Environmental Science and Ecotechnology 23 (2025) 100502

Contents lists available at ScienceDirect

Environmental Science and Ecotechnology

journal homepage: www.journals.elsevier.com/environmental-science-and-ecotechnology/

Original Research

Rapid identification of antibiotic resistance gene hosts by prescreening ARG-like reads

Zhiguo Su [a], April Z. Gu [b], Donghui Wen [c, *], Feifei Li [a], Bei Huang [d], Qinglin Mu [d], Lyujun Chen [a, **]

[a] School of Environment, Tsinghua University, Beijing, 100084, China
[b] School of Civil and Environmental Engineering, Cornell University, Ithaca, NY, 14853, United States
[c] College of Environmental Sciences and Engineering, Peking University, Beijing, 100871, China
[d] Marine Ecological Environmental Monitoring Center of Zhejiang Province, Zhoushan, 316021, China

ARTICLE INFO

Article history:
Received 20 June 2024
Received in revised form
23 October 2024
Accepted 27 October 2024

Keywords:
Antibiotic resistance genes (ARGs)
Host
Metagenomics
Wastewater treatment plants (WWTPs)
Coastal environment

ABSTRACT

Effective risk assessment and control of environmental antibiotic resistance depend on comprehensive information about antibiotic resistance genes (ARGs) and their microbial hosts. Advances in sequencing technologies and bioinformatics have enabled the identification of ARG hosts using metagenome-assembled contigs and genomes. However, these approaches often suffer from information loss and require extensive computational resources. Here we introduce a bioinformatic strategy that identifies ARG hosts by prescreening ARG-like reads (ALRs) directly from total metagenomic datasets. This ALR-based method offers several advantages: (1) it enables the detection of low-abundance ARG hosts with higher accuracy in complex environments; (2) it establishes a direct relationship between the abundance of ARGs and their hosts; and (3) it reduces computation time by approximately 44–96% compared to strategies relying on assembled contigs and genomes. We applied our ALR-based strategy alongside two traditional methods to investigate a typical human-impacted environment. The results were consistent across all methods, revealing that ARGs are predominantly carried by Gammaproteobacteria and Bacilli, and their distribution patterns may indicate the impact of wastewater discharge on coastal resistome. Our strategy provides rapid and accurate identification of antibiotic-resistant bacteria, offering valuable insights for the high-throughput surveillance of environmental antibiotic resistance. This study further expands our knowledge of ARG-related risk management in future.

1. Introduction

Global overuse of antibiotics has promoted the prevalence of antibiotic resistance genes (ARGs) and antibiotic-resistant bacteria (ARB), leading to the ineffectiveness of antibiotic treatments for pathogenic bacterial infections [1,2]. ARGs can proliferate in the microbial community via horizontal gene transfer (HGT) and further increase the number and diversity of ARB across phylogenetic boundaries, thus inducing an acutely growing concern [3,4]. Previous studies indicated that the environmental microbiome acts as the reservoir of ARGs that can be horizontally transferred back to human and animal pathogens [5–7]. Thus, it is essential to determine ARG hosts in the environments when evaluating the potential risk of antibiotic resistance to human health [8] and proposing feasible policies for ARG pollution control [9].

Cultivation-based methods for identifying ARB, common in clinical studies, have limitations due to only a small fraction of ARB isolated [10]. Recent advancements in molecular techniques have led to the exploration of culture-independent methods for identifying ARG hosts, such as correlation analysis, metagenomics, fluorescence-activated cell sorting (FACS), and single-cell fusion polymerase chain reaction (epicPCR) [9,11]. FACS and epicPCR can pinpoint ARG locations at the cellular level but are hindered by complexity and low throughput, limiting their use in large-scale environmental studies. Previous studies mainly utilized correlation analysis and metagenomics to explore ARG–host relationships

* Corresponding author.
** Corresponding author.
E-mail addresses: dhwen@pku.edu.cn (D. Wen), chenlj@tsinghua.edu.cn (L. Chen).

https://doi.org/10.1016/j.ese.2024.100502

in various environments [9]. Correlation analysis determines ARG hosts by assessing correlations between ARG abundance and specific microbial taxa [12,13]. However, while convenient and cheap, it has vital limitations such as high uncertainty or even spurious correlations of ARG–host links, which are susceptible to dataset sample size and normalization methods [9,14]. Metagenomics, involving shotgun sequencing of total DNA, offers comprehensive information on ARGs and microbial communities [15]. Generally, sequenced short reads will be assembled into long-read contigs or drafted into metagenome-assembled genomes (MAGs) to obtain a series of ARG-carrying contigs (ACCs) or MAGs via functional annotation; then taxonomies are assigned to them using phylogenetic biomarkers [16–18]. Nevertheless, this data processing with higher computational requirements is time-consuming, especially for mass environmental datasets. Additionally, the number of unused reads after assembly could increase with the complexity of the microbial community due to the limits of sequence coverage and depth, leading to insufficient recovery of low-abundance genomes [19]. Hence, we hypothesize that ARG-host information can be obtained directly from metagenomic short sequences, and to the best of our knowledge, such ideas and computation methods for identifying ARG hosts have not yet been explored.

