Science of the Total Environment 1002 (2025) 180607

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Genome-resolved metagenomics uncovers antimicrobial resistance gene carriers in hospital and municipal wastewater environments

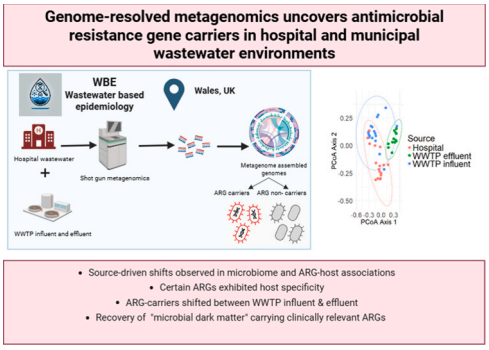
Reshma Silvester [a,c,*], Gordon Webster [b], William B. Perry [b], Kata Farkas [a,c], Laura Rushton [b], Noel Craine [d], Gareth Cross [e], Peter Kille [b], Andrew J. Weightman [b], Davey L. Jones [a,c]

[a] School of Environmental and Natural Sciences, Bangor University, Bangor, Gwynedd, LL57 2UW, United Kingdom
[b] School of Biosciences and Water Research Institute, Cardiff University, Cardiff, CF10 3AX, United Kingdom
[c] Verily Life Sciences LLC., Dallas, Texas, 75019, USA
[d] Public Health Wales, Microbiology Department, Ysbyty Gwynedd, Bangor, LL57 2PW, United Kingdom
[e] Science Evidence Advice Division, Health and Social Services Group, Welsh Government, Cathays Park, Cardiff, CF10 3NQ, United Kingdom

HIGHLIGHTS

- Genome-resolved insights into ARG dissemination in hospital and municipal wastewater.
- Accurate identification of ARG-carriers across complex wastewater environments.
- "Microbial dark matter" identified as reservoirs of clinically relevant ARGs.
- ARG-host dynamics shifted between WWTP influent and effluent.
- High-resolution foundational framework for integrated AMR monitoring.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Antimicrobial Resistance
Wastewater-based epidemiology
Metagenome-assembled genomes
ARG hosts
Public health surveillance
One Health

ABSTRACT

Wastewater-based epidemiology (WBE) is a powerful approach to study antimicrobial resistance (AMR) dynamics at the population level. Using genome-resolved metagenomics, we recovered 3978 metagenome-assembled genomes (MAGs) from archived metagenome sequences generated under the national wastewater surveillance programme across Wales, UK. Taxonomic profiling of MAGs revealed a diverse bacterial community, with significant compositional shifts observed across seasons and sample sources. Approximately 13.6 % of the MAGs carried one or more antimicrobial resistance genes (ARGs), with genes conferring resistance to tetracycline and oxacillin being the most prevalent within the wastewater microbiomes. We also recovered yet-uncultivated microbial genomes- often referred to as "microbial dark matter" harbouring clinically relevant ARGs, offering insights into previously uncharacterised resistance reservoirs in wastewater. ARG-host associations shifted between untreated influent and treated effluent, with effluent profiles also varying significantly between secondary and tertiary treatments, revealing the impact of treatment type on ARG host composition. This study represents the first comprehensive genome-resolved metagenomic characterisation of ARG carriers across both hospital and municipal wastewater in Wales, effectively bridging clinical and environmental compartments. Our findings

* Corresponding author.
E-mail address: r.silvester@bangor.ac.uk (R. Silvester).

https://doi.org/10.1016/j.scitotenv.2025.180607
Received 14 May 2025; Received in revised form 21 September 2025; Accepted 23 September 2025
Available online 1 October 2025

highlight the need to integrate high-resolution genome-resolved metagenomic surveillance into national AMR monitoring frameworks to track emerging threats, characterise ARG reservoirs and inform targeted public health interventions.

Abbreviations

ARG	Antibiotic-resistance genes
AMR	Antimicrobial resistance
CSO	Combined sewer overflow
MAG	Metagenome-assembled genomes
MDR	Multi-drug resistance
WBE	Wastewater-based epidemiology
WWTP	Wastewater treatment plant

1. Introduction

Antimicrobial resistance (AMR) is a global health crisis, projected to cause nearly 39 million deaths by 2050 (Naghavi et al., 2024). Wastewater, enriched with organic matter, antibiotics, and heavy metals, serves as a key reservoir for the emergence and spread of AMR (Perveen et al., 2023). The role of wastewater in disseminating AMR is a growing concern, as conventional treatment processes are not designed to mitigate AMR, leading to the continuous discharge of allochthonous bacteria carrying antimicrobial resistance genes (ARGs) into environmental reservoirs (Obayiuwana et al., 2025). Once introduced, these microbes may alter the native microbial communities and potentially spread AMR across ecosystems—posing public health risks through contaminated water, food chains, or direct environmental exposure (Silvester et al., 2025a). Additionally, combined sewer overflows (CSOs) worsen the problem by significantly increasing the discharge of untreated pollutants into recipient water bodies during heavy rainfall events (Silvester et al., 2025b).

Such dynamic and intermittent pollution events highlight the need for high-resolution, culture-independent approaches to characterise microbial communities and ARGs in complex wastewater matrices. Advances in metagenomic research have shed light on extensive microbial diversity that remains undetectable through culture-dependent methods (Kim et al., 2023; Rinke et al., 2013), reshaping our view of the microbial tree of life. Metagenome-assembled genomes (MAGs) have emerged as a powerful tool for exploring this hidden diversity, particularly among yet-uncultivated microorganisms. MAGs are reconstructed microbial genomes derived from metagenomic sequencing data. They are generated by assembling short reads into contigs and then binning those contigs based on sequence composition and coverage patterns, to identify genomic fragments and traits specific to individual microbial populations within a complex community (Zhu et al., 2019). MAG-based studies have significantly expanded our understanding of microbial ecology, uncovering novel taxa from diverse environments, including wastewater systems (Becsei et al., 2024; Busi et al., 2023; Singleton et al., 2021). In wastewater research, they can offer deeper insights into the composition, persistence, and transmission of AMR-associated bacteria within microbial communities inhabiting natural and engineered ecosystems.

While numerous studies have examined AMR dissemination in hospitals and wastewater treatment plants (WWTPs), most are limited in temporal and geographical scope. Also, despite extensive research on the prevalence and diversity of ARGs in wastewater using conventional shotgun metagenomics, most studies have relied on read-based analyses. These approaches, while useful for gene-level profiling, are limited in their ability to resolve ARG–host associations or capture the ecological context of resistance determinants. Hence, accurately identifying bacterial hosts carrying specific ARGs remains a critical and largely unresolved challenge. This is essential for estimating the actual risk of spreading antibiotic resistance from environmental microbes to human pathogens. Equally important is determining the genomic context of ARGs, such as their chromosomal or plasmid location and neighbouring elements, which informs their potential for horizontal gene transfer (HGT). Previously, we studied the resistome profiling in wastewater samples from hospitals and WWTPs in Wales using shotgun metagenomics read-based approach and predicted potential bacterial hosts using network-based prediction approach (Silvester et al., 2025b). However, these predictions may be influenced by unaccounted variables, such as putative relationships identified might be coincidental (Dai et al., 2022). To address these limitations, in this study, we employed a genome-resolved approach leveraging MAGs, which enables direct linkage of ARGs to their bacterial hosts within complex wastewater communities. This method offers a more accurate and holistic understanding of AMR transmission dynamics than conventional read-based or correlation-based strategies.

Using an archived metagenome dataset derived from wastewater samples collected from major hospitals and their associated receiving WWTPs across six diverse regions of Wales, United Kingdom (UK)-spanning urban, rural, and coastal settings- we aimed to reconstruct bacterial communities and identify the key ARG-hosting taxa within these environments and to identify the dominant ARG carriers within hospital and WWTP settings. Additionally, by comparing influent and effluent samples from WWTPs employing different treatment technologies, we evaluated how treatment processes can potentially influence shifts in ARG-host associations. Such analysis offers valuable insights into the antibiotic resistance transmission and its potential implications for public health.

Our genome-resolved MAG-based analysis has key advantages over the conventional read-based AMR surveillance approach relying on short-read shotgun metagenomic sequencing: (1) accurate identification of ARG-host relationships across complex wastewater environments and treatment stages, offering direct evidence of which organisms harbour ARGs rather than relying on indirect correlations; (2) recovery of genomes of both cultivable and uncultivable ARG-carrying microbes, overcoming the limitations of traditional culture-dependent methods. (3) Unlike PCR-based methods that target only known ARGs or culture-based approaches limited to specific pathogens, our whole-genome approach aims to provide a comprehensive characterisation of the entire resistome and its bacterial hosts. These data establish a foundation for targeted surveillance strategies, infection control measures, and evidence-based policy interventions in both clinical and environmental settings.

2. Methods

2.1. Selection of archived wastewater metagenomic samples

This study utilised existing metagenomic sequencing data generated from our Welsh Government's National Wastewater Monitoring program (Perry et al., 2024). In 2023, 24-h composite wastewater samples were collected using refrigerated autosamplers programmed to take aliquots every 15 min, from multiple sites across the nation of Wales following protocols and sample processing as described in Silvester et al. (2025a). Importantly, hospital and WWTP samples were collected during the same timestamp, ensuring synchronisation of sampling events across the wastewater pathways. The original dataset comprised samples from 47 wastewater treatment plants (WWTPs) and eight hospitals.

For this genome-resolved pilot study from Wales, we selected a representative subset of shotgun metagenome sequence samples to investigate hospital and community-associated AMR in greater detail.

The dataset included 48 metagenomes from hospital sites in North Wales (Bangor, Rhyl, Wrexham), Mid-Wales (Aberystwyth), and South Wales (Llandough, Swansea), along with their corresponding WWTP influent samples. Sequences were chosen from April and July 2023 to capture seasonal variability; for North Wales, additional metagenome were available from October and December 2023. WWTP effluent metagenomes were available only for North Wales sites. A summary of the selected hospitals and WWTP sites is provided in Supplementary DataS1.

2.2. Metagenomic assembly, binning, and quality assessment

We used the selected wastewater shotgun metagenome sequence datasets as mentioned above to reconstruct bacterial genomes, known as metagenome-assembled genomes (MAGs) (Krakau et al., 2022; https://nf-co.re/mag). Assembly of short reads into contigs was performed using MEGAHIT v1.2.9 (Li et al., 2016), and contigs were subsequently binned into MAGs. To maximise genome recovery, two binning tools were used: MetaBAT2 (Kang et al., 2019) and MaxBIN2 (Wu et al., 2016). DASTool was used to integrate and refine the resulting bins within each sample by selecting the highest-scoring non-redundant set. The quality of the recovered MAGs was assessed using CheckM (Parks et al., 2015), which evaluates genome completeness and contamination. MIMAG standards (Bowers et al., 2017), were used to choose high-quality MAGs. Low-quality MAGS (completeness <50 % and contamination >10 %) were excluded from downstream analysis.

2.3. Taxonomic assignment of the bins

Taxonomic classification of MAGs was performed using Genome Taxonomy Database toolkit (GTDB-Tk v2.1.1) based on 120 bacterial marker genes. We selected one representative MAG from each GTDB family per sample, based on the highest CheckM quality score (completeness - 5 × contamination), without the need for de-replication among samples since all analyses were executed in a sample-wise manner (Parks et al., 2015).

Species-level classification was cross-referenced against the Unified Human Gastrointestinal Genome (UHGG) collection (v2.0.2) (Almeida et al., 2021) to determine overlap with human gut-associated taxa.

2.4. ARG-carriers and associated plasmids

MAGs were screened for presence of ARGs using ABRicate v1.0.0 (Seemann, n.d., https://github.com/tseemann/abricate) with the MEGARes 2.0 database with default parameters. All ARG-carrying MAGs were further screened for the presence of plasmid replicons using ABRicate v1.0.1 with default parameters. Plasmid-associated contigs were then examined to identify co-localised ARGs, enabling the detection of plasmid-associated ARGs. *Escherichia coli* and *Klebsiella pneumoniae* MAGs were further screened for virulence determinants using ABRicate v1.0.1 with Virulence Factor Database (VFDB) and for plasmid-associated contigs using PlasmidFinder with default parameters.

2.5. Data analysis and visualisation

A phylogenetic tree of ARG-hosting bacterial MAGs was constructed using FastTree v2.1.10 based on concatenated alignments of single-copy marker genes extracted by GTDB-Tk. The resulting phylogeny was visualised in ITOL v.7.1.

An ARG–host bipartite network was created using the *igraph* package in R Studio version 4.1.2 (R Core Team, 2021). Visualizations were conducted in Gephi v0.10.1. The distribution of ARGs in bacterial genomes across sample types was examined using a presence-absence heatmap, generated with the *pheatmap* package in R. The data matrix was binarised, assigning values of 1 (presence) or 0 (absence) for ARGs within each MAG.

The microbial diversity of wastewater samples was evaluated using Shannon, Simpson, and Richness indices, providing insights into species evenness and richness across sample types. Diversity metrics were calculated using the *vegan* package in R. To assess whether sample type influenced microbial community composition, we conducted a Permutational Multivariate Analysis of Variance (PERMANOVA) using the adonis2 function in the vegan package. To further explore beta diversity and visualise sample clustering, Principal Coordinate Analysis (PCoA) was conducted using Bray-Curtis dissimilarity as the distance metric, implemented within the vegan package. The resulting ordination plots were generated using *ggplot2* package in R, providing an intuitive visual representation of inter-sample variation.

