TPM compared to 0.311 in hospital and 0.201 in community channels). *Acinetobacter*-associated *blaOXA* determinants accumulated predominantly in slaughterhouse outflow (SLW average: 6.72 TPM vs. 0.83 in community and 0.24 in hospital effluent). Across the full set of 631 genus–ARG pairs, no association exhibited statistically significant between-sector differences after FDR adjustment (minimum adjusted p = 0.794).

Mobile Genetic Element Profile

Mobilome Composition, Integron Content, and Temporal Stability

Hospital wastewater contained the largest and most integron-rich mobilome across the network. Across the six samples, we identified 37 MGE types spanning 1,236 individual element sequences (Supplementary Material: Fig. S5A, Table S5). Compositional heatmap profiling of Level 1 MGE types showed that transposases formed the largest contributor across all samples (72.7% of total abundance), dominated by *tnpA* variants, while IS91-family insertion sequences accounted for 9.6% and integrases, predominantly *intI1*, contributed 7.7% (Fig. 3B). A circos tracking analysis illustrated the distribution and interconnected flow of major MGE classes across the three drainage sectors, detailing shared mobilization platforms alongside hospital-specific enrichments (Fig. 3C). Total mobilome abundance averaged 7.775 CPC in hospital effluent, 1.85-fold higher than community (4.202 CPC) and 3.34-fold higher than slaughterhouse effluent (2.329 CPC). Integron-associated elements showed marked enrichment in hospital wastewater: *intI1* was 5.9-fold higher, *qacEΔ1* 4.4-fold, *tniA* 5.0-fold, *tniB* 5.5-fold, and *IS26* 9.7-fold higher than slaughterhouse means. Differential abundance testing across individual mobile element sequences showed no statistically significant separation between sectors following correction for multiplicity (lowest unadjusted $p = 0.091$; Supplementary Material: Fig. S3, Table S10). Hospital effluent exhibited the highest temporal stability across 30 days (Bray-Curtis similarity 0.911), followed by slaughterhouse (0.856). Community wastewater showed substantial turnover (similarity 0.469). This community shift coincided with sharp declines in IS91-family sequences (6.95-fold), *qacEΔ1* (5.93-fold), and *intI1* (3.84-fold) at the second timepoint.

Integrated Resistome–Mobilome Analysis

Genus-Level ARG Associations

Spearman rank correlation identified associations between bacterial genera and ARG families (Fig. 4A; Supplementary Material: Table S16a). Abundance of *Thauera* showed positive rank correlations with *GES* beta-lactamases (ρ = +0.886, exact p = 0.033), *VEB* beta-lactamases (ρ = +0.943, exact p = 0.017), and sulfonamide *sul* determinants (ρ = +0.886, exact p = 0.033). Conversely, *Thauera* correlated negatively with the *qnr* family (ρ = −0.886, exact p = 0.033), which was abundant in CW2. *Aliarcobacter* showed an inverse relationship with RND efflux pump abundance (ρ = −1.000, exact p = 0.003). *Acidovorax* also tracked positively with *TEM* beta-lactamase genes (ρ = +0.886, exact p = 0.033), whereas *Arcobacter* correlated negatively (ρ = −0.943, exact p = 0.017). *Pseudomonas* correlated negatively with *CMY* beta-lactamases (ρ = −0.943, exact p = 0.017). Despite nominal permutation p-values falling below the 0.05 threshold, no correlation survived Benjamini-Hochberg FDR adjustment across the 15 × 12 testing space (lowest FDR q = 0.46). These correlations represent exploratory rank effect sizes.

Integron Elements and Beta-Lactamase Families

Class 1 integron components correlated strongly with specific beta-lactamase families (Fig. 4B; Supplementary Material: Table S11, Table S16b). Nineteen MGE–ARG pairs reached $|\rho| \geq 0.943$ across the six samples. Components of the integron platform (integrase, *qacEΔ1*, *tniA*, *tniB*, and *ISCR*) correlated positively with *GES* family beta-lactamases (ρ = +0.886 to + 1.000) and *VEB* determinants (ρ = +0.886 to + 1.000). Perfect rank concordance (ρ = +1.000, exact p = 0.003) occurred for integrase–*GES*, *ISCR*–*GES*, and *qacEΔ1*–*VEB* pairs. Insertion sequence *IS26* co-varied positively with OXA beta-lactamases and MFS efflux systems (ρ = +0.943 each). In contrast, *tniA*, *tniB*, and tnp-*ISCR* correlated negatively with *qnr* genes (ρ = −0.943 each), and *Tn916* correlated negatively with RND efflux pumps (ρ = −0.886). None of these rank associations remained statistically significant after FDR adjustment across the 10 × 12 matrix (lowest FDR q = 0.11). However, the elements that co-varied with *GES* and *VEB* were the same four components elevated in hospital effluent.

Sector Structure in Resistome and Mobilome Ordinations

Constrained redundancy analysis (RDA) examined multivariate structuring in the resistome and mobilome. Detrended correspondence analysis supported linear redundancy analysis (gradient lengths 3.21 SD for ARGs, 1.92 SD for MGEs). Wastewater sector accounted for 36.6% of overall resistome variation (PERMANOVA F = 0.865, p = 0.733; Fig. 5A) and 42.4% of mobilome variance (F = 1.106, p = 0.467; Fig. 5B), though neither grouping reached statistical significance under this sample size. In the ARG ordination, hospital samples occupied negative RDA1 space (centroid − 4.15), distinct from community (+ 2.77) and slaughterhouse (+ 1.38). Top-loading vectors projecting toward hospital ordination space included *bla*VIM-1, *bla*NDM-5, *bla*OXA-427, and *bla*SHV-1. *bla*CTX-M-3 and colistin determinants *mcr-3* and *mcr-9* projected in the opposite direction. Sampling time accounted for substantial ordination spread in both profiles (ARG r^2 =

0.96; MGE r^2 = 0.93; permutation p = 0.10 in both), consistent with the marked temporal turnover observed at the community site.

Genus-Level ARG–MGE Co-occurrence Across Sectors

Cross-referencing taxonomically partitioned CARD and mobile element profiles revealed 44 bacterial genera that co-harbored both resistance and mobilisation genes across the drainage network (Fig. 5C; Supplementary Material: Table S6, Table S16c). Because their abundance signals were restricted to the isolated CW2 bloom event, *Vibrio* and *Salmonella* were omitted from correlation evaluations. Sixteen genera reached $\rho \geq 0.8$. *Acinetobacter* and *Thauera* showed the strongest abundance associations (ρ = 1.000 each; exact p = 0.003). *Acinetobacter* dominated slaughterhouse effluent for both ARG carriage (SLW mean: 41.00 TPM) and MGE burden (615.9 TPM in slaughterhouse, 2,112.2 TPM dataset total). *Thauera* predominated in hospital effluent (HW mean: 3.07 TPM for ARGs; 126.9 TPM for MGEs). *Pseudomonas* showed strong coupling among clinical genera (ρ = 0.943; exact p = 0.017), predominating in hospital effluent and ranking second in overall MGE burden. *Escherichia* also showed co-occurrence (ρ = 0.829; exact p = 0.058), predominating in slaughterhouse wastewater. None of these correlations survived FDR correction across the 43 genera (minimum q = 0.06), remaining hypothesis-generating observations.