In this study, we developed and implemented a metagenomic strategy based on prescreening ARG-like reads (ALRs) to profile the overall composition of ARG hosts in the environment. We used the synthetic and actual metagenomic sequencing data to evaluate and validate the proposed method for the ARG-host identifications. The actual datasets came from engineered and natural systems we have studied before, including wastewater treatment plants (WWTPs), coastal effluent-receiving areas (ERAs), and Hangzhou Bay (HZB), a heavily polluted bay in the East China Sea. Both WWTPs and coastal environments are regarded as hotspots for ARG spread in microbial communities [20]. However, previous studies mainly focused on the distribution and composition of ARGs from WWTPs to coastal environments, while host information on ARGs remains unclear. Uncovering the characteristics of ARG hosts more efficiently could greatly help evaluate human activities' impact on environmental resistome and its potential risk to human health [21]. Based on the ARG detection in our previous studies [22,23], we employed metagenomic sequencing for 11 wastewater and 42 coastal sediment samples. More importantly, we proposed a first-of-its-kind fast bioinformatic pipeline to obtain the composition of ARG hosts directly from metagenomic short reads. Finally, we evaluated the validity of the proposed method by comparing the three ARG-host identification strategies.

2. Materials and methods

2.1. Sample collection and metagenomic sequencing

The sampling sites in the WWTPs and HZB are shown in Figs. S1 and S2 (Supplementary Materials). W1 and W2 treat domestic and industrial wastewater from Shangyu County (the south coast) and Jiaxing City (the north coast). Shangyu (SY) and Jiaxing (JX) ERAs directly receive the treated wastewater from W1 and W2, respectively. HB1–HB10 are sampling sites from the inner bay to the outer bay. In October 2020, 11 wastewater samples were collected and filtered through 0.22 μm membranes for biomass collection. Wastewater samples are directly linked to coastal environments, which represent the potential ecological impacts of WWTPs. In April 2018, 42 surficial sediment samples (0–5 cm deep, single sample at each SY and JX site, triplicate samples at each HB site) were collected using a grab sampler (Van Veen, Hydro-Bios, Germany). All the samples were stored at −80 °C for further analysis. The total DNA extraction and its quality determination were conducted according to a previous study [23]. High-quality DNA was sent to the company (Guangdong Magigene Biotechnology Co., Ltd.) for library construction and 150 bp pair-end shotgun sequencing on an Illumina HiSeq 2500 platform. We obtained 186 Gb (gigabase pairs) of raw data for wastewater samples and 510 Gb for sediment samples. Raw reads were filtered with the KneadData pipeline [24] to obtain 29.7–66.2 million clean reads in these metagenomic datasets.

2.2. ARG-host identification based on prescreening ARG-like reads from metagenomic datasets

We proposed a new strategy for identifying the ARG hosts based on prescreening ALRs from metagenomic short reads (Fig. 1). This strategy included assembly-free (ALR1) and assembly (ALR2) analysis pipelines.

(1) ALR1 pipeline: Clean reads were first searched against the Structured Antibiotic Resistance Genes (SARG, v2.2) database [25] to identify the potential matched reads using UBLAST [26,27] (e-value $\leq 10^{-5}$). Then, these matched reads were further aligned against the SARG database using BLASTX with the default parameters (e-value $\leq 10^{-7}$, sequence identity $\geq$80%, and hit length $\geq$75%) for identifying target reads and ARG classification [25]. The target reads were taxonomically assigned by using Kraken2 (v2.0.8-beta) [28] with the GTDB database (r89), which depended on exact *k-mer* matching and lowest common ancestor (LCA) algorithms. Candidate ARG-carrying taxa with more than ten sequences were retained for further analysis.

(2) ALR2 pipeline: The potential matched reads obtained in the ALR1 pipeline were assembled to contigs (>500 bp) by using MEGAHIT (v1.1.3) [29] with recommended parameters. Prodigal (v2.6.3) [30] with a meta-model was used to predict the open reading frames (ORFs). The protein sequences of ORFs were searched against the SARG database with BLASTP (v2.6.0) under an e-value $\leq 10^{-5}$. The alignment results with identity $\geq$80% and query coverage $\geq$70% were identified as ARG-like ORFs [17]. The relative abundance of genes (Transcripts Per Kilobase Million, TPM) was generated by Salmon (v0.13.1). A contig was considered an ACC if it carried at least one ARG-like ORF. Kraken2 (v2.0.8-beta) [31] with the GTDB database (r89) was used to obtain the taxonomic annotation of the ACCs. CoverM (v0.6.1) was employed to calculate the relative abundance of each ACC (average sequence number).

2.3. ARG-host identification based on metagenomic read assembly

Clean reads of each metagenomic dataset were directly assembled to contigs (>500 bp) via MEGAHIT (v1.1.3). The subsequent steps concerning gene prediction, annotation, quantification, and taxonomic assignment were the same with the ALR2 pipeline.