Relative abundance of MAGs at phylum and genus levels was visualised using stacked bar plots faceted by site with *ggplot2* package in R. UpSet plots were generated using the *UpSetR* package in R to analyse the overlap of the ARG-hosting community across different sites. This method represents shared and unique ARG hosts across multiple sites.

Sankey diagrams were generated in R using the *networkD3* package to highlight shifts in ARG host composition between raw WWTP influents and treated WWTP effluent wastewater.

3. Results

A total of 3978 medium-to-high-quality MAGs were recovered from 48 wastewater metagenome samples collected from hospitals and WWTPs across six sites in Wales. Of these, 1240 MAGs met the MIMAG criteria (Bowers et al., 2017) for high-quality (≥90 % completeness and ≤ 5 % contamination), while the remaining were classified as medium quality (≥50 % completeness and ≤ 10 % contamination) (DataS2). Of the 1240 high-quality MAGs, MetaBAT2 generated 796 (64.2 %), while MaxBin2 produced 444 (35.8 %) (Fig. S1). Similarly, for the 2738 medium-quality MAGs, MetaBAT2 recovered 1996 (72.8 %), whereas MaxBin2 retrieved 743 (27 %), highlighting MetaBAT2's higher efficiency in good-quality genome reconstruction.

The average (±SEM) number of raw reads per sample across the 48 samples selected for this study was 125 (±6.3) million reads, with a range of 43 to 208 million reads (DataS3). A moderate positive correlation was observed between depth and number of MAGs ($R^2 = 0.296$), with an estimated gain of approximately 0.47 MAGs per million reads. Based on our findings, a sequencing depth of approximately 150–170 million reads/sample appears optimal for achieving high-quality genome-resolved metagenomic data in complex wastewater samples. However, further studies with larger sample sizes are needed to confirm this threshold.

3.1. Taxonomic classification and distribution of MAGs

Among the 3978 bacterial MAGs recovered, 70.9 % ($n = 2821$) were classified up to the genus level distributed over 27 phyla. Pseudomonadota (25 %; $n = 997$), Bacteroidota (15.4 %; $n = 615$), Bacillota (12 %; $n = 479$), and Actinomycetota (11.5 %; $n = 457$) represented the dominant phyla (Fig. 1a). Nine MAGs remained unclassified at the phylum level. Additionally, 243 MAGs were unclassified at the genus level, 30 at the family level, and 4 at the order level.

Less frequent phyla included Acidobacteriota, Armatimonadota, Bdellovibrionota, Campylobacterota, Chloroflexota, Cyanobacteria, Deinococcota, Dependentiae, Desulfobacterota, Eisenbacteria, Elusimicrobiota, Eremiobacterota, Fibrobacterota, Fusobacteriota, Gemmatimonadota, Latescibacterota, Margulisbacteria, Myxococcota, Nitrospirota, Patescibacteria, Planctomycetota, Spirochaetota, Synergistota, and Verrucomicrobiota. These low-detected groups contribute to the overall taxonomic diversity within the dataset. The MAGs in the Dependentiae phylum are classified at class, partially at order, and family levels but entirely unclassified at the genus level.

PERMANOVA results indicated a significant effect of source on wastewater microbial community composition ($p = 0.001$). This

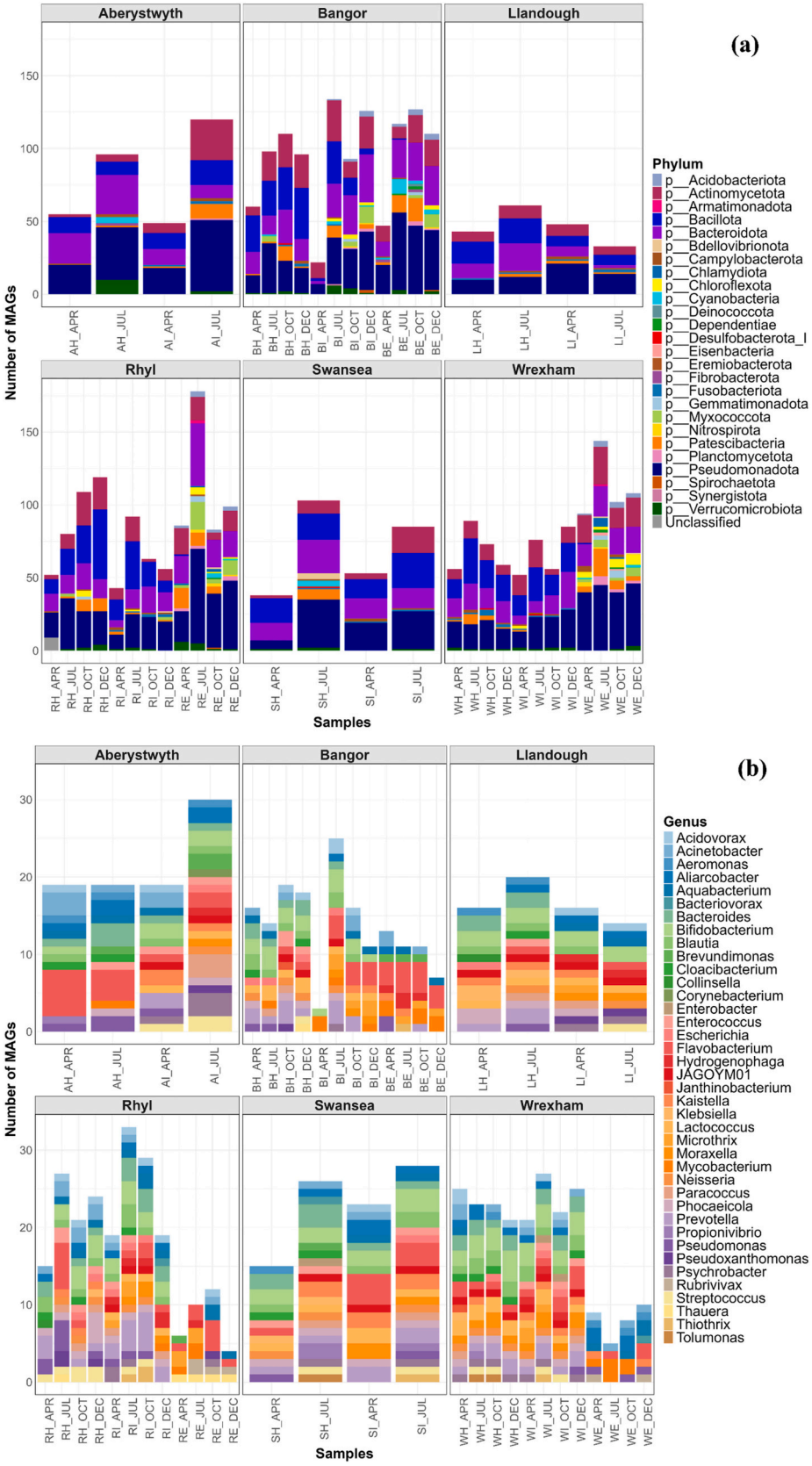


Fig. 1. Bar plots showing the number of (a) phylum level and (b) top 40 genus level bacterial MAGs recovered in wastewater collected from six major urban sites across Wales. Principal coordinate analysis (PCoA) plots illustrating the grouping of wastewater samples according to season-wise composition of (c) total bacterial genomes (d) ARG-carrying genomes. Each point in Fig. 2 c/d represents a sample, coloured by site, with the ellipses indicating the 95 % confidence interval for each site group. Site labels: H = hospital wastewater; E = WWTP effluent; I = WWTP influent, A = Aberystwyth, B = Bangor, L = Llandough, R = Rhyl, S = Swansea, W = Wrexham.

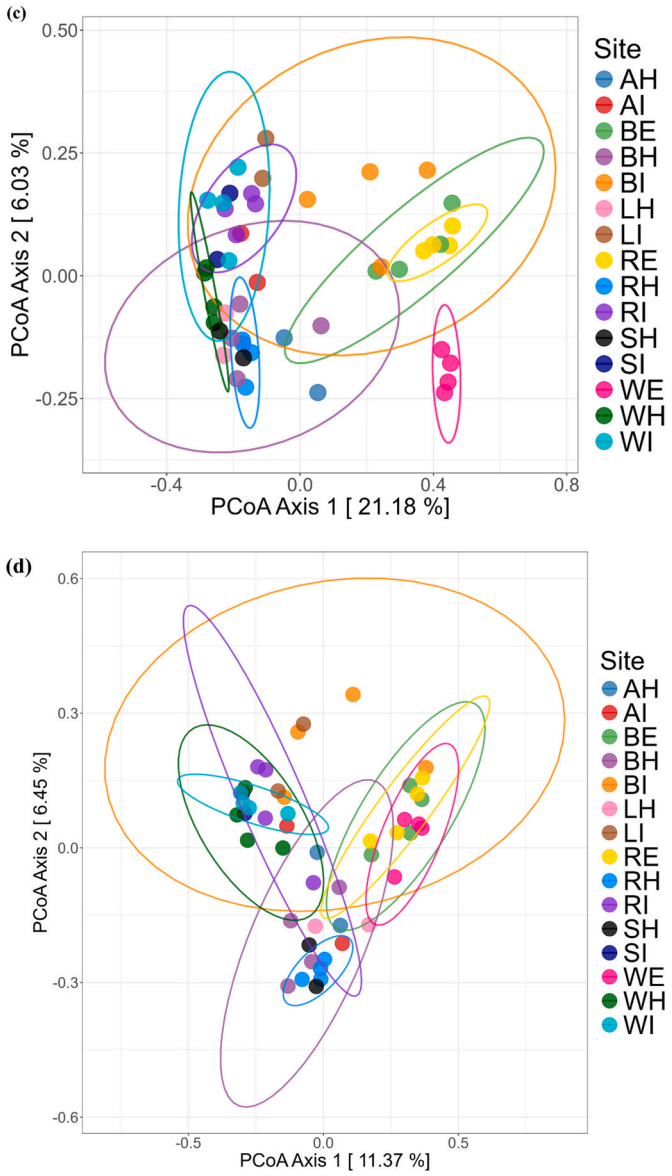


Fig. 1. (*continued*).

suggests that microbial community structures differ significantly between untreated hospital effluent, WWTP raw influent water, and WWTP treated effluent (Fig. S2). Sewage microbial community composition varied seasonally across all sites and across sample types (WWTP influent, WWTP effluent, hospital), with Pseudomonadota consistently dominant. Phyla belonging to Armatimonadota, Chlamydiota, Dependentiae, and Eisenbacteria, were unique to the effluent from WWTPs, highlighting their persistence and growth in the treatment system.

Bifidobacterium, *Ruminococcus*, *Alistipes*, *Bacteroides*, *Gemmiger*, and *Agathobacter* were the predominant genera retrieved from untreated hospital wastewater (Fig. 1b). In WWTP raw influent water, *Bifidobacterium*, *Blautia*, *Flavobacterium*, *Gemmiger*, and *Prevotella* had relatively higher abundance, whereas *Flavobacterium*, PHOS-HE28, *Azonexus*, JJ008, and *Mycobacterium* dominated in the treated WWTP effluent samples. PCoA plot shows the clustering of samples based on genus-level bacterial MAG composition (Fig. 1c). A statistically significant difference was found in microbial community composition between treatment types i.e. activated sludge-treated (Rhyl) and tertiary- UV treated (Bangor and Wrexham) effluents (PERMANOVA, $p = 0.02$). UV treatment exhibited a better bacterial community removal efficiency.

3.2. ARG carriers in wastewater revealed from MAGs

We further screened the MAGs for ARGs to accurately identify the ARG carriers. The distribution of ARGs in the MAGs from 6 sites are shown using presence-absence heatmaps (Fig. S3 a-f). ARGs were identified in 13.6 % ($n = 541$) of the MAGs(DataS4). The phylogenetic tree of the ARG-hosting MAGs recovered from six sites is given below (Fig. 2). These spanned 12 bacterial phyla, with Pseudomonadota (204 bins) and Bacteroidota (158 bins) being the most dominant, followed by Actinomycetota (81 bins) and Firmicutes (now reclassified as Bacillota) (63 bins). Less abundant host phyla included Campylobacterota (6), Chloroflexota (5), Cyanobacteria (1), Fibrobacterota (1), Fusobacteriota (8), Gemmatimonadota (2), Myxococcota (3), and Patescibacteria (5), highlighting the diversity of ARG-hosting taxa. About 30.7 % ($n = 167$) of ARG-hosting MAGs were classified up to the genus level.

The distribution of ARG-hosting MAGs varied between sites, with numerous unique bacterial species/genera identified at each site (Fig. 3a-c). Approximately 11 % of the ARG hosts were potentially novel or uncultured/yet-to-be cultivated and carried clinically relevant ARGs encoding resistance to vancomycin, fosfomycin, beta-lactams, tetracyclines, macrolides, aminoglycosides, and other antibiotic classes.