Class 1 Integron Platforms and Resistome–Mobilome Coupling

We evaluated class 1 integron elements by summing *intI1*, *qacEΔ1*, *tniA*, and *tniB* into a composite index (Table 1; Fig. 6A, B). The hospital sector mean reached 1.982 CPC, 5.20-fold higher than slaughterhouse (0.381 CPC) and 2.87-fold higher than community effluent (0.690 CPC). Expressed as a percentage of total mobilome burden, class 1 integron components accounted for 24.3% and 26.9% in the two hospital samples, compared with 17.9%, 12.2%, 17.8%, and 15.2% in non-hospital samples. Sector means were 25.6% for hospital, 16.5% for slaughterhouse, and 15.0% for community wastewater. Both hospital samples exceeded every non-hospital sample. An exact permutation test across all 15 possible assignments of two samples to one group yielded a one-sided p = 0.067. Because 1/15 = 0.067, this represents the smallest attainable one-sided p-value for a two-versus-four design. The enrichment was disproportionate: total mobilome burden was 3.3-fold higher in hospital than slaughterhouse effluent, whereas integron components were 4.4- to 5.9-fold higher (and *IS26* was 9.7-fold higher; Fig. 6C).

Resistance gene burden and mobilisation capacity were uncoupled across the sampled drains. Total ARG burden and total mobilome burden showed no statistically significant correlation across the six samples (Spearman ρ = 0.429; exact permutation p = 0.419; Fig. 6D). Procrustes analysis of ARG and MGE ordinations showed weak, non-significant congruence (m^2 = 0.744, p = 0.756; Supplementary Material: Fig. S7, Table S12). The ratio of ARG to MGE burden separated CW2 from all other samples (0.797 versus 0.117–0.179 elsewhere). Normalising ARG burden to *intI1* gave a corresponding separation: CW2 reached 16.3 compared with 1.2–3.6 in other samples. At that second community timepoint, ARG burden increased 2.3-fold while total mobilome burden fell 2.88-fold. The *Vibrio*-dominated sample thus combined a heavy resistance load with reduced mobilisation capacity, consistent with clonal expansion rather than broad horizontal dissemination.

Table 1
Class 1 integron elements across the six Peshawar wastewater samples. Abundances in copies per cell (CPC) for the four components of the class 1 integron platform, their composite sum, each sample's total mobile genetic element burden, and the composite expressed as a percentage of that total

Sample	Sector	*intI1*	*qacEΔ1*	*tniA*	*tniB*	Composite	Total MGE	% of mobilome[a]
CW1	Community	0.398	0.341	0.230	0.147	1.116	6.237	17.9
CW2	Community	0.106	0.058	0.072	0.029	0.264	2.168	12.2
HW1	Hospital	0.764	0.534	0.442	0.284	2.025	8.338	24.3
HW2	Hospital	0.688	0.480	0.487	0.285	1.940	7.213	26.9
SLW1	Slaughterhouse	0.119	0.121	0.084	0.047	0.370	2.075	17.8
SLW2	Slaughterhouse	0.126	0.108	0.101	0.057	0.392	2.582	15.2
HW mean	**Hospital**	**0.726**	**0.507**	**0.465**	**0.284**	**1.982**	**7.775**	**25.6**
SLW mean	**Slaughterhouse**	**0.122**	**0.114**	**0.093**	**0.052**	**0.381**	**2.329**	**16.5**
CW mean	**Community**	**0.252**	**0.199**	**0.151**	**0.088**	**0.690**	**4.202**	**15.0**

a Sector values in this column are the mean of the two per-sample percentages rather than the ratio of the sector-mean composite to the sector-mean total; the two differ where the paired samples are dissimilar.

Discussion

In this network, wastewater sectors differed in mobilome architecture and specific bacterial taxa rather than overall ARG richness. Global PERMANOVA and individual feature tests showed no significant sector separation after multiple testing correction. This outcome reflects low statistical power at n = 2 per sector. Nonetheless, the class 1 integron composite index separated hospital effluent from the other drains. Prior surveys similarly report sector-level differences in wastewater resistomes [4, 15–17]. Broader geographic sampling is needed to test how well these local patterns generalise.

Hospital effluent was the only sampled source carrying the critical carbapenemases *blaNDM* and *blaVIM* and the glycopeptide genes *vanH* and *vanR*, determinants that define WHO Critical Priority pathogens. Hospital sewage is a known reservoir for carbapenemase-producing organisms. In hospital sewers across Wales, Silvester et al. [16] recovered several carbapenemase families, including KPC, IMP, VIM, NDM and OXA-48. Rodriguez-Nava et al. [18] observed persistent NDM and VIM genes in hospital wastewater even without matching clinical cases. In Pakistan, Sattar et al. [19] isolated *blaNDM*-5- and *blaOXA*-181-bearing bacteria from wastewater. The restriction of these genes to hospital effluent is consistent with antibiotic selection in hospitals.

Mobilome composition separated hospital from non-hospital effluent most clearly. Class 1 integrons capture and express resistance gene cassettes [20]. A typical platform has an *intI1* integrase, a recombination site, transposition genes (*tniA*/*tniB*) and the disinfectant resistance marker *qacEΔ1*. The *intI1* gene is used as a molecular proxy for anthropogenic pollution [21]. Every component of this platform was elevated in hospital effluent (25.6% of total mobilome burden). Carbapenemase genes and integron elements co-occurred in this discharge. The effluent enters open channels untreated, so downstream waterways are exposed to mobile, clinically relevant resistance determinants.

Slaughterhouse wastewater had a different profile. *Escherichia* and *Acinetobacter* carried their highest ARG burdens at this site, and ARG–MGE co-variation was strong in *Acinetobacter*. Carbapenem-resistant *Acinetobacter baumannii* is a critical international pathogen. In Germany, Savin et al. [22] recovered carbapenem-resistant *Klebsiella* from slaughterhouse and municipal sewer outfalls. Our samples come from a region with limited veterinary antibiotic regulation [23] and no effluent treatment. Even conventional municipal treatment plants do not always remove carbapenemases [24].

Community wastewater changed markedly over 30 days. The Bray-Curtis distance between T1 and T2 at this site was comparable to distances between different sectors. Other sewer networks show similar temporal shifts [25]. Bacterial blooms can skew wastewater profiles. Becsei et al. [26] reported sewage blooms in Europe in which single genera reached ~ 95% relative abundance, similar to the *Vibrio* surge in CW2. We had no negative extraction controls, so a biological bloom cannot be distinguished from contamination. A single grab sample from a municipal channel cannot give a stable baseline.