2.4. ARG-host identification based on metagenomic draft genome assembly

Following the assembly, the MAGs were recovered from contig groups via the metaWRAP pipeline (v1.2.1) [32]. Consolidated MAGs (completeness >50% and contamination <10%) were produced from all the recovered MAGs. dRep (v2.6.2) [33] was applied to obtain a non-redundant MAG set with recommended parameters (-sa 0.95, -nc 0.30). The ARGs carried by MAGs were identified with BLASTP (v2.6.0) and the SARG database (e-value $\leq 10^{-5}$, identity $\geq$80%, and query coverage $\geq$70%). Finally, the taxonomic

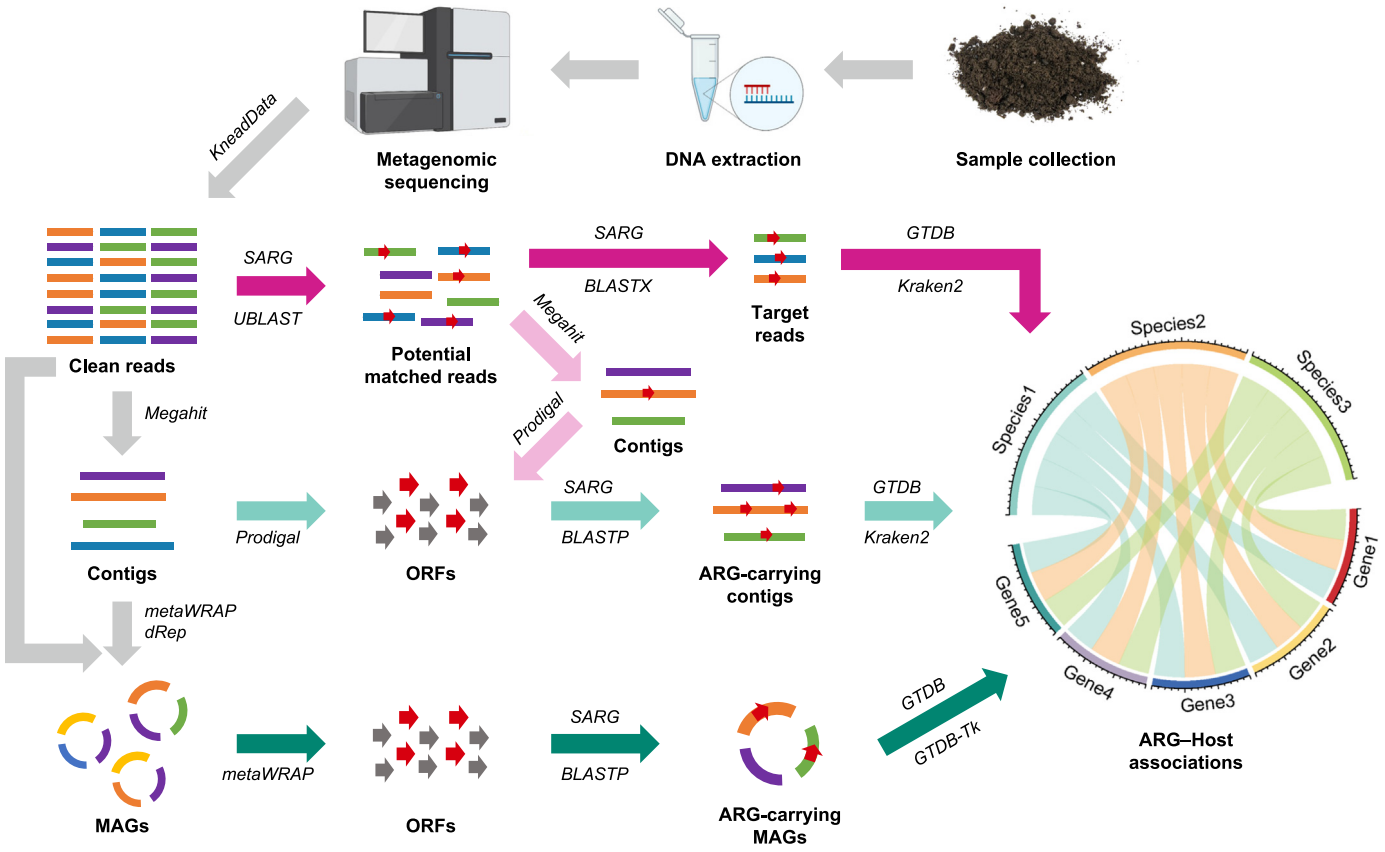

Fig. 1. Analysis pipeline for identifying ARG hosts based on different metagenomic strategies. MAGs, metagenome-assembled genomes. ORFs, open reading frames.

assignment of ARG-carrying MAGs (ACMs) was performed by the GTDB-Tk (v1.0.2) [34]. The relative abundance of ACMs (genome copies/ppm reads) was calculated by the *Quant_bins* command in the metaWRAP pipeline (v1.2.1).

2.5. Identification of pathogenic antibiotic-resistant bacteria

To identify the potential pathogen in the ARG-carrying bacterial community, the taxonomic results of different pipelines were assigned to a bacterial pathogen database, which contained 538 pathogenic species [16].

2.6. Synthetic metagenomic datasets and assessment of different strategies for identifying ARG hosts

Synthetic metagenomic datasets of varying complexity from the Critical Assessment of Metagenomic Interpretation study were employed to test different strategies. The source genomes from "low" (RL), "medium" (RM), and "high" (RH) diversity datasets were first conducted to ARG and taxonomic annotation. These genomes carrying ARGs were regarded as the "gold standard" for ARG host identification pipelines in this study. Then, we analyzed the short reads generated from the RL, RM, and RH datasets through different pipelines to obtain ARG host information. We adopted the following terms and definitions to determine the quality of results: true positive (TP), true negative (TN), false positive (FP), and false negative (FN); positives (P) = TP + FP, negatives (N) = TN + FN, sensitivity = TP/(TP + FN), specificity = TP/P, and accuracy = TP + TN/(P + N).