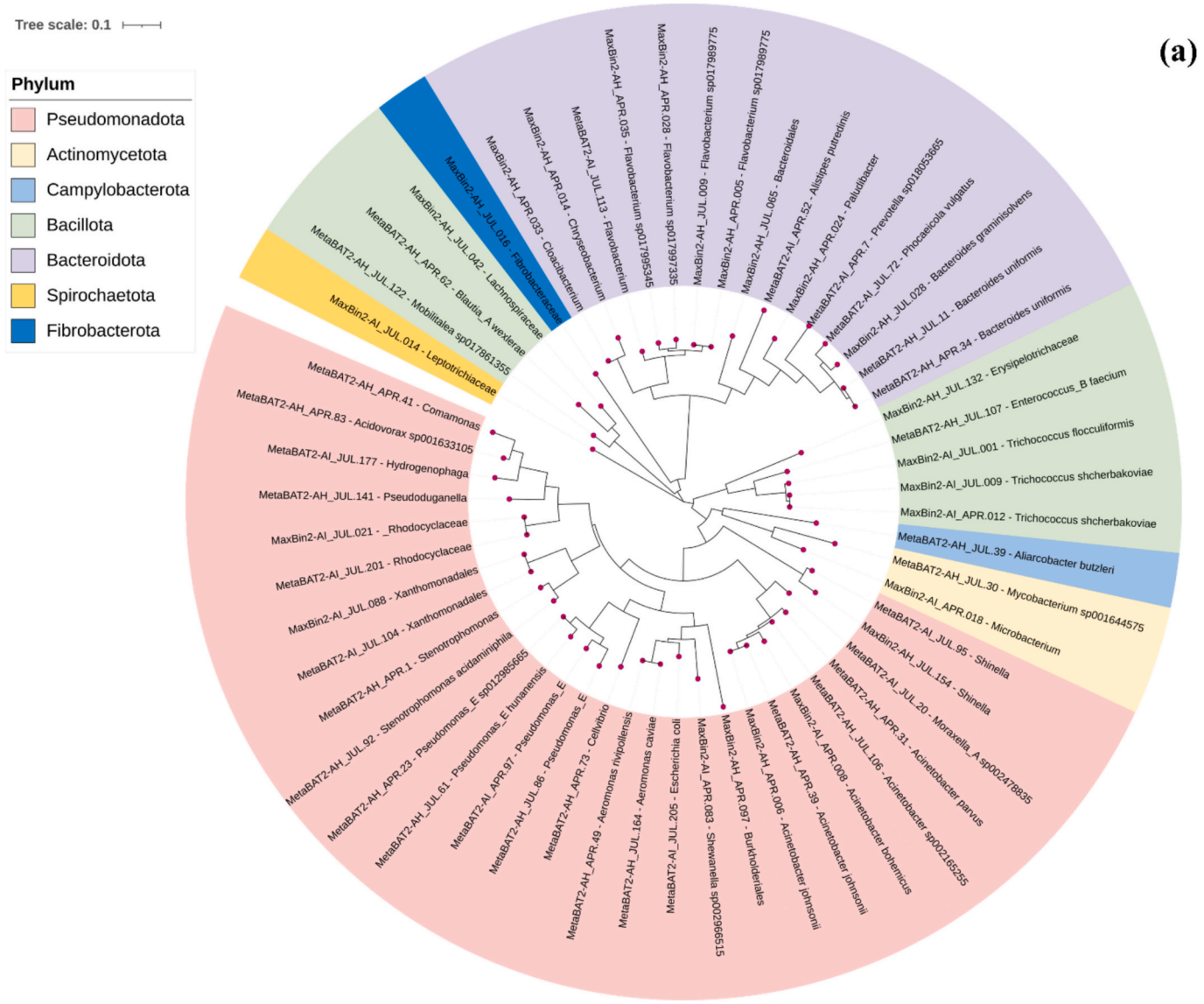


Fig. 2. Phylogenetic trees of metagenome-assembled genomes (MAGs) carrying antimicrobial resistance genes (ARGs) across six major urban wastewater sites in Wales to provide an evolutionary perspective on ARG-carrying microbial communities: (a) Aberystwyth; (b) Bangor; (c) Llandough; (d) Rhyl; (e) Swansea; (f) Wrexham. Clades colour-coded to indicate the phyla. Phylogenetic tree was constructed using FastTree v2.1.10 based on concatenated alignments of single-copy marker genes extracted by GTDB-Tk and visualised using iTOL.

We compared ARG-hosting MAGs with the Unified Human Gastrointestinal Genome (UHGG) database to identify the human gut-associated taxa, and 20 % of them were classified as human gut microbes, dominated by *Bacteroides uniformis* and *E. coli (*DataS5).

About 14.6 % ARG-hosting MAGs simultaneously harboured plasmids as well (DataS6). However, upon further specific screening the plasmid-associated contigs of each MAG for ARGs, only four were found to contain plasmid-associated ARGs; two Col3M_1 plasmids carrying quinolone resistance encoding *qnrD1_1*, IncQ2_1 with tetracycline resistance gene *tet(C)_3*, and rep22_1b_repB with aminoglycosidase encoding *aph(2″)-Ic_1*.

3.2.1. *ARG-hosts in hospital wastewater*

Pseudomonadota and Bacteroidota were the top ARG carriers in hospital wastewater (Fig. 2). MAGs belonging to opportunistic pathogenic species such as *Escherichia coli*, *Klebsiella* spp., *Serratia fonticola*, and *Enterococcus faecium* carried multiple ARGs conferring resistance to three or more classes of antibiotics, reflecting a multidrug-resistant profile. (Fig. 4a). Clinically relevant vancomycin resistance genes were present in 26 MAGs, with *Microbacterium* being the predominant host. OXA-type β-lactamase genes were detected in 37 MAGs, while tetracycline resistance genes were found in 40 MAGs, with multiple taxa acting as hosts. *Bacteroides uniformis*, carrying the cephalosporinase gene bla_{CBLA}, was detected across all hospital sites (Fig. 4a).

3.2.2. *ARG hosts shifted from influent to effluent in WWTPs*

Phyla belonging to Pseudomonadota, Bacteroidota and Bacillota (earlier Firmicutes) were the top ARG carriers in WWTP influents, with *Moraxella* sp., *Phocaeicola vulgatus* and *Flavobacterium* sp. 017997335 being the top hosting genera (Fig. 2). *Phocaeicola vulgatus* MAGs carried ARGs belonging to different antibiotic classes, such as aminoglycosides (*ant*), β-lactams ($bla_{CBLA,}$ bla_{CFX}), macrolides (*mef*), lincosamides (*lnu*), tetracyclines (*tet*), and colistin resistance (*mcr*) (Fig. 4b). All *Moraxella* MAGs harboured gene for intrinsic colistin resistance *(ICR)*. The *Flavobacterium* MAGs in this dataset carried resistance genes targeting β-lactams (bla_{OXA}, bla_{VEB}), aminoglycosides (*ant6*, *aph2*-dprime),

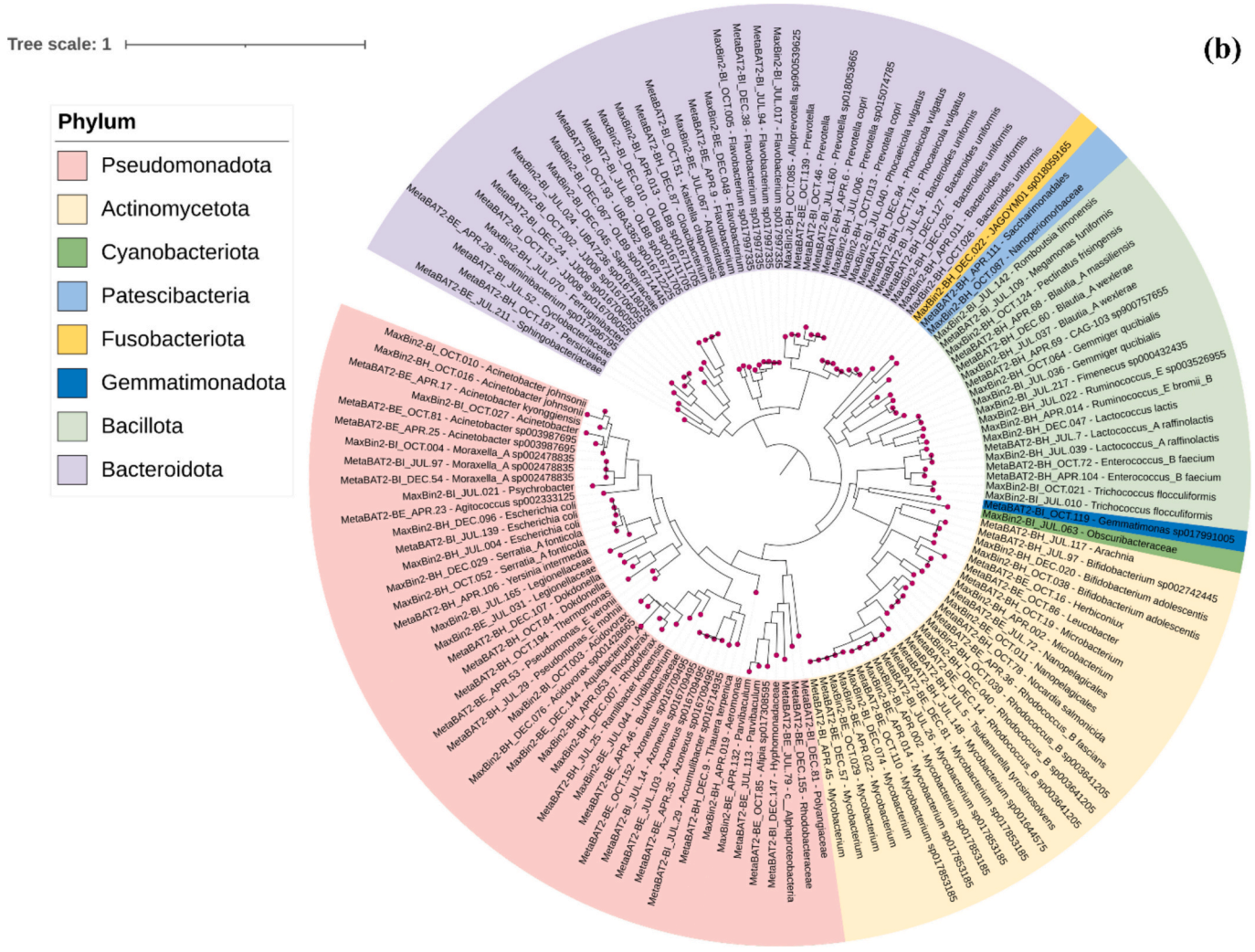


Fig. 2. (*continued*).

macrolides (*ereD*, *ermB*, *ermF*), lincosamides (*lnuG*), phenicols (*mca*, *mphG*), and tetracycline (*tetM*, *tetU*, *tetX*).

Pseudomonadota, Bacteroidota and Actinomycetota were dominant ARG hosts in the WWTP effluents from the studied North Wales sites. WWTP effluent ARG hosts were less diverse. *Mycobacterium* being the dominant genus with resistance determinants, carrying *mtrAD* and *rbpA* (17.7 %; Fig. 4c). *MtrAD* is multi-drug RND efflux regulator and *rbpA* RNA-polymerase binding protein A (RBPA) involved in the intrinsic rifampicin resistance in *Mycobacterium* (Sirichoat et al., 2023). Macrolide and oxacillin resistance genes were also frequently detected in effluent MAGs. A statistically significant difference was observed in the composition of effluent ARG hosts between WWTPs using varied treatment; activated sludge-treated (Rhyl) and UV-treated (Bangor and Wrexham) WWTP effluents ($R^2 = 0.133$, $F = 1.53$, $p = 0.005$).

A shift in the composition of ARG-hosting taxa was observed when comparing WWTP influent and effluent samples, indicating changes in the bacterial hosts of ARGs between untreated influent and treated effluent (Fig. 4b, c). Sankey plots clearly indicate this shift from influent to effluent (Fig. 5a, b). The total number of ARG-hosting MAGs also varied monthly and between influent and WWTP effluent samples (Fig. 1d), with higher counts and diversity observed in influent samples during certain months, while WWTP effluent samples generally showed lower counts. MAGs with multi-drug resistance determinants and opportunistic pathogen MAGs, such as *E. coli*, were mostly identified in influent samples compared to effluents.

3.3. Distinct ARG-host associations revealed

The *tet* and *bla*$_{OXA}$ type genes that encode resistance to tetracycline and beta-lactam oxacillin were the most frequently identified ARGs in the MAGs (Fig. 4a-c). The *bla*$_{OXA}$ type genes were primarily associated with *Cloacibacterium*, *Aliarcobacter*, *Comamonas denitrificans*, *Aeromonas media*, *Acinetobacter*, *Shewanella*, and *Flavobacterium* (Fig. 4a-c, DataS4). The *tet* genes exhibited diverse subtypes, each associated with distinct host organisms; *tetQ* was linked to *Bacteroides graminisolvens*, *tet33* to *Microbacterium*, *tet40* to *Gemmiger quercibialis*, *tetA* to *Bifidobacterium adolescentis*, *tet32* to *Blautia wexlerae*, and *tetW* to *Trichococcus*.

Certain ARGs were found to be consistently associated with specific taxa majority being intrinsic. Some notable associations included intrinsic colistin resistance *(ICR)* in *Moraxella* sp., *mtraD* and *rbpA* (intrinsic resistance to rifampicin) in *Mycobacterium*, with *rbpA* also found in *Rhodococcus*. *CblA* was unique to *Bacteroides uniformis*, vancomycin resistance gene *vanRO* was mostly detected in *Microbacterium*, while *ereD* was found in *Flavobacterium*.