Weak Procrustes congruence [27] does not exclude physical linkage between resistance genes and mobile elements. Several pairs correlated strongly, including integrase–GES and ISCR–GES ($\rho = 1.000$, exact $p = 0.003$). Metagenomic surveys in other regions also connect integrases, ISCR elements and transposases to sewage ARGs [28–30]. The integrase–GES correlation agrees with molecular evidence for GES cassette integration into class 1 integrons [20].

Taxonomic profiling linked ARGs and MGEs across 44 bacterial genera. *Pseudomonas*, *Acinetobacter* and *Thauera* showed the strongest abundance co-variation. *Pseudomonas* and *Acinetobacter* are naturally competent and undergo environmental gene transfer [31]. *Thauera* correlated with GES, VEB and sul genes ($\rho = 0.886–0.943$). Ecological correlation across bulk metagenomes does not prove intracellular carriage, however, because abundances can co-vary when taxa respond to shared conditions. Xu et al. [32] found that 7.1% to 31.0% of ARGs in hospital wastewater treatment plant (WWTP) metagenomes were flanked by MGEs on assembled contigs. Here, insertion sequence IS26 correlated with OXA beta-lactamases and MFS pumps ($\rho = 0.943$), consistent with integron- and insertion sequence-mediated mobilisation [21, 33]. At $n = 6$, these leads remain preliminary.

This study has design limitations. Each sector had one sampling location, visited twice, so site identity and sector type are confounded. The data describe these drainage channels, not urban wastewater across Pakistan. Ward outbreaks, facility cleaning cycles or slaughter schedules could explain differences between sites. Statistical power at $n = 2$ per sector is very low. The minimum attainable permutation p-value for between-group PERMANOVA is $1/15 = 0.067$, which precludes formal significance at $\alpha = 0.05$. The same

floor applies to the two-versus-four integron permutation test. The four index components (*intI1*, *qacEΔ1*, *tniA*, *tniB*) had nearly identical rank profiles across samples. They form one aggregated signal, not independent lines of evidence. No correlation remained significant at FDR $q < 0.05$. Extraction blanks and negative controls were not included. Contig-level physical linkage was not tested, and short reads cannot distinguish plasmid from chromosomal genes [30]. Sampling across 30 days misses seasonal shifts. Water chemistry was not measured. CARD v3.2.7 detects only catalogued resistance genes.

Future studies should add spatial replication across several hospitals, abattoirs and municipal drains in the province. Metatranscriptomics could show whether co-occurring ARGs and mobilisation elements are expressed in wastewater. Long-read sequencing (Oxford Nanopore or PacBio) is needed to confirm whether hospital carbapenemases are physically integrated into class 1 integron cassettes.

Conclusion

Shotgun metagenomic profiling showed clear differences in mobilome composition among the three drainage sectors. Hospital effluent was the only sampled source carrying the critical carbapenemases *blaNDM* and *blaVIM* and the glycopeptide genes *vanH* and *vanR*. Class 1 integrons predominated in hospital wastewater. Non-hospital drains carried far lower integron loads. Site and sector identity were confounded, so these observations need validation across independent facilities.

Two findings bear on wastewater surveillance in this region. First, resistance gene burden and mobilisation capacity were uncoupled across drains. Gene abundance alone therefore does not capture mobilisation potential. Second, hospital effluent was stable over 30 days, whereas community wastewater shifted substantially. Surveillance programmes will therefore need repeated sampling rather than single grab collections.

Treating hospital effluent before discharge would remove carbapenemases and class 1 integrons at their primary source. Point-source intervention would interrupt transmission. In parallel, monitoring abattoir drains tracks livestock resistance reservoirs that hospital sampling alone will miss.

Declarations

Acknowledgements

Not applicable.

Declaration of Competing 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.

Funding Sources

This work was supported by the National Natural Science Foundation of China for International Young Scientists [grant number 42150410383] and the SUMC Scientific Research Initiation Grant (SRIG) [grant number 009-510858073].

CRediT Author Statement

Ishaq Khan: Conceptualization, Methodology, Formal analysis, Project administration, Supervision, Writing - original draft. Ihtisham Naeem: Formal analysis, Investigation, Visualization, Writing - original draft. Syed Shujait Ali: Validation, Data curation, Writing - review & editing. Murad Ali Rahat and Shafee Ur Rehman: Validation, Writing - review & editing. Arshad Iqbal: Validation, Supervision, Writing - review & editing. Muhammad Shafiq: Software, Resources, Funding acquisition, Supervision, Writing - review & editing. All authors read and approved the final manuscript.

Declaration of Generative AI and AI-assisted Technologies in the Manuscript Preparation Process

During the preparation of this work the authors used GitHub Copilot in order to assist in assembling and formatting initial code and pipeline scripts. After using this tool, the authors reviewed, tested, and verified all analytical code and manuscript content, and take full scholarly responsibility for the publication.

Research Data and Data Availability Statement

The raw sequencing reads generated in this study have been deposited in the NCBI Sequence Read Archive (SRA) under BioProject PRJNA1463247. The data will be made publicly available upon publication of the manuscript. Reviewers and the editorial office may use the link below to access the data during review. Reviewer access link: https://dataview.ncbi.nlm.nih.gov/object/PRJNA1463247?reviewer=q1l9a701ebcdqij0oa5imnoset

All downstream analysis code, taxonomic profiles, and mobilome matrices are freely available at GitHub (https://github.com/ihtishamnaeem36/Hospital-Wastewater-Integrons-and-Carbapenemases) under the MIT License, and are archived at Zenodo (https://doi.org/10.5281/zenodo.22725582).

Sex and Gender-Based Analyses (SAGER)

Sex and gender dimensions are not applicable to this study, as the research was conducted exclusively on environmental wastewater metagenomic samples without human or animal participants.

Ethics Approval and Consent to Participate

Not applicable. This study utilized environmental wastewater drainage samples and did not involve human participants, clinical specimens, or animal experimentation.

References

1. Antimicrobial Resistance Collaborators. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet* **399** (10325), 629–655. https://doi.org/10.1016/S0140-6736(21)02724-0

(2022).