To further test the performance of different pipelines for identifying the target taxa, different reads were generated from a methicillin-resistant *Staphylococcus aureus* (strain MRSA252, NCBI Accession BX571856) at different genome coverages (100X, 50X, 20X, 10X, 5X, 2X, and 1X) by InSilicoSeq (v1.3.1) [35]. These reads were added to the RH metagenomes and reanalyzed through different pipelines. Three methicillin resistance ARGs (*mecA*, *mecI*, and *mecR1*) in the chromosomal genome of MRSA were detected for host assignment in each pipeline, and the above three were not observed in the original RH metagenomes.

2.7. Visualization

Circos (http://circos.ca/) was applied to visualize the correspondence between the taxa and carried ARGs. OrthoFinder (v2.3.12) [36] was adopted to identify the homology relationships between sequences and to obtain the homologous genes of ACMs. We constructed three maximum likelihood phylogenetic trees using IQ-TREE (v1.6.12) [37] and visualized the trees using iTOL (v5). The co-occurrence network was visualized via Gephi (v0.9.2). Other figures were mainly drawn via the ggplot2 package in R (v3.6.3).

3. Results

3.1. Efficiency of different metagenomic strategies for identifying ARG hosts from synthetic datasets

We assessed standard metagenomic datasets with varying microbial diversities using four ARG-host identification pipelines. Comparisons were made between the pipelines' ARG-host annotations and known information to evaluate identification efficiency. We observed decreasing identification efficiency for ARG hosts as taxonomic annotation levels increased from phylum to species

(Fig. 2). The ALR1 pipeline showed high sensitivity, particularly in low-diversity datasets (RL), with 100% sensitivity at various taxonomic levels. However, its specificity was low due to a high number of false positives, reducing accuracy. In contrast, the ACM pipeline had high specificity but low sensitivity, affecting accuracy. The ALR2 and ACC pipelines demonstrated high sensitivity, specificity, and accuracy for ARG-host identification. Notably, the ALR2 pipeline had the highest accuracy (83.9–88.9%) for ARG-host identification in high-diversity datasets (RH) across different taxonomic levels. To assess the identification efficiency of different pipelines for target ARG hosts, we created a test dataset with ARB of varying abundance levels and analyzed them using four pipelines. Table S1 (Supplementary Material) showed that the ALR1, ALR2, and ACC pipelines annotated the target bacterium (*Staphylococcus aureus*) at phylum, genus, and mainly species levels, while the ACM strategy only provided species-level annotation. The ALR1 pipeline exhibited some identification efficiency for the target bacterium at extremely low abundance (1X), whereas the ALR2 and ACC pipelines detected them at 2X genome coverage, with detections stabilizing after 5X. The ACM pipeline only detected them at 5X coverage. The number of target bacterium detected by ALR1 correlated closely with its genome coverage.

3.2. Composition of ARG hosts in the wastewater and coastal sediment samples

We focused on HZB and surrounding WWTPs to study the connection between ARGs and hosts in wastewater and sediment samples using various metagenomic approaches. With the ALR1 pipeline, we found that potential ARG-matched reads accounted for approximately 0.020–0.171% of all clean sequences in the samples (Supplementary Material Table S2). Using the ALR2 pipeline, we identified 943 contigs with ARGs out of 4540 contigs (Supplementary Material Table S3), and with the ACC pipeline, 1656 contigs out of 2,717,142 contained ARGs (Supplementary Material Table S4). The ACM pipeline revealed 44 MAGs with ARGs out of 1391 non-redundant MAGs (Supplementary Material Table S5). Analysis at the short-read level showed dominant resistance genes like multidrug, bacitracin, macrolide–lincosamide–streptogramin (MLS), sulfonamide, aminoglycoside, β-lactam, and tetracycline, with a notable presence of unclassified resistance genes in ERA and HZB (Supplementary Material Fig. S3). The WWTP samples exhibited significantly higher ARG richness and abundance than the ERA and HZB samples. The ALR1, ALR2, ACC, and ACM pipelines identified 370, 140, 331, and 40 ARG host taxa across all samples, respectively (Supplementary Material Fig. S4). ALR2 had the lowest proportion of uniquely identified taxa (42.9%). The number of ARG hosts identified by the four pipelines in the WWTPs was higher than that of ERAs and HZB, indicating that WWTPs had a higher diversity of ARB (Supplementary Material Table S6). As shown in Table S7 (Supplementary Material), all the ARG-carrying taxa in different samples could be annotated at the phylum level by four pipelines, and at least 61.11% of taxa were annotated at the genus level (HZB samples, ALR2 pipeline). The ALR1, ALR2, and ACC pipelines identified Gammaproteobacteria and Bacilli as primary ARG hosts in various samples (Fig. 3). Gammaproteobacteria were dominant in WWTP, while Bacilli were more prevalent in coastal sediments. Gammaproteobacteria decreased from WWTP to ERA and HZB, while Bacilli increased in ERA according to the ACC pipeline. Many reads in the WWTP samples lacked genus-level