The multidrug efflux system regulator *mexT* was consistently linked to *Acinetobacter*, while macrolide resistance-encoding *mefC* and *mphG* were associated with *Flavobacterium*, whereas *mexEF* efflux pump were identified in *Pseudomonas_E*. *Lnu*A and *mefE* were identified in majority of *Phocaeicola vulgatus*, and both beta-lactamases *bla*$_{MOX}$ and *bla*$_{OXA}$ were linked to *Aeromonas media*. Additionally, *aac6-prime*, *efmA*, and *msrC* were detected in *Enterococcus_B faecium*.

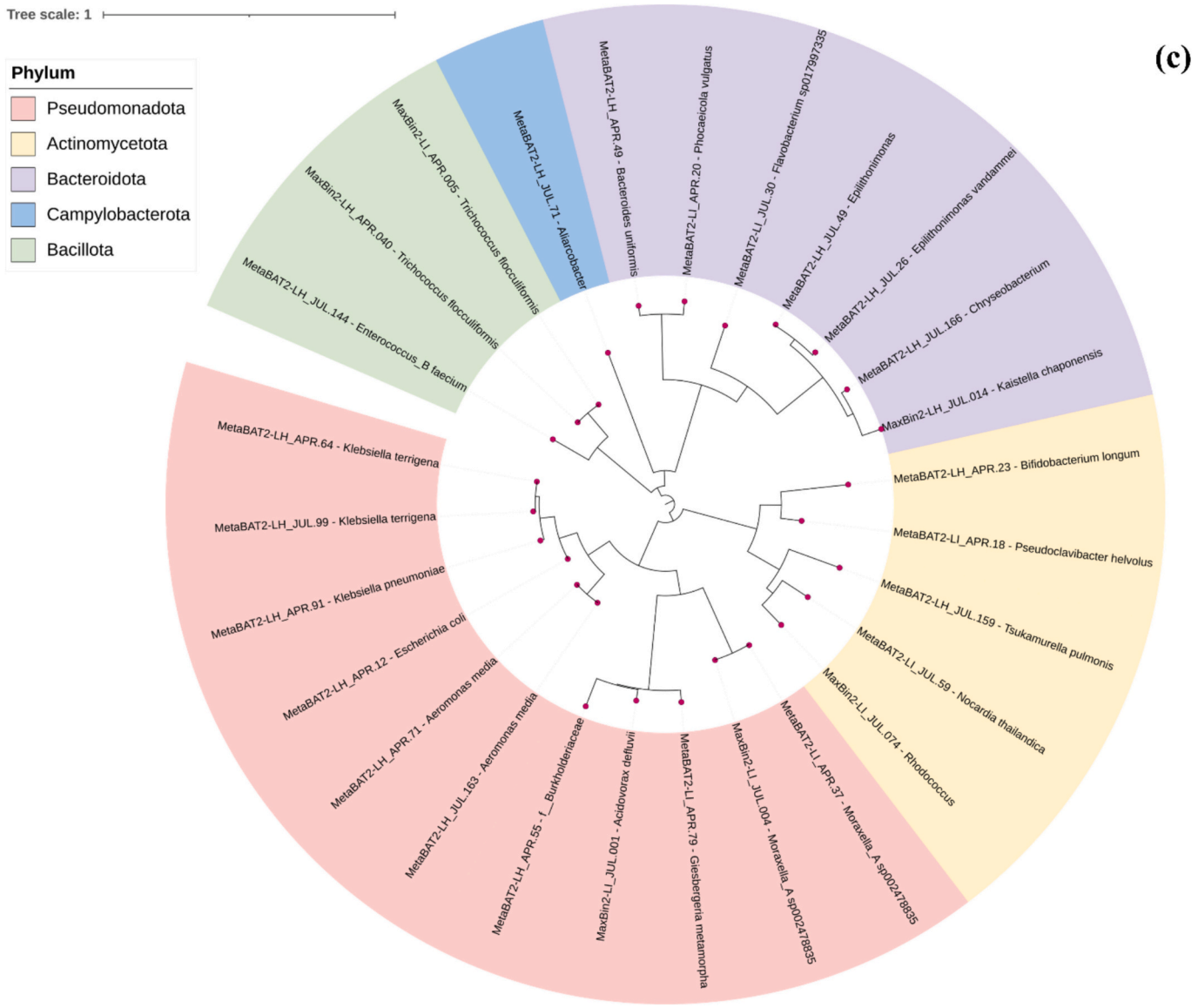


Fig. 2. (*continued*).

3.4. ARGs, virulence determinants and plasmids associated with *E. coli* and *Klebsiella pneumoniae* genomes

A total of 17 recovered MAGs belonged to *E. coli*, the majority of which originated from hospitals in North Wales, with none detected in WWTP effluent samples. Additionally, two from hospitals were identified as *Klebsiella pneumoniae*. These MAGs are submitted to NCBI under the Bioproject accession number PRJNA1235106.

A total of eight *E. coli* MAGs were identified across hospital-associated metagenomic assemblies. Each genome exhibited a unique profile of ARGs, virulence-associated factors, and plasmid replicon types (Supplementary Table 1). The identified ARGs were associated with multidrug efflux pumps (*acrA, acrB, acrD, acrE, mdtE, mdtG, mdtI, mdtJ, msbA*), β-lactam resistance (bla_{EC}, $bla_{TEM\text{-}1}$, $bla_{OXA\text{-}2}$, bla_{CTX}), aminoglycoside resistance (*ant2-dprime*), macrolide resistance (*mphB, ermX*), and tetracycline resistance (*tetW, tetX*), among others. Virulence factors associated with motility and adhesion (*fli, flg, fim, ecp*), iron acquisition (*chu, fep, iro, iucD, iutA*), and capsular polysaccharide synthesis (*kpsD, sitA, rcsB*) were widely distributed. Type III Secretion System (T3SS) and effector-like protein-encoding genes (*esp*) were also found in a few *E. coli* bins. This secretion system is an effector delivery system that allows Gram-negative bacteria to inject toxic effectors into host cells and competing bacteria (Zhou et al., 2014). The plasmid types identified from hospitals belong to various incompatibility (Inc) groups and Col plasmid types.

A total of nine *E. coli* were identified across assemblies from WWTP influents (Supplementary Table 1). Each exhibited a unique profile of ARGs, virulence-associated factors, and plasmid replicon types. Across the identified bins, multidrug efflux pumps (*acrA, acrB, acrD, acrE, mdtE, mdtG, mdtI, mdtJ, msbA, emrA, emrR, emrD*) were common. Aminoglycoside resistance genes (*aph(3′)-prime),* bacitracin *(bacA),* β-lactam resistance genes (bla_{EC}, $bla_{CTX\text{-}M\text{-}177}$, $bla_{TEM\text{-}1}$), macrolide resistance genes *(mphB),* colistin resistance gene (*mcr9.1),* and tetracycline resistance genes *(tet40)* were also present in a few *E. coli* genomes. Plasmid replicon typing revealed the presence of diverse incompatibility (Inc) groups, including IncF, and Col-like plasmids. The presence of multiple plasmid types within the same genome suggests plasmid co-existence.

The *K. pneumoniae* MAG recovered from Rhyl Hospital harboured a diverse array of ARGs associated with resistance to multiple antibiotic classes (Supplementary Table 1). Notably, genes encoding multi-drug efflux pumps (*acrA, acrD, emrB, oqxA, oqxB, kdeA*), β-lactamase ($bla_{SHV\text{-}120}$), fosfomycin resistance (*fosA6*), and efflux-pump regulatory

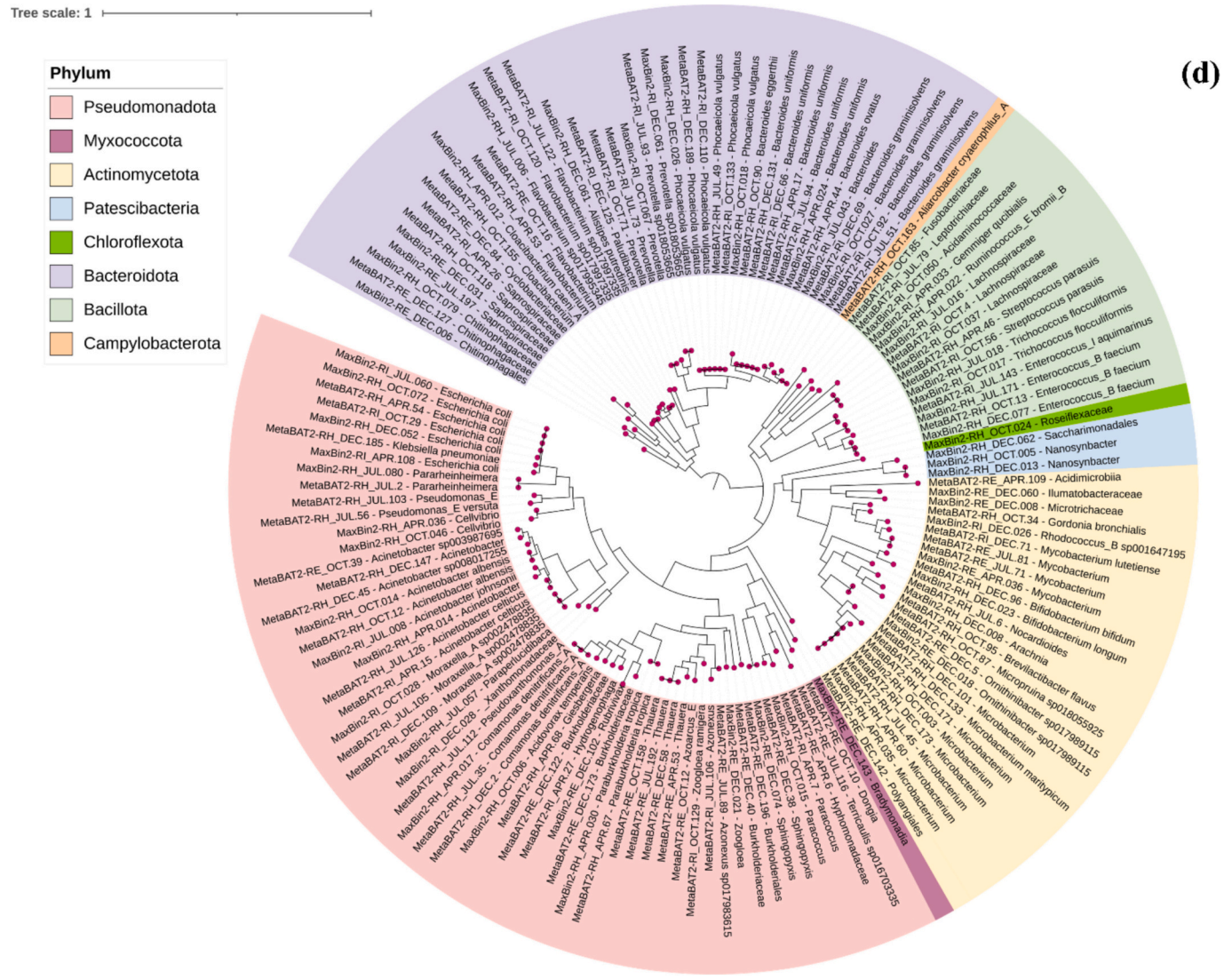


Fig. 2. (*continued*).

elements (*marA, ramA, soxS*) involved in multidrug resistance mechanisms were detected. The presence of the *omp37* gene, which has been linked to reduced permeability and carbapenem resistance, further highlights the multi-drug resistance potential of this strain. In addition to ARGs, the genome contained several virulence determinants, including genes related to iron acquisition (*entA-F, fepB-G, fes, iroE*), fimbriae and adhesion (*fimB-K, mrkF-J, ompA*), and stress response (*rpoS, fur, rcsB*). The Type VI secretion system (*tssF, tssG, sciN, vipA-C*) was also present, which may contribute to bacterial competition and persistence in the hospital environment.

K. pneumoniae from Llandough hospital harboured ARGs associated with β-lactam resistance (bla_{SHV-1}), efflux pump-mediated resistance genes (*kdeA*), global transcriptional regulators involved in multidrug resistance (*ramA, soxS*), and stress response regulators (*cpxAR*). The presence of *ampH*, a β-lactamase regulatory gene, suggests potential resistance to penicillins. Several virulence-associated genes were identified, including those involved in adhesion and biofilm formation (*fimA, fimI, fimC, fimG, fimF*), iron acquisition (*fes, irp2, entA, entB, fepC*), and secretion system components (*tssG, hcp/tssD, sciN/tssJ*). The presence of *rcsB*, a key regulator of capsule biosynthesis, suggests the strain may possess enhanced survival capabilities under host immune pressure. Additionally, *yagW/ecpD* and *yagV/ecpE* indicate the presence of *ecp*-encoded fimbrial structures, which contribute to bacterial adhesion and colonisation.