2. Naghavi, M. et al. Global burden of bacterial antimicrobial resistance 1990–2021: a systematic analysis with forecasts to 2050. *Lancet* **404** (10459), 1199–1226. https://doi.org/10.1016/S0140-6736(24)01867-1 (2024).
3. World Health Organization. *WHO bacterial priority pathogens list, 2024: bacterial pathogens of public health importance to guide research, development and strategies to prevent and control antimicrobial resistance* (World Health Organization, 2024).
4. Hendriksen, R. S. et al. Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. *Nat. Commun.* **10** (1), 1124. https://doi.org/10.1038/s41467-019-08853-3 (2019).
5. Chaudhry, T. H. et al. Emergence of blaNDM-1 harboring Klebsiella pneumoniae ST29 and ST11 in veterinary settings and waste of Pakistan. *Infect. Drug Resist.* **13**, 3033–3043. https://doi.org/10.2147/IDR.S248091 (2020).
6. Walsh, T. R., Weeks, J., Livermore, D. M. & Toleman, M. A. Dissemination of NDM-1 positive bacteria in the New Delhi environment and its implications for human health. *Lancet Infect. Dis.* **11** (5), 355–362. https://doi.org/10.1016/S1473-3099(11)70059-7 (2011).
7. Atif, M., Azeem, M., Saqib, A. & Scahill, S. Investigation of antimicrobial use at a tertiary care hospital in Southern Punjab, Pakistan using WHO methodology. *Antimicrob. Resist. Infect. Control*. **6**, 41. https://doi.org/10.1186/s13756-017-0199-7 (2017).
8. Saleem, Z. et al. Pattern of inappropriate antibiotic use among hospitalized patients in Pakistan: a longitudinal surveillance and implications. *Antimicrob. Resist. Infect. Control*. **8**, 146. https://doi.org/10.1186/s13756-019-0649-5 (2019).
9. Shams, K. et al. Highly drug-resistant Escherichia coli from hospital wastewater with several evolutionary mutations: an integrated insights from molecular, computational, and biophysics. *Mol. Biotechnol.* **68** (2), 828–845. https://doi.org/10.1007/s12033-025-01410-y (2025).
10. Zaman, S. et al. Metagenomic profiling reveals extensive bacterial diversity in chicken manure and associated contaminated wastewater. *Int. J. Mol. Sci.* **27** (9), 3741. https://doi.org/10.3390/ijms27093741 (2026).
11. Khan, I. et al. Metagenomic surveillance identifies a high-risk antibiotic resistance profile in community wastewater: a pilot study from Pakistan. *Naunyn Schmiedebergs Arch. Pharmacol.* https://doi.org/10.1007/s00210-026-05471-x (2026).
12. Gao, Y., Zhang, G., Jiang, S. & Liu, Y. X. Wekemo Bioincloud: a user-friendly platform for meta-omics data analyses. iMeta. ;3(1):e175. (2024). https://doi.org/10.1002/imt2.175
13. Yin, X. et al. -OAP v3.0: Antibiotic-Resistance Gene Database Curation and Analysis Pipeline *Optim. Eng.* ;**27**:204–213. https://doi.org/10.1016/j.eng.2022.10.011 (2023).
14. Frost, L. S., Leplae, R., Summers, A. O. & Toussaint, A. Mobile genetic elements: the agents of open source evolution. *Nat. Rev. Microbiol.* **3** (9), 722–732. https://doi.org/10.1038/nrmicro1235 (2005).
15. Pärnänen, K. M. M. et al. Antibiotic resistance in European wastewater treatment plants mirrors the pattern of clinical antibiotic resistance prevalence. *Sci. Adv.* **5** (3), eaau9124. https://doi.org/10.1126/sciadv.aau9124 (2019).

16. Silvester, R. et al. Metagenomic profiling of hospital wastewater: a comprehensive national scale analysis of antimicrobial resistance genes and opportunistic pathogens. *J. Infect.* **90**, 106503. https://doi.org/10.1016/j.jinf.2025.106503 (2025).
17. Acosta, N. et al. Metagenomic analysis after selective culture enrichment of hospital and community wastewater enhances antimicrobial resistance gene detection. *mBio* **16** (7), e01672–e01625. https://doi.org/10.1128/mbio.01672-25 (2025).
18. Rodriguez-Nava, G. et al. Implementation of wastewater-based epidemiology in a Northern California acute care hospital: a pilot study, June–July 2024. *J. Hosp. Infect.* **172**, 8–17. https://doi.org/10.1016/j.jhin.2026.03.007 (2026).
19. Sattar, F. et al. Analyzing antibiotic resistance in bacteria from wastewater in Pakistan using whole-genome sequencing. *Antibiotics* **13** (10), 937. https://doi.org/10.3390/antibiotics13100937 (2024).
20. Gillings, M. R. Integrons: past, present, and future. *Microbiol. Mol. Biol. Rev.* **78** (2), 257–277. https://doi.org/10.1128/MMBR.00056-13 (2014).
21. Gillings, M. R. et al. Using the class 1 integron-integrase gene as a proxy for anthropogenic pollution. *ISME J.* **9** (6), 1269–1279. https://doi.org/10.1038/ismej.2014.226 (2015).
22. Savin, M. et al. Genetic characterisation of carbapenem-resistant Klebsiella spp. from municipal and slaughterhouse wastewater. *Antibiotics* **11** (4), 435. https://doi.org/10.3390/antibiotics11040435 (2022).
23. Van Boeckel, T. P. et al. Global trends in antimicrobial resistance in animals in low- and middle-income countries. *Science* **365** (6459), eaaw1944. https://doi.org/10.1126/science.aaw1944 (2019).
24. Makowska, N., Koczura, R. & Mokracka, J. Metagenomic analysis of beta-lactamase and carbapenemase genes in the wastewater resistome. *Water Res.* **170**, 115277. https://doi.org/10.1016/j.watres.2019.115277 (2019).
25. Fierer, N. et al. A metagenomic investigation of spatial and temporal changes in sewage microbiomes across a university campus. *mSystems* **7** (5), e00651–e00622. https://doi.org/10.1128/msystems.00651-22 (2022).
26. 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. https://doi.org/10.1038/s41467-024-51957-8 (2024).
27. Peres-Neto, P. R. & Jackson, D. A. How well do multivariate data sets match? The advantages of a Procrustean superimposition approach over the Mantel test. *Oecologia* **129** (2), 169–178. https://doi.org/10.1007/s004420100720 (2001).
28. Lekunberri, I., Subirats, J., Borrego, C. M. & Balcázar, J. L. Metagenomic exploration reveals a marked change in the river resistome and mobilome after treated wastewater discharges. *Environ. Pollut*. **234**, 538–542. https://doi.org/10.1016/j.envpol.2017.12.001 (2018).
29. Zhang, J. et al. Metagenomics insights into the profiles of antibiotic resistome in combined sewage overflows from reads to metagenome assembly genomes. *J. Hazard. Mater.* **429**, 128277. https://doi.org/10.1016/j.jhazmat.2022.128277 (2022).

30. Liu, S. et al. Close interactions between prokaryotes and plasmids or viruses highlight a pivotal role of horizontal gene transfer in shaping antibiotic/metal(loid) resistome and their prokaryotic supercarriers in untreated hospital sewage. *Water Res.* **286**, 124178. https://doi.org/10.1016/j.watres.2025.124178 (2025).
31. Smillie, C. S. et al. Ecology drives a global network of gene exchange connecting the human microbiome. *Nature* **480** (7376), 241–244. https://doi.org/10.1038/nature10571 (2011).
32. Xu, C. et al. Antibiotic resistance genes risks in relation to host pathogenicity and mobility in a typical hospital wastewater treatment process. *Environ. Res.* **259**, 119554. https://doi.org/10.1016/j.envres.2024.119554 (2024).
33. Larsson, D. G. J. & Flach, C. F. Antibiotic resistance in the environment. *Nat. Rev. Microbiol.* **20** (5), 257–269. https://doi.org/10.1038/s41579-021-00649-x (2022).