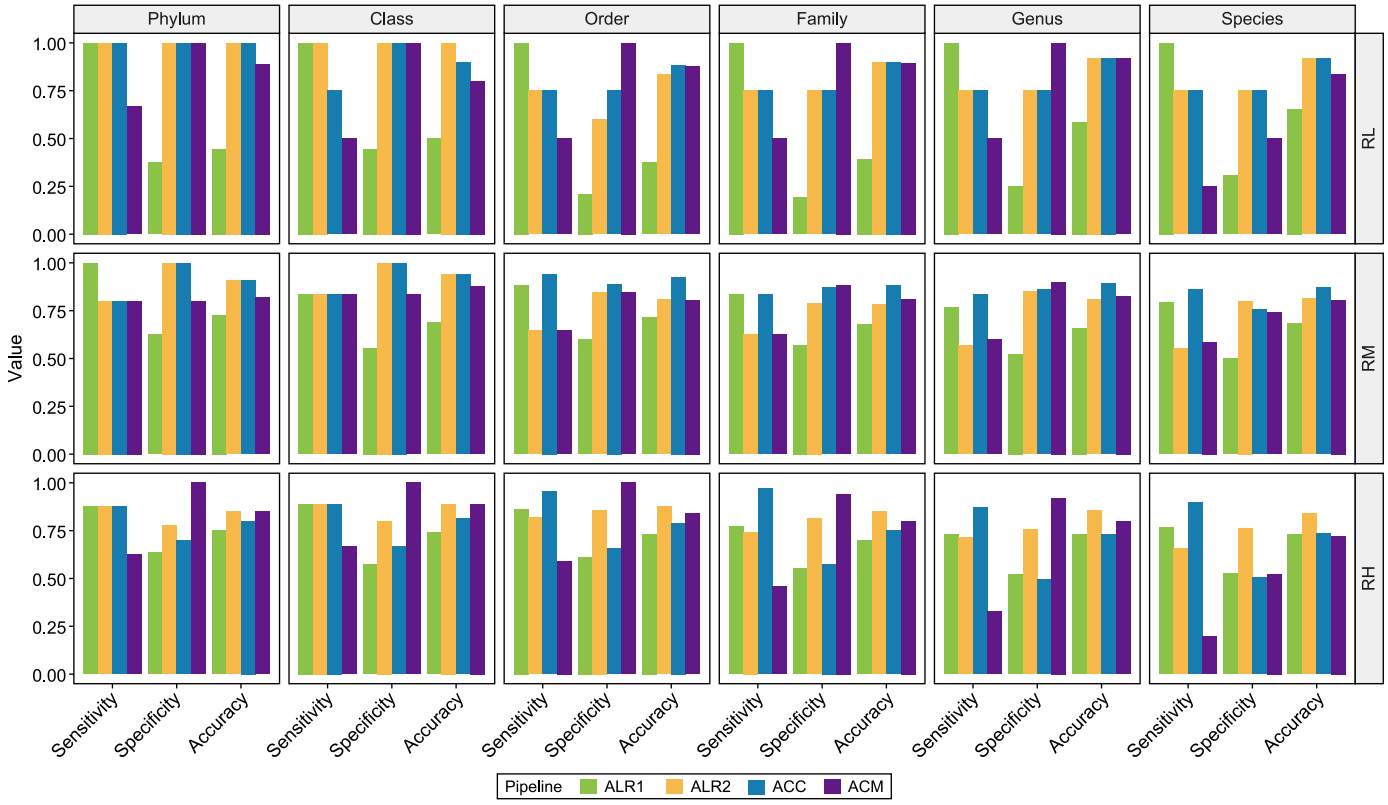

Fig. 2. Performance of different analysis pipelines on identifying ARG hosts at different taxonomic levels from Critical Assessment of Metagenomic Interpretation datasets. RL, RM, and RH are "low," "medium," and "high" diversity datasets, respectively. ALR1, assembly-free analysis pipeline based on prescreening ARG-like reads. ALR2, assembly analysis pipeline based on prescreening ARG-like reads. ACC, analysis pipeline based on assembling ARG-carrying contigs. ACM, analysis pipeline based on binning ARG-carrying metagenome-assembled genomes.

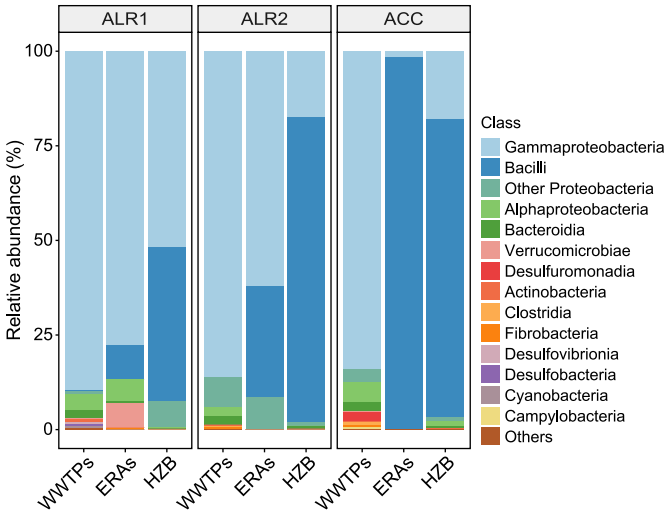

Fig. 3. Composition of ARG hosts at class level in the wastewater and coastal sediment samples based on ALR1, ALR2, and ACC pipeline. WWTPs, wastewater samples in wastewater treatment plants. ERAs, sediment samples in effluent-receiving areas. HZB, sediment samples in Hangzhou Bay. Others, other hosts. ALR1, assembly-free analysis pipeline based on prescreening ARG-like reads. ALR2, assembly analysis pipeline based on prescreening ARG-like reads. ACC, analysis pipeline based on assembling ARG-carrying contigs.