4. Discussion

This is the first study to integrate genome-resolved metagenomic characterisation of AMR in hospital and municipal wastewater across Wales, offering a foundational framework for national AMR surveillance. By characterising AMR at high resolution, this work provides critical insight into how ARGs and their bacterial hosts persist and spread in human-impacted environments. This study represents a substantial conceptual and methodological advancement over our previously published wastewater-based epidemiology (WBE) study (Silvester et al., 2025a). Unlike our previous read-based analysis, which relied on relative abundance and co-occurrence patterns, using our genome-resolved metagenomics approach, here we reconstructed near-complete bacterial genomes (MAGs) from short-read metagenomic sequencing data to directly identify the microbial hosts of antimicrobial resistance (AMR) genes. This genome-centric framework overcomes the limitations of read-based analyses by enabling the direct linkage of resistance determinants to specific bacterial genomes, significantly enhancing the interpretability and epidemiological relevance of the findings. Crucially, this allows for the identification of high-risk bacterial reservoirs and potential transmission vectors within wastewater

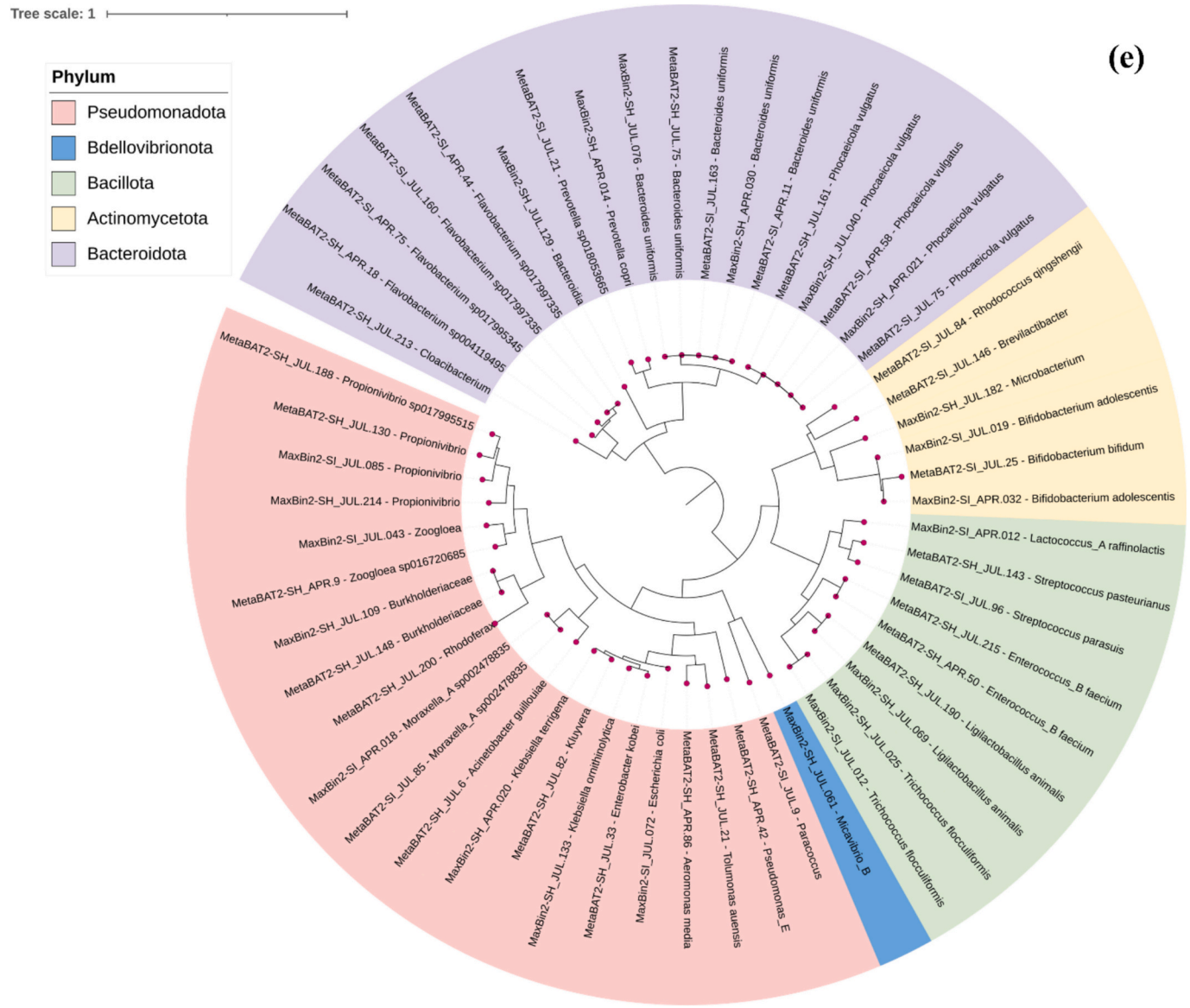


Fig. 2. (*continued*).

environments. This work contributes uniquely to environmental AMR surveillance by integrating spatial, seasonal, and genome-level data, laying the foundation for risk-based monitoring and One Health interventions. It thus advances the field beyond descriptive profiling, offering a framework for targeted action in both environmental management and public health policy.

However, the approach is not without limitations. A key limitation of this study is the exclusion of a substantial number of low-quality MAGs due to insufficient completeness or high contamination during stringent filtering. While this was necessary to ensure analytical robustness, it may have led to the loss of potentially relevant but partial genomes, particularly from rare or low-abundance taxa. Also, the reliance on reference-based classification (e.g., GTDB-Tk) may limit the identification of underrepresented taxa. Hence, some aspects of microbial and resistome diversity may have been underrepresented in the final dataset.

4.1. MAGs uncover microbiomes and ARG carriers in wastewater environments

The wastewater microbiomes were dominated by Pseudomonadota, Bacteroidota, Bacillota (formerly Firmicutes), and Actinomycetota, consistent with previous studies (Oluseyi Osunmakinde et al., 2019; Zheng et al., 2023). The prevalence of Bacillota and Bacteroidota, both common in the human gut, reflects substantial human-derived microbial input. As the gut microbiome is a known reservoir of ARGs, particularly in Bacillota (Crits-Christoph et al., 2022), our findings clearly suggest that gut-associated taxa in wastewater can carry clinically relevant ARGs.

Approximately 6 % of the recovered MAGs in our study could not be classified at the genus level, indicating the presence of potentially novel or poorly characterised microbial lineages. Similarly, Singleton et al. (2021) reported that around 7 % of MAGs recovered from activated sludge in Danish WWTPs were unclassified.

Around 13.6 % of MAGs recovered from wastewater carried ARGs, whereas a previous study from the Caspian aquatic environment found that nearly 24 % of genomes carried ARGs, highlighting source-wise variation (Goodarzi et al., 2022). The bla_{OXA} and *tet* genes, encoding resistance to oxacillin and tetracycline, respectively, were the most frequently identified ARG types in the wastewater microbiome.

Some ARGs were found to exhibit host specificity, including a few associated with intrinsic resistance traits. For example, *Moraxella* species carried intrinsic colistin resistance (ICR) genes (Wei et al., 2018),

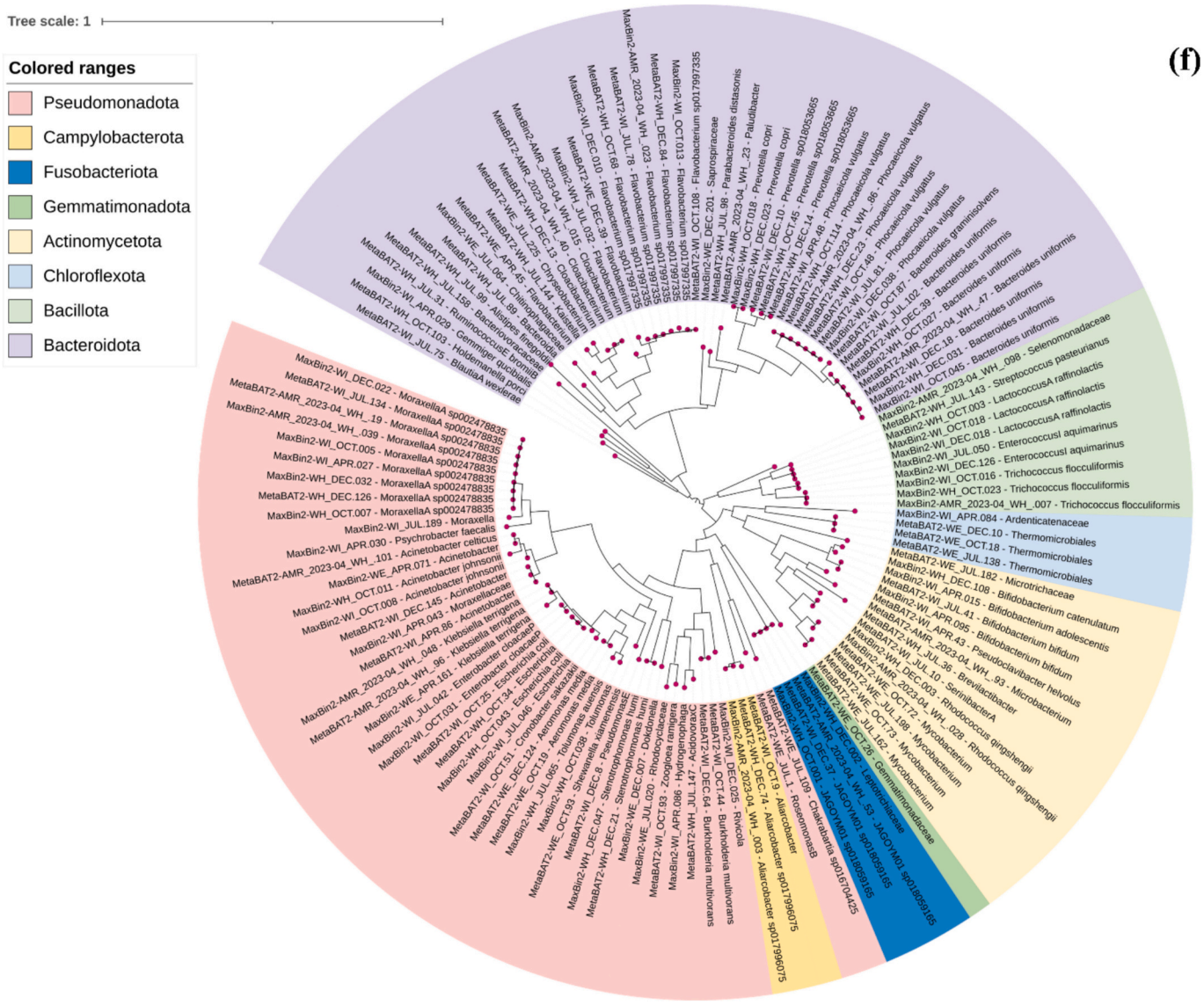


Fig. 2. (*continued*).

and the cephalosporinase gene *cblA*, was intrinsic to *Bacteroides uniformis* (Diebold et al., 2023). *Mycobacterium* hosted *mtrA* and *rbpA* genes, both involved in intrinsic resistance to rifampicin (Hu et al., 2012). The *rbpA* gene encodes RNA polymerase-binding protein A, an intrinsic factor that enhances rifampicin tolerance in *Mycobacterium* species. *MtrA* in *Mycobacterium* is a key regulator involved in the intrinsic resistance to some antimycobacterial drugs and the *mtrD* gene is part of the multidrug efflux pump system, specifically within the MtrAB two-component regulatory system (Sirichoat et al., 2023).

Conversely, the repeated identification of certain ARGs in MAGs of specific taxa, but not in all MAGs belonging to those taxa, may indicate likely acquisition via horizontal gene transfer rather than intrinsic origin. For instance, the gene *lnuA*, which encodes a lincosamide nucleotidyltransferase, was consistently detected in multiple *Phocaeicola vulgatus* MAGs, suggesting potential acquired resistance. These associations offer valuable insights into the mechanisms of antibiotic resistance and highlight the role of specific bacterial taxa as important reservoirs for both intrinsic and acquired ARGs.

E. coli and *K. pneumoniae* genomes were mainly recovered from hospital wastewater and, to a lesser extent, from WWTP influents. These genomes carried diverse ARGs, virulence factors, and plasmid replicons. While no plasmid-associated ARGs were detected in these genomes, the coexistence of multiple plasmids suggests the potential for future ARG acquisition, especially in hospital environments where selective pressures are high. This aligns with previous reports of plasmid-mediated ARG transfer between *K. pneumoniae* and *E. coli* in patients (Goren et al., 2010), highlighting hospitals as AMR hotspots and potential point sources for environmental dissemination (Hassoun-Kheir et al., 2020 We suggest that these organisms could be specifically targeted within public health surveillance programmes due to their high potential risk for causing MDR infections. Although we screened for plasmid-associated ARGs, a comprehensive genome-wide analysis of all mobile genetic elements (MGEs) was beyond the scope of this study. We plan to incorporate this into future investigations to examine the roles of integrons, plasmids, and transposons in the dissemination of AMR across microbial populations.