Figures

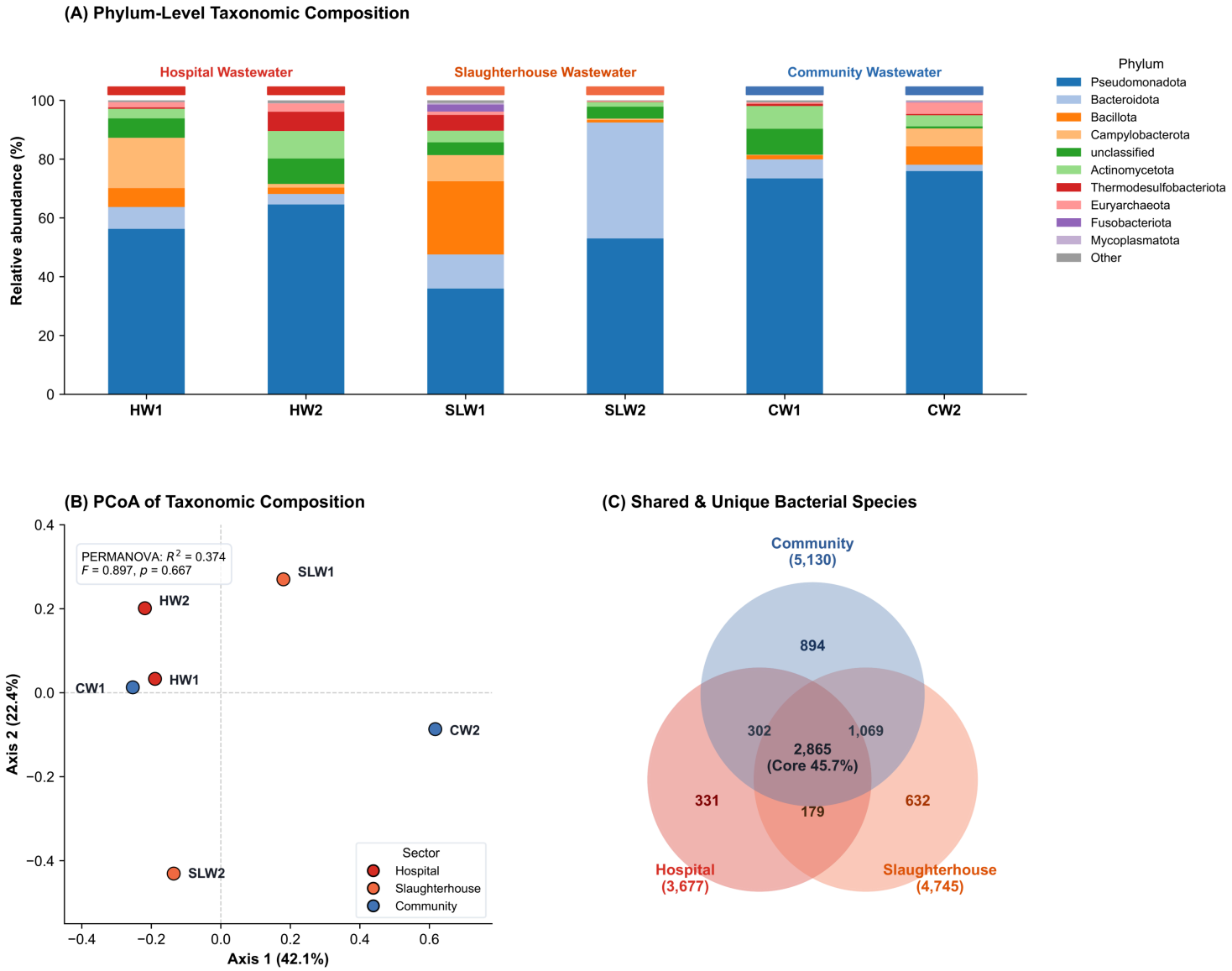


Figure 1

Microbial community composition across hospital (HW), slaughterhouse (SLW), and community (CW) wastewater sites in Peshawar, Pakistan. (A) Phylum-level relative abundance stacked bar chart across all six samples; top phyla colour-coded. (B) Principal coordinates analysis (PCoA) of Bray-Curtis dissimilarities among the bacterial taxa; Axis 1 explains 42.1% and Axis 2 explains 22.4% of total variance; sample points colour-coded by sector, shown without confidence ellipses since n = 2 per sector is too small to estimate a stable ellipse; the PERMANOVA result for sector on the taxonomic profile is given in the panel title (R^2 = 0.374, F = 0.897, p = 0.667). Sector colour keys are panel-specific and are drawn in each panel. (C) Venn/flower diagram showing shared and unique bacterial species across the three wastewater sectors; core microbiome shared by all three sectors = 2,865 species (45.7% of total detected species)