taxonomic information. *Pseudomonas_E* was the major ARG-carrying genus in the WWTPs, while *Staphylococcus* was dominant in coastal sediments (Supplementary Material Fig. S5). The ACM pipeline detected 44 ACMs, with 27 being Gammaproteobacteria, mainly in WWTPs (Fig. 5a; Supplementary Material Table S8). Two ACMs identified as Bacilli in ERA and HZB carried the highest number of ARGs (17 and 15, respectively).

3.3. The corresponding relationships between ARGs and their hosts

Fig. 4 illustrates the relationship between ARG composition and host information in different environmental samples. The ALR1, ALR2, and ACC pipelines observed a decrease in the number of ARGs carried by Gammaproteobacteria from WWTPs to ERAs and HZB, while Bacilli showed an increasing trend. In WWTPs, Gammaproteobacteria predominantly carried genes for multidrug, MLS, beta-lactam, aminoglycoside, sulfonamide, and bacitracin resistance. In ERAs and HZB, Gammaproteobacteria mainly carried multidrug resistance genes, while Bacilli carried multidrug, beta-lactam, tetracycline, and bacitracin resistance genes. Assembly-based methods (ALR2 and ACC pipelines) effectively identified host information for aminoglycoside and sulfonamide resistance genes in the WWTP samples. The ACM pipeline detected 63 ARG subtypes in 44 ACMs, with multidrug, beta-lactam, aminoglycoside, MLS, and tetracycline resistance genes being the most common. The *bacA* gene, conferring bacitracin resistance, was most frequently detected among the 44 ACMs (40.9%, Fig. 5b), indicating its stable presence in various ARB [38]. Multiple ACMs in WWTP effluent carried multidrug resistance genes like *multidrug-ABC-transporter*, *mexD*, and *mexF*, along with high-risk resistance genes such as *tetX*, *mcr-5*, *aph(6)-I*, *catB*, and *floR*. Coastal sediment *Staphylococcus aureus* genomes harbored beta-lactam resistance genes *mecA*, *mecI*, and *mecR1*, suggesting resistance to methicillin antibiotics. These ACMs also carried other high-risk resistance genes (*emrB-qacA*, *mepA*, *norA*, *tetM*, and *ermB*), posing a potential "superbug" (MRSA) threat to human health.

3.4. Identification of potential pathogenic antibiotic-resistant bacteria

Pathogenic antibiotic-resistant bacteria (PARB) enriched with ARGs in the environment pose a significant threat to human health when they spread to clinical settings [39]. This study compared ARG-host annotation information at the species level with a list of common pathogenic bacteria to identify potential PARB in various samples. Fig. 6a exhibits 26 potential PARB across different samples, primarily belonging to Gammaproteobacteria. The ALR1 and ACC pipelines identified the highest number of PARB, followed by the ALR2 and ACM pipelines. The highest number of PARB was identified in WWTPs, followed by ERAs and HZB. The main PARB in WWTPs included *Ralstonia pickettii*, *Aeromonas caviae*, *Legionella pneumophila*, *Klebsiella pneumoniae*, *Bacteroides uniformis*, and *Aeromonas hydrophila*. In ERAs, *Staphylococcus aureus* and *Enterobacter cloacae* were prominent, while *Staphylococcus aureus* was prevalent in HZB. The study also analyzed the relative abundance of PARB and their carried ARGs in different samples. Fig. 6b shows that the relative proportion of PARB and their ARGs in WWTPs was relatively low across all four analysis pipelines. The ALR1 and ALR2 pipelines indicated an increase in PARB from WWTPs to ERAs and HZB, along with a similar trend in the relative abundance of carried ARGs. The ACC pipeline revealed that while PARB in ERAs constituted 98.1% of the total ARG-host abundance, only 11.5% carried ARGs. The relative abundance of PARB and ARGs they carried in HZB accounted for more than 50%. The ACM pipeline highlighted that PARB and their carried ARGs in ERA samples accounted for 100%, as only one ACM identified was PARB.