Not all ARG-carrying hosts were opportunistic pathogens. Many were environmental or gut-associated bacteria, including commensals and facultative anaerobes. While not inherently pathogenic, these organisms can act as important reservoirs of ARGs, enabling HGT to pathogenic bacteria within wastewater systems and downstream environments (Paul and Das, 2022). This highlights the broader ecological impacts of ARG dissemination, extending beyond hospital waste into larger water ecosystems. This broadens the scope of AMR surveillance

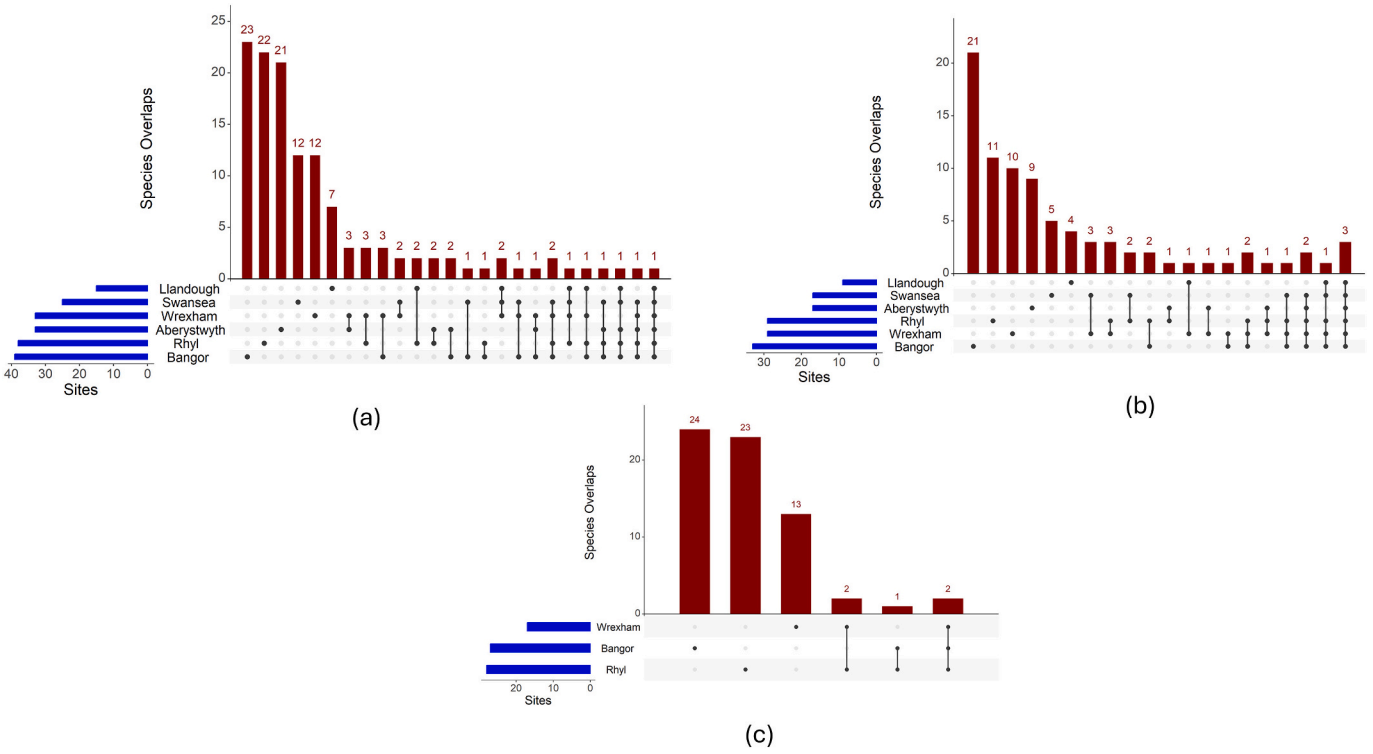


Fig. 3. UpSet plots showing the distribution and overlap of ARG-hosts across (a) hospital, (b) WWTP untreated influent, and (c) WWTP treated effluent samples. Black points connected by lines represent unique intersections, indicating species/genera that are shared across different sites, while vertical red bars depict the number of species in each intersection. Set size bars on the left indicate the total number of ARG-hosting species detected in each category.

beyond clinically significant species to include the wider microbial community.

Recent advances in sequencing technologies and computational methods have helped in successfully deciphering the uncultivated microbes, often referred to as "microbial dark matter" from complex environments (Sood et al., 2024). Around 11 % of the recovered genomes belonged to uncultivated bacteria, in line with a previous study (Kim et al., 2023). Notably, superphylum *Patescibacteria*, a highly diverse but largely uncultivated group of ultrasmall bacteria with small genomes carried ARGs conferring resistance to aminoglycosides, rifamycins, quinolones, β-lactams, sulfonamides, and fosfomycin in both hospital and WWTP-associated wastewater (Busi et al., 2023; Kim et al., 2023). These findings shed light on ARGs residing in the microbial dark matter and raise concerns about their role as previously overlooked resistance reservoirs. Unclassified bacteria have been previously reported as major hosts for bacitracin ARGs in sludge (Zhao et al., 2025). This study has thus revealed untapped sources of resistance from wastewater that are yet to be explored in traditional surveillance systems. Our recovery of multiple MAGs from uncultured ARG-hosting bacteria highlights the limitations of traditional culture-based surveillance and the importance of including yet-uncultivated microbes in AMR monitoring frameworks. These taxa may harbour resistance genes that have not yet been described and may act as vectors for ARG transfer to clinically relevant bacteria. Their persistence in WWTPs suggests the need for targeted monitoring approaches beyond known pathogens (Becsei et al., 2024). Future surveillance and interventions should thus prioritise both culturable and uncultured taxa to gain a comprehensive understanding of ARG dynamics in wastewater.

By leveraging a cultivation-independent, MAG-based approach, our study effectively links ARGs to specific bacterial hosts, spanning both recognised pathogens and previously uncharacterised lineages. This approach also uncovered novel reservoirs and transmission pathways of resistance, providing critical insights into AMR distribution and transmission in wastewater environments. The diversity of ARG-hosts unique to each site further reveals the need for tailored AMR monitoring and control measures based on regional and site-specific microbial ecology.

4.2. Shifts in composition from WWTP influent to effluent

While MAGs enabled high-resolution characterisation of ARG hosts across hospital and WWTP wastewater, it is equally important to understand how these ARG-hosting taxa and resistance determinants are affected by wastewater treatment processes. A key component of AMR mitigation is understanding the efficiency of WWTPs in removing ARGs, MGEs, and antibiotic-resistant bacteria (ARBs). Our findings highlight the need to assess WWTP performance in eliminating these elements, especially considering the risk that treated effluents may still contribute to downstream environmental contamination. Equally important is the characterisation of microbial communities in influent entering WWTPs, particularly given the role of combined sewer overflows (CSOs), which can release untreated or partially treated wastewater directly into receiving environments.

We observed a shift in overall microbial composition between WWTP influent and effluent, with selective enrichment of certain bacterial taxa during wastewater treatment. The microbial composition also varied according to the treatment methods. While Rhyl employed conventional activated sludge (AS), Bangor and Wrexham used tertiary UV treatment. AS-treated samples retained more unique genera, suggesting lower removal efficiency than UV treated systems. This aligns with previous findings, showing that AS systems support higher microbial diversity (Fang et al., 2024). Our previous work confirmed higher bacterial removal efficiency in UV-treated effluents (Silvester et al., 2025b), supporting broader evidence that UV effectively inactivates bacteria, including ARBs to some extent, thereby significantly reducing microbial loads (Oh et al., 2014).

However, WWTPs are not specifically designed to eliminate ARGs, MGEs, or ARBs. The detection of ARBs, even in UV-treated effluents, suggests the need for more advanced quaternary treatment strategies.

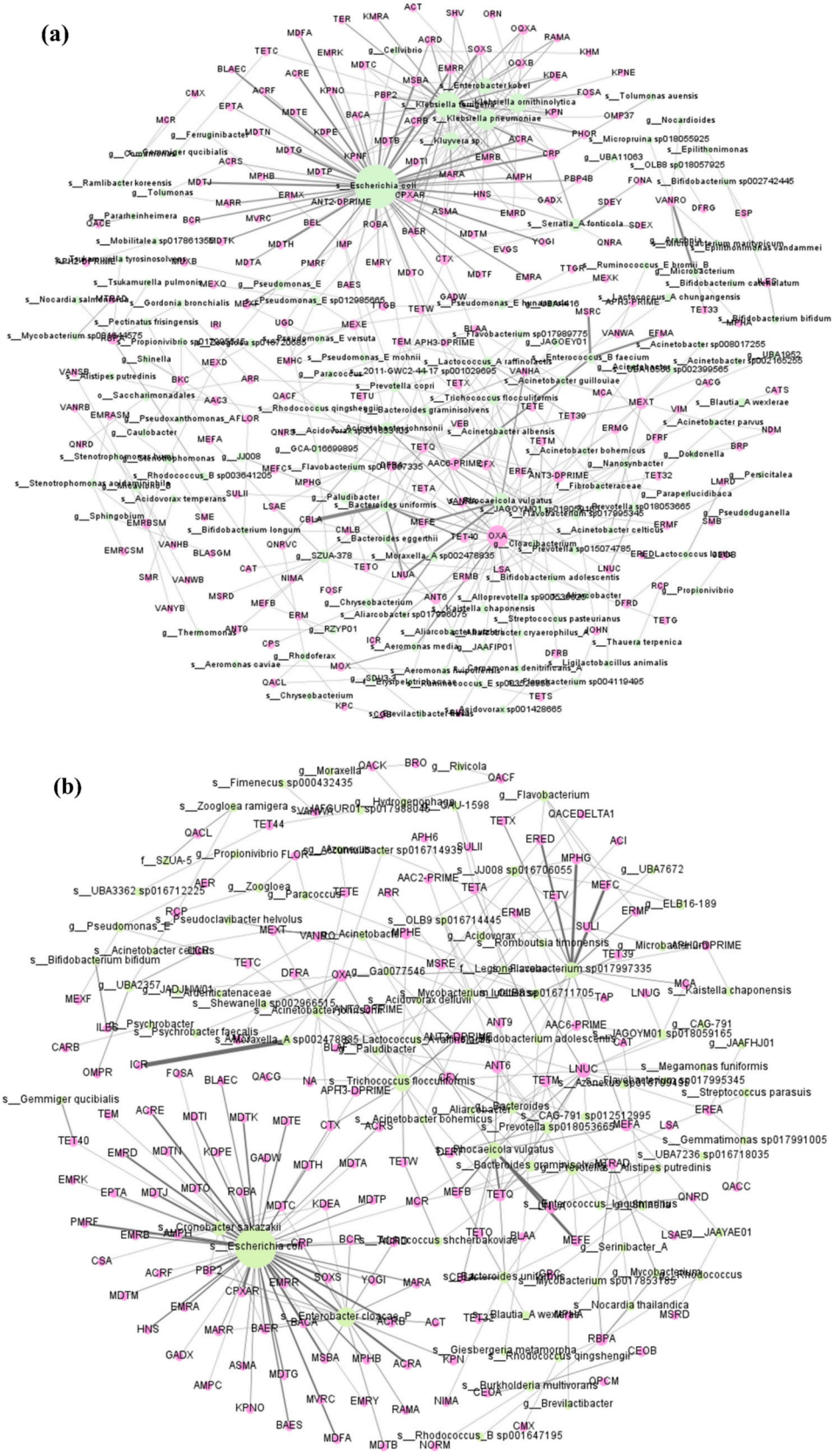


Fig. 4. Bipartite network diagrams illustrate the associations between antimicrobial resistance genes (ARGs) and their bacterial hosts (MAGs) in (a) hospital wastewater samples, (b) WWTP influent, and (c) WWTP effluent. Green nodes represent MAGs and pink nodes represent ARGs, with edges indicating the presence of specific ARGs within a given host.

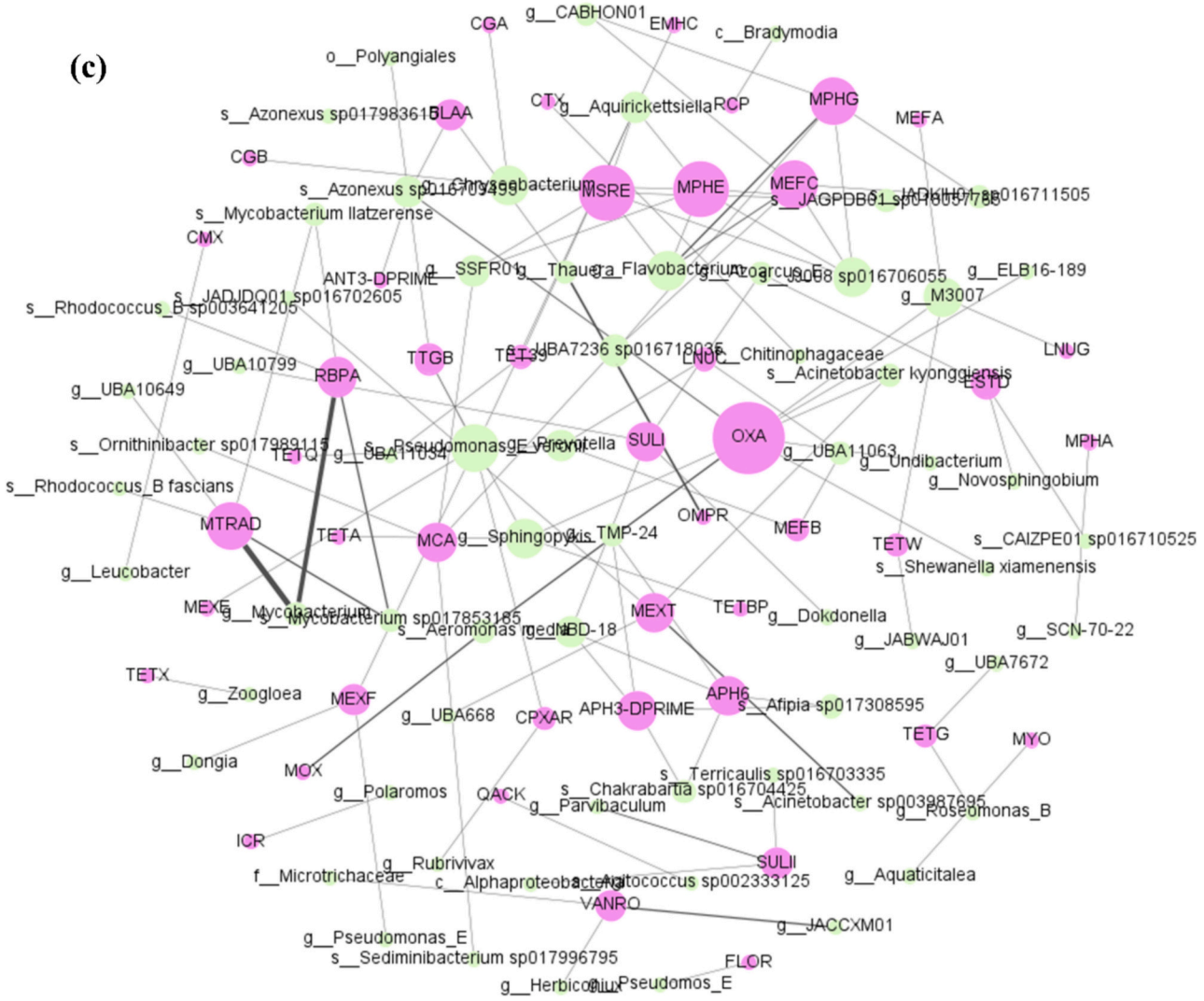


Fig. 4. (*continued*).