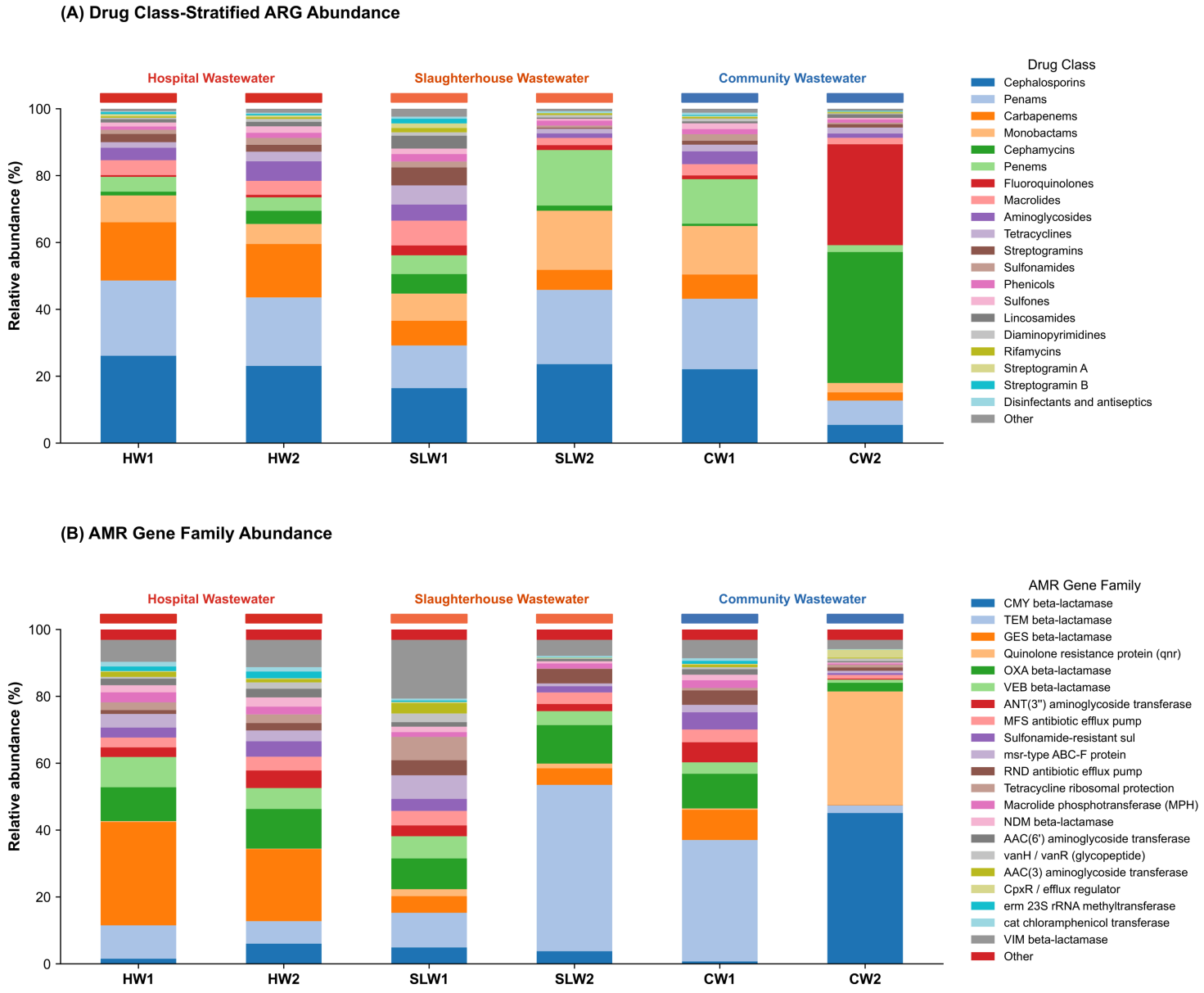


Figure 2

Resistome architecture across the six Peshawar wastewater samples. (A) Drug class-stratified ARG abundance, each bar the percentage composition of that sample's drug-class-assigned abundance; genes

assigned by CARD to more than one class contribute to each, and classes below the display threshold are pooled as "Other". (B) AMR gene family abundance, each bar the percentage composition of that sample's family-assigned abundance; families ranked by total TPM across all samples, the remainder pooled as "Other". Sector colour bars above the panels follow the scheme used in Figs. 5 and 6. The circos plot of the top 30 ARGs by abundance is given in Supplementary Material: Fig. S6

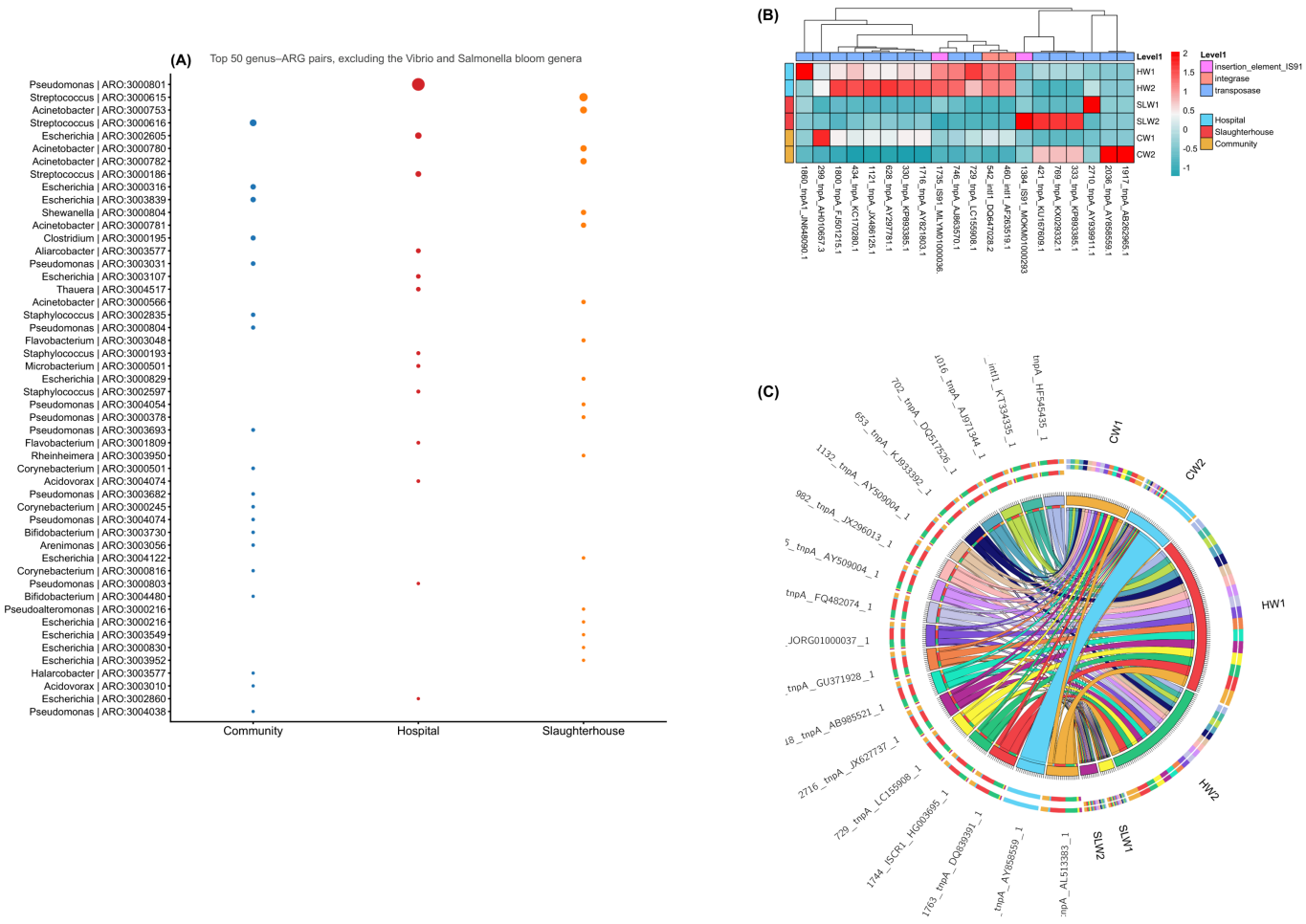


Figure 3

WHO priority pathogen–resistance gene co-occurrence and mobile genetic element (MGE) composition across Peshawar wastewater sectors. (A) Dot plot showing genus–ARG associations for the top 50 taxon–resistance gene pairs after excluding *Vibrio* and Salmonella, both of which are confined to the single CW2 sampling event rather than distributed across the network; dot size indicates relative abundance (TPM); dot colour indicates sector of dominant enrichment. (B) Level 1 MGE type composition heatmap showing transposases, integrases, insertion sequences, and other element types per sample, across the six Peshawar wastewater samples. (C) Circos plot showing MGE distribution and flow across sectors; ribbon thickness proportional to TPM