4. Discussion

4.1. Advantages of ALR-based metagenomic strategies for ARG-host identification

Compared to the established ACC- and ACM-based methods, our new strategy based on ALRs showed the following three advantages: First, although all three strategies can quantify ARG host abundance, only the ALR-based strategy enables the covering of ARG hosts with low abundance. The assembly process of sequencing data and the complexity of the dataset significantly influence the performance of four identification pipelines. ALR1 bypasses the assembly process and directly annotates ARG-like sequences, enabling the detection of extremely low-abundance hosts with high sensitivity (Fig. 2; Supplementary Material Table S1). However, this pipeline also exhibits higher false positives, reducing specificity in host identification. On the other hand, the metagenomic binning-based pipeline (ACM) aims to reconstruct complete bacterial genomes, resulting in relatively conservative host identification with high specificity but limited comprehensive detection of hosts, particularly in samples with high microbial diversity. This pipeline faces challenges in identifying low-abundance species or genes due to the limitation of sequencing depth [15], thereby impacting the final accuracy of host identification. The ALR2 and ACC pipelines offer more balanced solutions; both assemble short sequences into longer ones (contigs) to optimize sensitivity and specificity for host identification, yielding high identification accuracy (Fig. 2). Nevertheless, ALR2 mitigates the influence of numerous irrelevant sequences before assembly, not only saving computational time but also enhancing detection accuracy, as evidenced in RH datasets (Fig. 2). Therefore, the ALR2 pipeline may provide a comprehensive understanding of the ARG-host information in the complex environmental samples.

Second, our ALR-based strategy directly correlates ARG abundance with ARG hosts. In our study, the ACC-based pipeline in the

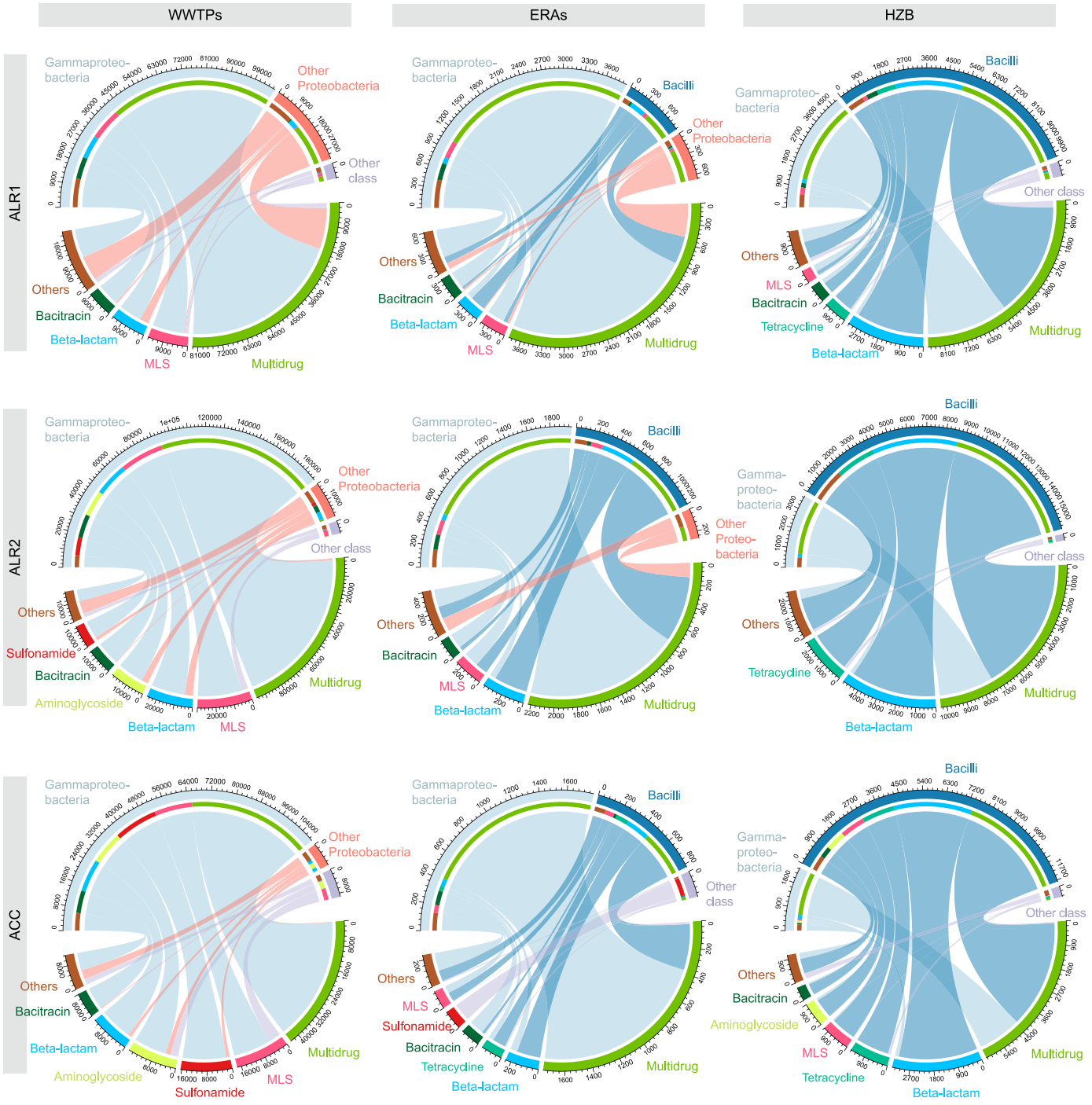

Fig. 4. The corresponding relationship between taxonomic (class level) and functional annotation in the wastewater and coastal sediment samples based on ALR1, ALR2, and ACC pipeline. MLS, macrolide–lincosamide–streptogramin. Others, other ARG types. WWTPs, wastewater samples in wastewater treatment plants. ERAs, sediment samples in effluent-receiving areas. HZB, sediment samples in Hangzhou Bay. ALR1, assembly-free analysis pipeline based on prescreening ARG-like reads. ALR2, assembly analysis pipeline based on prescreening ARG-like reads. ACC, analysis pipeline based on assembling ARG-carrying contigs.