While the studied WWTPs reduced the overall diversity of ARG-hosting taxa, several ARGs and their hosts persisted post-treatment, consistent with the findings of Hultman et al. (2018). Notably, *Mycobacterium* emerged as a dominant ARG host in effluent, as previously observed (Wang et al., 2016). The persistence of bla_{OXA} and *tet* genes across treatment stages suggests their potential as indicator genes for monitoring residual AMR burden and assessing the efficiency of WWTPs.

Although pathogenic bacteria and associated ARGs were reduced in WWTP effluent, certain ARGs were still carried by taxa unique to or enriched in effluents, such as Chloroflexota. This phylum, frequently found in AS systems, plays a role in biofloc formation and can thrive under both anoxic and oxic conditions, facilitating its persistence in WWTPs (Petriglieri et al., 2023).

Moreover, AS treatment can enrich specific ARG subtypes, allowing them to persist in sludge water, often linked to antibiotic residues and ARGs from community wastewater (Zhao et al., 2025). This highlights the dual role of WWTPs as both filters and reservoirs for AMR dissemination. WWTPs create a conducive environment for HGT between native environmental microbes and pathogenic bacteria. Studies have shown an increase in the relative abundance of MGEs-associated ARGs in treated WWTP effluent compared to influent (Petrovich et al., 2018), with HGT events detected in 50 % of MAGs recovered from AS in Hong Kong (Fang et al., 2024). This suggests that treatment processes may inadvertently enrich mobile ARGs, enhancing their persistence. Therefore, a reduction in ARG-hosts as seen in our study does not necessarily indicate a decrease in AMR post-treatment. Future work should focus on characterising MGE-associated ARGs to better understand their mobility and public health implications.

Bacteroides uniformis, all of which carried the bla_{CBLA} gene was potentially completely removed during wastewater treatment, as it was undetected in effluent samples, though detected in both hospital and WWTP influent samples. Opportunistic pathogens such as *E. coli* were also not detected in effluents, possibly due to degradation or dilution effects. However, bacteria more resilient to WWTP effluent conditions may persist, complicating treatment outcomes (Sivalingam et al., 2023). A higher persistence of ARG carriers in influents compared to effluents offers valuable insights into the fate of ARGs in wastewater treatment systems. The presence of ARG carriers in WWTP influent is equally concerning given the number of CSO discharge events in Wales. During periods of heavy rainfall, these treatment facilities can reach their capacity, leading to CSOs. This results in the direct discharge of untreated sewage, containing all these ARGs and ARG carriers into rivers and coastal waters (Perry et al., 2024). The scale of the issue is considerable; in 2024 alone, England and Wales experienced 563,730 sewage spills, which is an average of 1544 per day (Rivers Trust, 2024). At the 3 WWTPs under study, annual CSO events were recorded at 35 (Wrexham), 65 (Kinmel Bay), and 99 (Bangor), lasting cumulatively between 13.6 and 39.1 days, representing 3.7–10.7 % of the year (Rivers Trust,

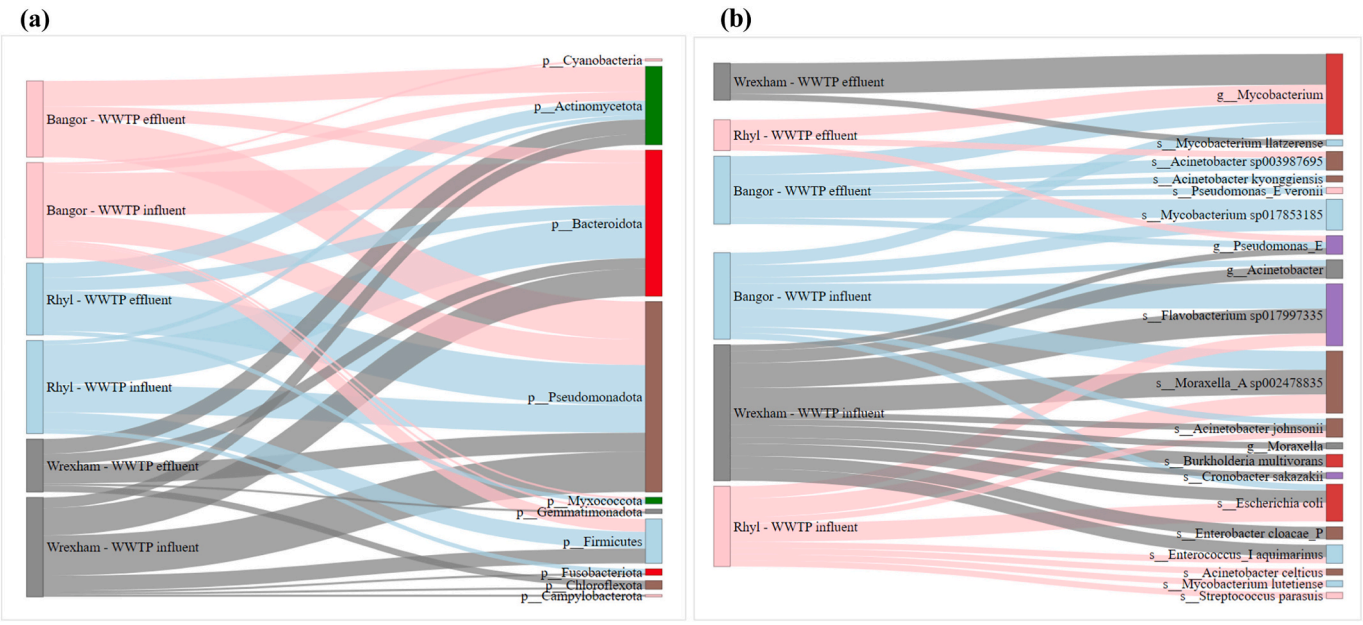


Fig. 5. Sankey plots illustrate shifts in antimicrobial resistance gene (ARG) hosts at the (a) phylum level and (b) clinically relevant human pathogens in WWTP influent and WWTP effluent wastewater samples from three sites. These plots reveal how wastewater treatment influences the persistence and removal of ARG-associated pathogens, providing insights into potential transmission risks and treatment efficacy.

2024). With climate change expected to increase the frequency of extreme rainfall events, CSO discharges and the environmental dissemination of AMR and pathogens are likely to rise.

Seasonal variation in microbial diversity was also observed in both WWTP influent and effluent samples, though the findings are limited by the number of sampling months. Such variations could be influenced by environmental changes, extreme rainfall events, shifts in wastewater composition, or differences in treatment efficiency. Even minor environmental fluctuations are known to significantly impact microbial taxa and ARG profiles (Javvadi and Mohan, 2024). Although our sampling spanned multiple hospitals and WWTPs, it was constrained temporally (limited months) and geographically for effluent samples (limited to North Wales). These limitations have restricted our ability to generalise treatment effects or detect robust seasonal trends. Future work should prioritise long-term, spatially expansive sampling to understand temporal and regional patterns in ARG emergence and resistance dynamics across diverse wastewater infrastructures.

4.3. Integrating genome-resolved metagenomics with wastewater-based epidemiology for national AMR surveillance: Strengths, limitations, and future directions

This study highlights the significant strengths of WBE when integrated with genome-resolved metagenomics for AMR surveillance. This combined approach enables comprehensive detection of a wide range of bacterial species, including unculturable and unclassified taxa, while allowing direct linkage of ARGs to their specific microbial hosts within a single sample. Unlike traditional culture-based or PCR-based methods, genome-resolved metagenomics offers a culture-independent, untargeted, and unbiased profiling of microbial communities and their associated resistomes, providing a more holistic view of AMR dynamics in wastewater environments. The study relied on short-read metagenomic sequences for MAG recovery, which offers high accuracy and depth but has limitations such as incomplete recovery of low-abundance taxa. In the future, incorporating long-read sequencing, either alone or in combination with short-reads, could enable the reconstruction of far more complete genomes.

By integrating spatial, seasonal, and genome-level data across diverse hospital and municipal wastewater systems, this study demonstrates the power of genome-resolved WBE to identify site-specific ARG-host associations, reveal uncultured reservoirs of resistance, and assess treatment efficacy. For example, ARGs hosted by *Mycobacterium* persisted even after tertiary treatment. Similarly, resistance genes such as bla_{OXA} and *tet* were consistently detected across treatment stages, suggesting that both specific taxa and ARG classes could serve as robust indicators of residual AMR burden in treated effluents. This approach also enables the differentiation between intrinsic and acquired resistance within key bacterial taxa. Furthermore, the detection of ARGs in unclassified and ultra-small bacterial lineages highlights the value of this approach in uncovering resistance reservoirs that are inaccessible to conventional monitoring. These findings collectively support the inclusion of such methods in regional risk assessments and targeted AMR mitigation strategies.

Despite these strengths, several practical limitations currently hinder the routine integration of this approach into national AMR monitoring programmes. Key challenges include high sequencing costs, lack of standardised protocols, intensive computational requirements, and the need for specialised bioinformatics expertise.

Each step from sample collection to sequencing and data analysis requires careful optimisation to ensure consistent results across time and space. In particular, wastewater sample processing and DNA extraction methods must be standardised across sites to ensure comparability of metagenomic outputs.

Another critical limitation is that DNA-based metagenomic approaches cannot distinguish between viable and non-viable cells. This limits the interpretation of risk from a public health perspective, as the presence of ARGs may not reflect live, infectious organisms. To address this, future studies should consider integrating viability techniques, such as propidium monoazide (PMA) dye method with sequencing, to better assess the presence of live microbial threats.

To fully realise the potential of WBE coupled with genome-resolved metagenomics for AMR surveillance, there is a need to invest in bioinformatics infrastructure, standardise methodologies across surveillance sites and promote training and capacity-building in computational biology. Such steps will facilitate the scaling and harmonisation of WBE-metagenomics-based AMR monitoring, ensuring it can complement and integrate into the ongoing clinical AMR surveillance frameworks.

While practical and technical barriers remain, the integration of this

method into AMR policy frameworks holds great promise for advancing One Health surveillance and protecting public health.

5. Conclusion

The study demonstrates the potential of genome-resolved metagenomics as a scalable, high-resolution approach for AMR surveillance in wastewater environments and provides a foundation for monitoring strategies in Wales and beyond. Our genome-centric framework enables non-invasive, population-level monitoring of AMR, offering a robust tool that bridges clinical and environmental surveillance. Ultimately, our findings highlight the critical role of WBE in national AMR monitoring frameworks, paving the way for targeted interventions, treatment plant upgrades, and the development of One Health strategies to mitigate the spread of resistance across healthcare and community settings.

Our key findings:

- Genome-resolved metagenomics enabled direct linkage of ARGs to specific bacterial hosts in wastewater, including clinically relevant pathogens, commensals, and uncultivated environmental taxa.
- Approximately 14 % of recovered MAGs carried ARGs, with bla_{OXA} and *tet* genes being the most prevalent and persistent through wastewater treatment.
- ARGs exhibited host specificity, with certain resistance genes consistently associated with distinct bacterial taxa.
- *E. coli* and *Klebsiella* spp. were among the key ARG carriers in hospital wastewater, harbouring multiple drug resistance genes, virulence factors, and plasmid replicons.
- Uncultivated bacteria, including members of Patescibacteria, were found to carry ARGs, highlighting overlooked reservoirs of resistance in microbial dark matter.
- A shift in ARG host composition was observed from WWTP influent to effluent, indicating selective pressures imposed by treatment processes.
- Despite some reductions in ARGs and their hosts during treatment, their persistence in effluent and the increasing frequency of CSOs that discharge untreated influents containing ARG hosts pose a continuing environmental risk, exacerbated by climate change.