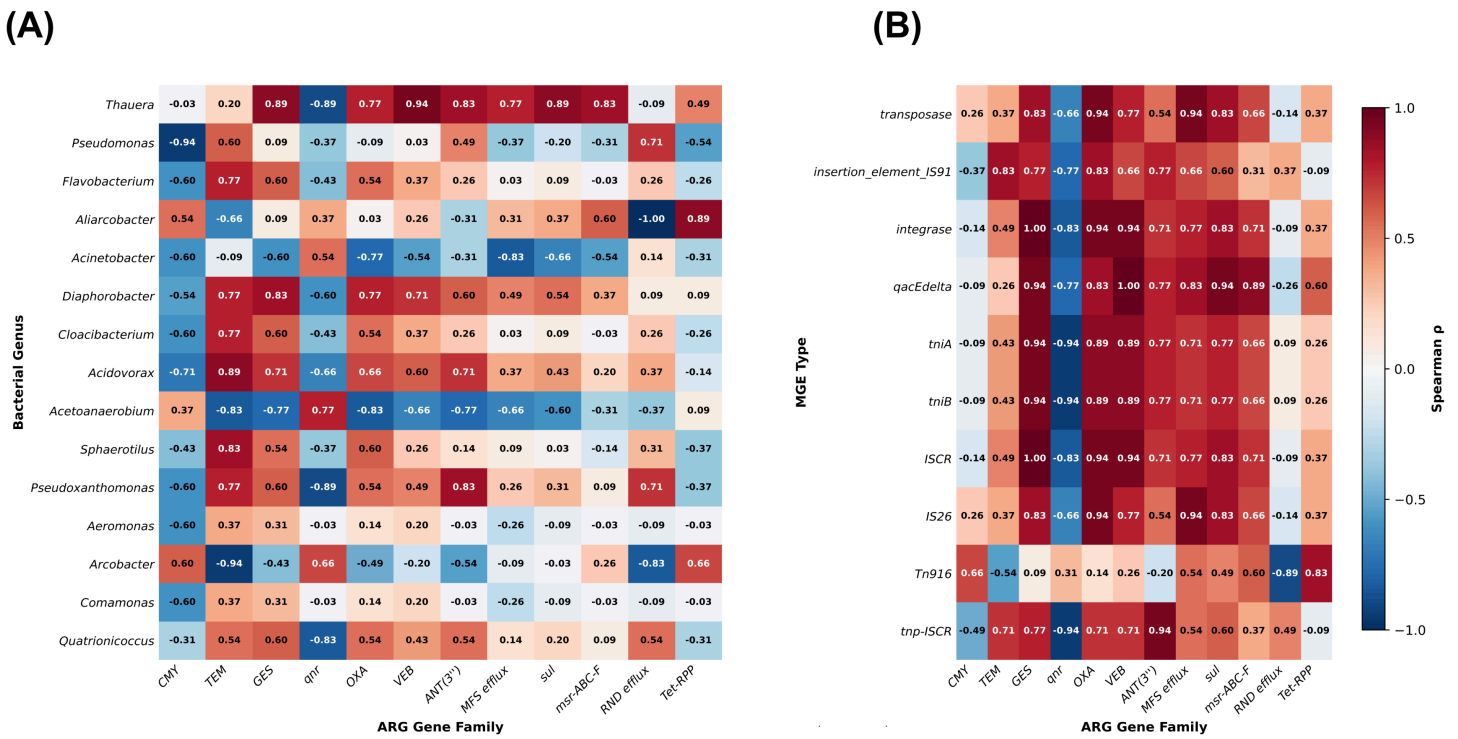


Figure 4

Integrated resistome–mobilome correlation across the six Peshawar wastewater samples. (A) Correlation heatmap between bacterial genera and ARG gene families; cell values shown. (B) Correlation heatmap between MGE types and ARG families; cell values shown. Both panels share the colour scale drawn at the right, representing Spearman ρ from −1 to +1. Rows with identical values across all families (*tniA*/*tniB*, integrase/ISCR, transposase/IS26) have coincident rank profiles across the six samples and are not independent. *qacEΔ1*is labelled *qacEdelta* in panel B. The Procrustes superimposition of ARG and MGE ordinations is given in Supplementary Material: Fig. S7