Future AMR surveillance frameworks should incorporate genome-resolved metagenomics along with traditional monitoring methods to capture the entire spectrum of resistance, including their microbial reservoirs, inform targeted interventions, and guide evidence-based wastewater management and policy.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2025.180607.

CRediT authorship contribution statement

Reshma Silvester: Writing – original draft, Writing – review & editing, Visualization, Project administration, Investigation, Formal analysis, Conceptualization. **Gordon Webster:** Writing – review & editing, Methodology. **William B. Perry:** Writing – review & editing, Data curation. **Kata Farkas:** Writing – review & editing. **Laura Rushton:** Writing – review & editing, Methodology. **Noel Craine:** Writing – review & editing. **Gareth Cross:** Writing – review & editing, Funding acquisition. **Peter Kille:** Writing – review & editing. **Andrew J. Weightman:** Writing – review & editing, Supervision, Funding acquisition. **Davey L. Jones:** Writing – review & editing, Supervision, Funding acquisition.

Funding

The project was funded by the Welsh Government under the Wastewater Monitoring Programme in Wales programme (Contract Number C525/2021–2024), the HM Treasury Shared Outcome Fund under the Pathogen Surveillance in Agriculture, Food and Environment (PATH-SAFE) program Workstream FBD3, and EPSRC Digital Health Hub for Antimicrobial Resistance Grant (EP/X031276/1).

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.

Acknowledgements

We thank Tony Harrington and Ian Trick at Dŵr Cymru Welsh Water for access to sampling sites and subsequent sample collection. We thank all the team members of the Genome hub at Cardiff University for sample processing and Wales Gene Park for help with metagenomic sequencing. We thank the staff at the Wastewater Research Centre Wales for their assistance in sampling and sample processing. This work would not have been possible without the help of the estate departments at the hospitals evaluated in this study and the Health Boards. We thank David Fidler (School of Environmental and Natural Sciences, Bangor University) for the initial guidance on configuring Nextflow to run the pipelines on the Supercomputing Wales (SCW) platform, Daniel Pass (Compass Bioinformatics) for help with initial data processing, Rachel Williams (School of Environmental and Natural Sciences, Bangor University) for help with the data transfer and data curation.

Data availability

Data will be made available on request.

References

Almeida, A., Nayfach, S., Boland, M., Strozzi, F., Beracochea, M., Shi, Z.J., Finn, R.D., 2021. A unified catalog of 204,938 reference genomes from the human gut microbiome. Nat. Biotechnol. 39 (1), 105–114.

Becsei, Á., Fuschi, A., Otani, S., Kant, R., Weinstein, I., Alba, P., Munk, P., 2024. Time-series sewage metagenomics distinguishes seasonal, human-derived and environmental microbial communities potentially allowing source-attributed surveillance. Nat. Commun. 15 (1), 7551.

Bowers, R.M., Kyrpides, N.C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T. B.K., Woyke, T., 2017. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. Nat. Biotechnol. 35 (8), 725–731.

Busi, S.B., de Nies, L., Pramateftaki, P., Bourquin, M., Kohler, T.J., Ezzat, L., Battin, T., 2023. Glacier-fed stream biofilms harbor diverse resistomes and biosynthetic gene clusters. Microbiology spectrum 11 (1) e04069–22.

Crits-Christoph, A., Hallowell, H.A., Koutouvalis, K., Suez, J., 2022. Good microbes, bad genes? The dissemination of antimicrobial resistance in the human microbiome. Gut Microbes 14 (1), 2055944.

Dai, D., Brown, C., Bürgmann, H., Larsson, D.J., Nambi, I., Zhang, T., Vikesland, P.J., 2022. Long-read metagenomic sequencing reveals shifts in associations of antibiotic resistance genes with mobile genetic elements from sewage to activated sludge. Microbiome 10 (1), 20.

Diebold, P.J., Rhee, M.W., Shi, Q., Trung, N.V., Umrani, F., Ahmed, S., Brito, I.L., 2023. Clinically relevant antibiotic resistance genes are linked to a limited set of taxa within gut microbiome worldwide. Nat. Commun. 14 (1), 7366.

Fang, G.Y., Liu, X.Q., Jiang, Y.J., Mu, X.J., Huang, B.W., 2024. Horizontal gene transfer in activated sludge enhances microbial antimicrobial resistance and virulence. Sci. Total Environ. 912, 168908.

Goodarzi, Z., Asad, S., Mehrshad, M., 2022. Genome-resolved insight into the reservoir of antibiotic resistance genes in aquatic microbial community. Sci. Rep. 12 (1), 21047.

Goren, M.G., Carmeli, Y., Schwaber, M.J., Chmelnitsky, I., Schechner, V., Navon-Venezia, S., 2010. Transfer of carbapenem-resistant plasmid from Klebsiella pneumoniae ST258 to Escherichia coli in patient. Emerg. Infect. Dis. 16 (6), 1014.

Hassoun-Kheir, N., Stabholz, Y., Kreft, J.U., De La Cruz, R., Romalde, J.L., Nesme, J., Paul, M., 2020. Comparison of antibiotic-resistant bacteria and antibiotic resistance genes abundance in hospital and community wastewater: a systematic review. Sci. Total Environ. 743, 140804.

Hu, Y., Morichaud, Z., Chen, S., Leonetti, J.P., Brodolin, K., 2012. Mycobacterium tuberculosis RbpA protein is a new type of transcriptional activator that stabilizes the σ A-containing RNA polymerase holoenzyme. Nucleic Acids Res. 40 (14), 6547–6557.

Hultman, J., Tamminen, M., Pärnänen, K., Cairns, J., Karkman, A., Virta, M., 2018. Host range of antibiotic resistance genes in wastewater treatment plant influent and effluent. FEMS Microbiol. Ecol. 94 (4) fiy038.

Javvadi, Y., Mohan, S.V., 2024. Temporal dynamics and persistence of resistance genes to broad spectrum antibiotics in an urban community. npj Clean Water 7 (1), 56.
Kang, D.D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., Wang, Z., 2019. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ 7, e7359.
Kim, J.J., Seong, H.J., Johnson, T.A., Cha, C.J., Sul, W.J., Chae, J.C., 2023. Persistence of antibiotic resistance from animal agricultural effluents to surface water revealed by genome-centric metagenomics. J. Hazard. Mater. 457, 131761.
Krakau, S., Straub, D., Gourlé, H., Gabernet, G., Nahnsen, S., 2022. Nf-core/mag: a best-practice pipeline for metagenome hybrid assembly and binning. NAR Genomics Bioinform. 4 (1), lqac007.
Li, D., Luo, R., Liu, C.M., Leung, C.M., Ting, H.F., Sadakane, K., Lam, T.W., 2016. MEGAHIT v1. 0: a fast and scalable metagenome assembler driven by advanced methodologies and community practices. Methods 102, 3–11.
Naghavi, M., Vollset, S.E., Ikuta, K.S., Swetschinski, L.R., Gray, A.P., Wool, E.E., Dekker, D.M., 2024. Global burden of bacterial antimicrobial resistance 1990–2021: a systematic analysis with forecasts to 2050. Lancet 404 (10459), 1199–1226.
Obayiuwana, A., Ibekwe, A.M., Eze, C., 2025. Resistome Profile of Treated Wastewater Using Metagenomic Approach. Water 17 (6), 867.
Oh, J., Salcedo, D.E., Medriano, C.A., Kim, S., 2014. Comparison of different disinfection processes in the effective removal of antibiotic-resistant bacteria and genes. J. Environ. Sci. 26 (6), 1238–1242. https://doi.org/10.1016/S1001-0742(13)60594-X.
Oluseyi Osunmakinde, C., Selvarajan, R., Mamba, B.B., Msagati, T.A., 2019. Profiling bacterial diversity and potential pathogens in wastewater treatment plants using high-throughput sequencing analysis. Microorganisms 7 (11), 506.
Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., Tyson, G.W., 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 25 (7), 1043–1055.
Paul, D., Das, B., 2022. Gut microbiome in the emergence of antibiotic-resistant bacterial pathogens. Prog. Mol. Biol. Transl. Sci. 192 (1), 1–31.
Perry, W.B., Chrispim, M.C., Barbosa, M.R.F., de Souza Lauretto, M., Razzolini, M.T.P., Nardocci, A.C., Durance, I., 2024. Cross-continental comparative experiences of wastewater surveillance and a vision for the 21st century. Sci. Total Environ. 919, 170842.
Perveen, S., Pablos, C., Reynolds, K., Stanley, S., Marugán, J., 2023. Growth and prevalence of antibiotic-resistant bacteria in microplastic biofilm from wastewater treatment plant effluents. Sci. Total Environ. 856, 159024.
Petriglieri, F., Kondrotaite, Z., Singleton, C., Nierychlo, M., Dueholm, M.K., Nielsen, P.H., 2023. A comprehensive overview of the Chloroflexota community in wastewater treatment plants worldwide. MSystems 8 (6) e00667–23.
Petrovich, M., Chu, B., Wright, D., Griffin, J., Elfeki, M., Murphy, B.T., Wells, G., 2018. Antibiotic resistance genes show enhanced mobilization through suspended growth and biofilm-based wastewater treatment processes. FEMS Microbiol. Ecol. 94 (5) fiy041.
R Core Team, 2021. *R: A language and environment for statistical computing*. Version 4.1.2. R Foundation for Statistical Computing. Vienna, 2021.
Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N.N., Anderson, I.J., Cheng, J.F., Woyke, T., 2013. Insights into the phylogeny and coding potential of microbial dark matter. Nature 499 (7459), 431–437.
Rivers Trust, 2024. Unpacking the 2024 Annual Sewage Spill Data. Available at https://theriverstrust.org/about-us/news/2024-sewage-spill-cso-data#:~:text=But%2C%20looking%20at%20the%20CSO,a%20spill%20was%203.6%20hours. Accessed 01/05/2025.
Seemann, T. Abricate. GitHub https://github.com/tseemann/abricate.
Silvester, R., Perry, W.B., Webster, G., Rushton, L., Baldwin, A., Pass, D.A., Jones, D.L., 2025a. Metagenomics unveils the role of hospitals and wastewater treatment plants on the environmental burden of antibiotic resistance genes and opportunistic pathogens. Sci. Total Environ. 961, 178403.
Silvester, R., Woodhall, N., Nurmi, W., Muziasari, W., Cross, G., Malham, S.K., Jones, D. L., 2025b. High-throughput qPCR profiling of antimicrobial resistance genes and bacterial loads in wastewater and receiving environments. Environ. Pollut., 126096
Singleton, C.M., Petriglieri, F., Kristensen, J.M., Kirkegaard, R.H., Michaelsen, T.Y., Andersen, M.H., Albertsen, M., 15 May 2025. Connecting structure to function with the recovery of over 1000 high-quality metagenome-assembled genomes from activated sludge using long-read sequencing. Nat. Commun. 373, 126096.
Sirichoat, A., Kaewprasert, O., Hinwan, Y., Faksri, K., 2023. Phenotypic drug-susceptibility profiles and genetic analysis based on whole-genome sequencing of Mycobacterium avium complex isolates in Thailand. PLoS One 18 (11), e0294677.
Sivalingam, P., Sabatino, R., Sbaffi, T., Fontaneto, D., Corno, G., Di Cesare, A., 2023. Extracellular DNA includes an important fraction of high-risk antibiotic resistance genes in treated wastewaters. Environ. Pollut. 323, 121325.
Sood, U., Hira, P., Garg, G., Lal, R., Shakarad, M., 2024. Deciphering the microbial dark matter using metagenome-assembled genomes, Culturomics, and Seqcode. In: Microbial Diversity in the Genomic Era. Academic Press, pp. 747–757.
Wang, P., Yu, Z., Qi, R., Zhang, H., 2016. Detailed comparison of bacterial communities during seasonal sludge bulking in a municipal wastewater treatment plant. Water Res. 105, 157–166.
Wei, W., Srinivas, S., Lin, J., Tang, Z., Wang, S., Ullah, S., Feng, Y., 2018. Defining ICR-Mo, an intrinsic colistin resistance determinant from Moraxella osloensis. PLoS Genet. 14 (5), e1007389.
Wu, Y.W., Simmons, B.A., Singer, S.W., 2016. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics 32 (4), 605–607.
Zhao, B., Zhang, R., Jin, B., Yu, Z., Wen, W., Zhao, T., Zhou, J., 2025. Sludge water: a potential pathway for the spread of antibiotic resistance and pathogenic bacteria from hospitals to the environment. Front. Microbiol. 16, 1492128.
Zheng, X., Zhong, Z., Xu, Y., Lin, X., Cao, Z., Yan, Q., 2023. Response of heavy-metal and antibiotic resistance genes and their related microbe in rice paddy irrigated with treated municipal wastewaters. Sci. Total Environ. 896, 165249.
Zhou, M., Guo, Z., Duan, Q., Hardwidge, P.R., Zhu, G., 2014. Escherichia coli type III secretion system 2: a new kind of T3SS? Vet. Res. 45, 1–5.
Zhu, Z., Ren, J., Michail, S., Sun, F., 2019. MicroPro: using metagenomic unmapped reads to provide insights into human microbiota and disease associations. Genome Biol. 20 (1), 154.