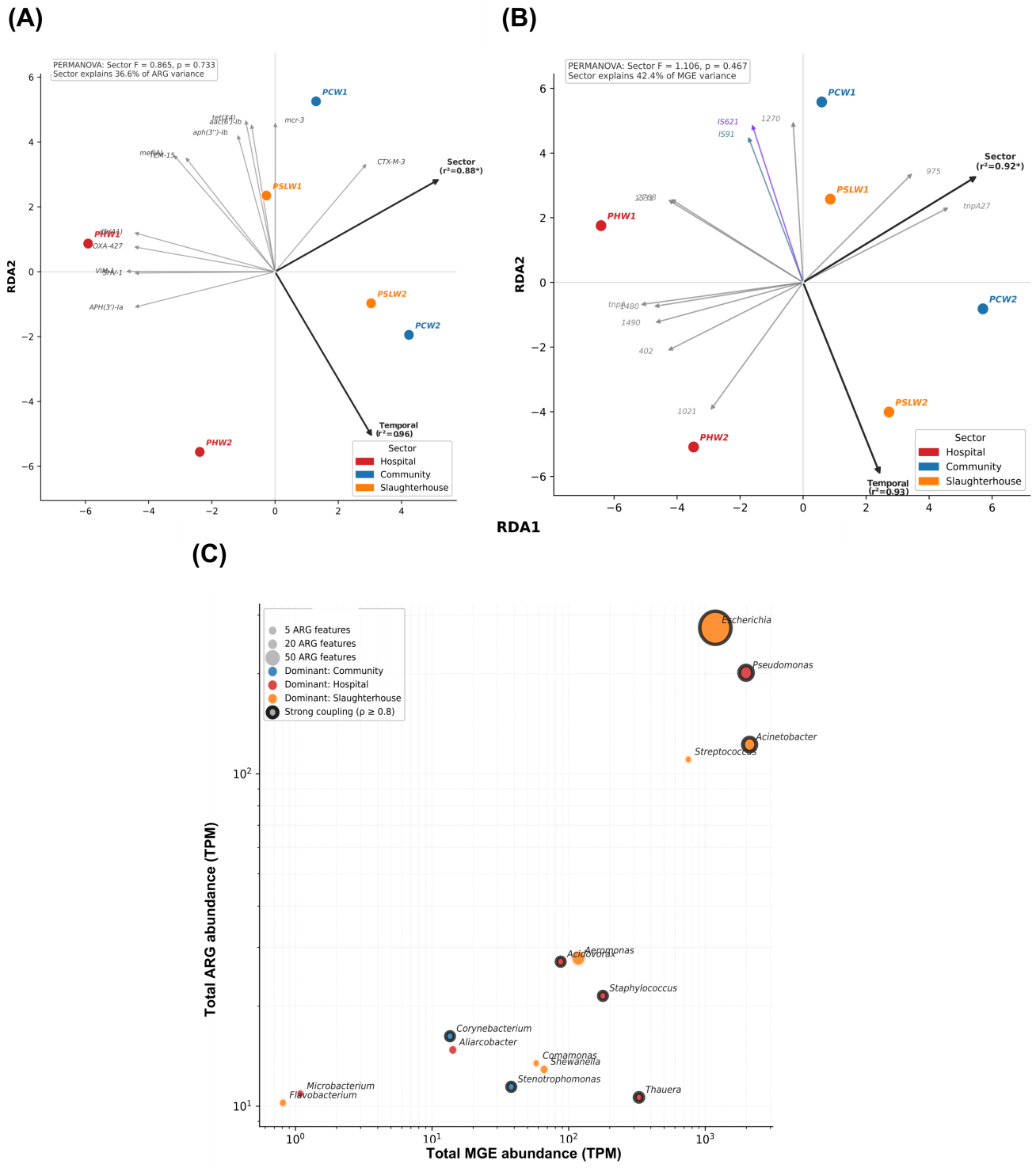


Figure 5

Sector structure of the resistome and mobilome across the six Peshawar wastewater samples. (A) RDA biplot for ARG profiles constrained by sector; grey arrows indicate top-loading ARG features and black arrows the fitted sector and sampling-time gradients (source and sampling time coded as numeric gradients, sector envfit r² = 0.88, p = 0.022, the smallest value attainable for this design, and sampling-time envfit r² = 0.96, p = 0.10); the PERMANOVA result for sector is given in the inset box. (B) RDA biplot for MGE

profiles constrained by sector, same construction (envfit sector $r^2 = 0.92$, time $r^2 = 0.93$); MGE features are labelled by their Level 2 family identifier. Confidence ellipses are not drawn, since they cannot be estimated from two samples per sector. (C) Genus-stratified total ARG abundance versus total MGE abundance (TPM, log-log scale) for genera carrying both; point size reflects the number of ARG features carried; point colour indicates the sector of dominant carriage (community, hospital, slaughterhouse); black-outlined points denote strong ARG–MGE coupling (Spearman $\rho \geq 0.8$)

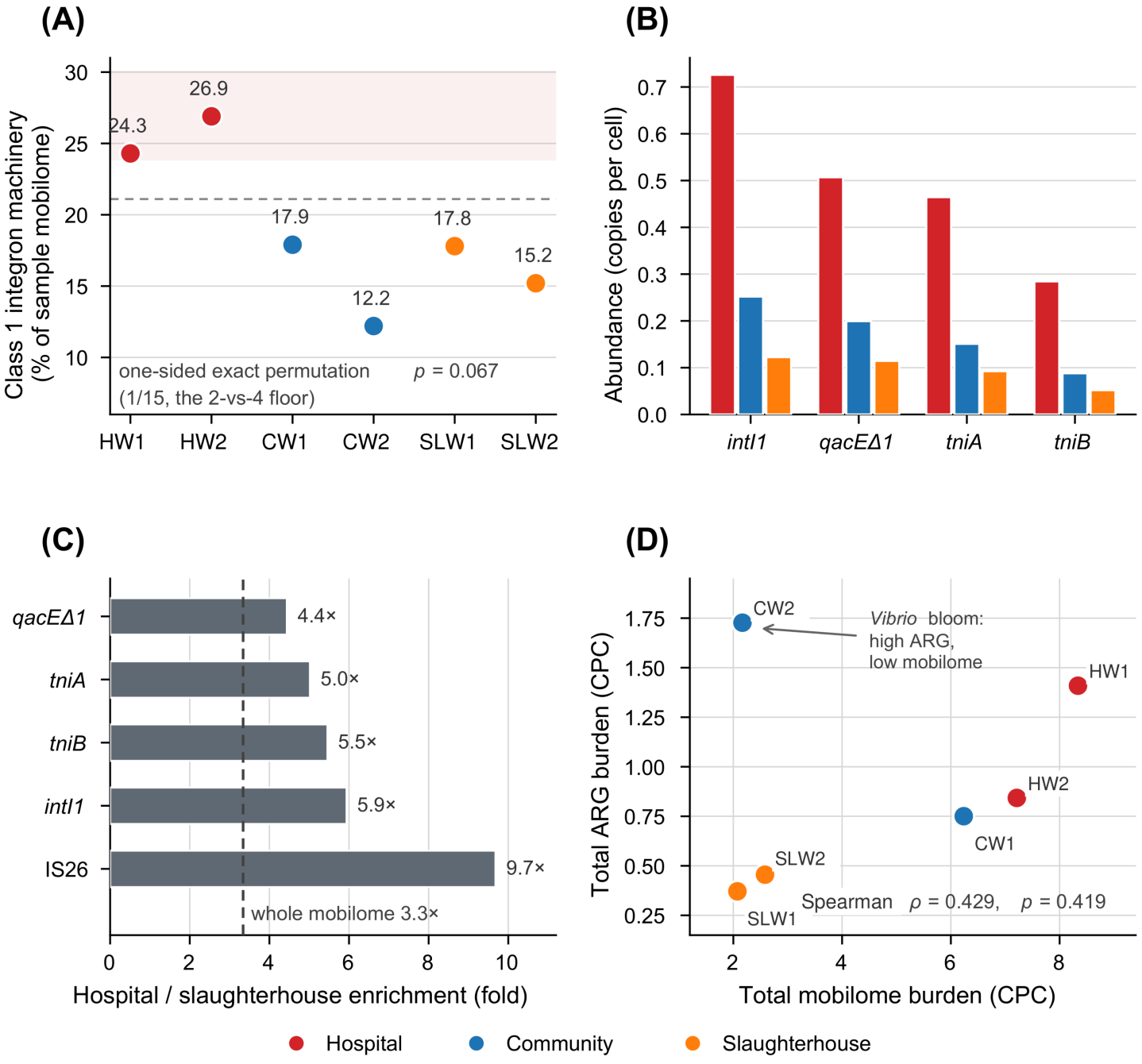


Figure 6

Class 1 integron elements across the six Peshawar wastewater samples. (A) The composite index (*intI1* + *qacEΔ1* + *tniA* + *tniB*) as a percentage of each sample's total mobile genetic element burden; both hospital

samples exceed every non-hospital sample, and the dashed line marks the separation. (B) Sector means of the four components in copies per cell. (C) Hospital/slaughterhouse enrichment for each element, with the dashed line marking the 3.3-fold enrichment of the mobilome as a whole. (D) Total ARG burden against total mobilome burden per sample; the *Vibrio*-dominated CW2 sample carries a high resistance load with a low mobilome burden

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementary.pdf