ERA samples identified potential PARB accounting for 98.1% of total ARG hosts, but only 11.5% of total ARG abundance was attributed to ARGs carried by PARB (Fig. 6b). This mismatched result indicated that some host bacteria with high relative abundance might not carry a higher abundance of ARGs. Thus, assembling many unrelated sequences in this pipeline might prevent us from gaining insight into the actual contribution of the host to ARGs. By annotating ARG sequences without interference from unrelated information, our approach avoids this problem to a certain extent, as it captures as many hosts as possible, while the relative abundance of hosts also reflects the occurrence of ARGs.

Third, the ALR-based strategy requires significantly less computational time and resources than the ACC and ACM strategies. In this study, to extract ARG-host information from synthetic and actual datasets, ALR-based pipelines (especially ALR2) showed lower computation hours requirements than the ACC and ACM pipelines (Supplementary Material Table S9). We also found that ALR-based pipelines demonstrated greater advantages in high-

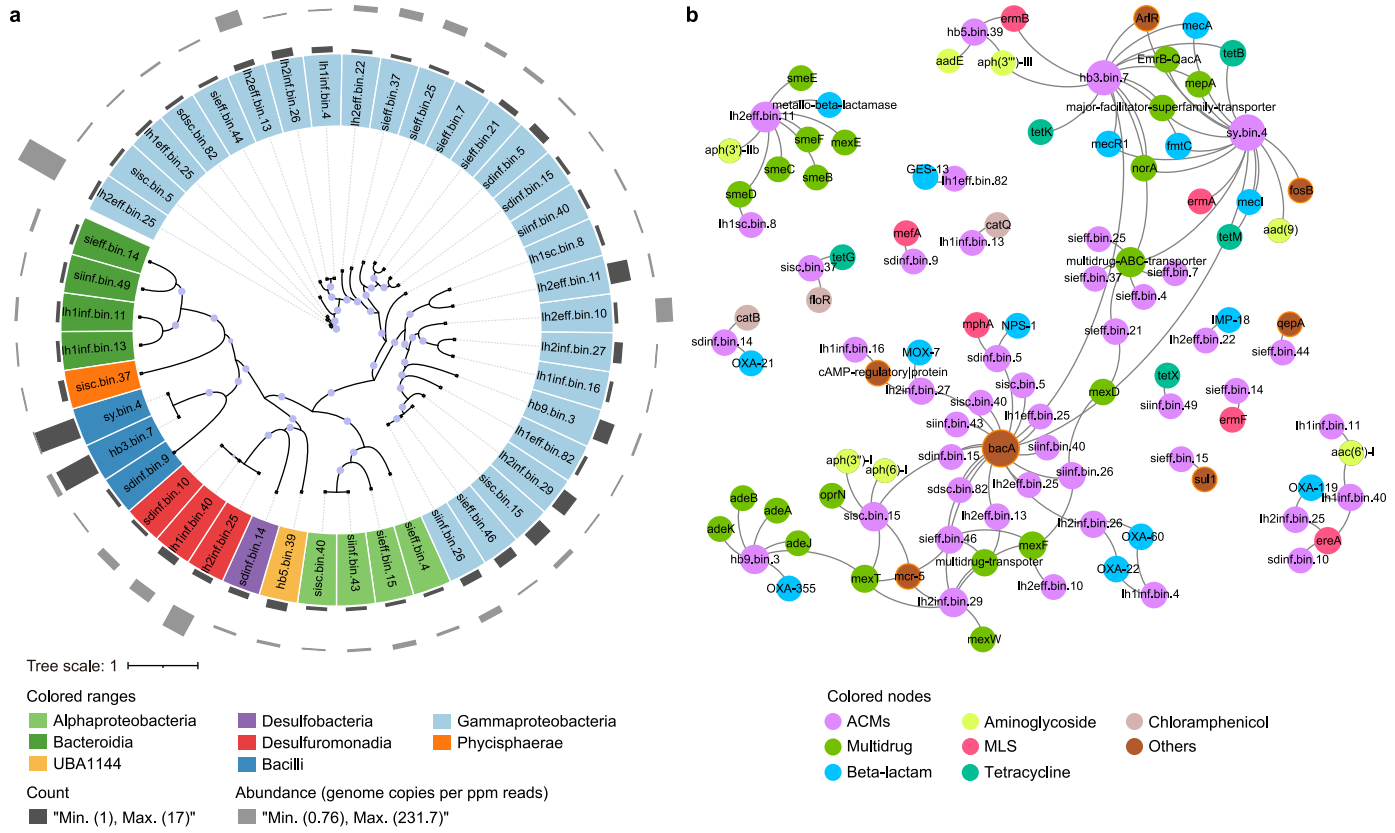

Fig. 5. **a**, Phylogenetic assignment of ARG hosts based on ACM pipeline. **b**, Co-occurrence network between ACMs and their carried ARGs. The nodes with different colors represent ACMs and different ARG types. The node size represents the node degree. The connecting edge between two nodes indicates the affiliation. For instance, *sul1* was linked to sieff.bin.15, indicating that sieff.bin.15 harbors the resistance gene *sul1*. The width of the edge corresponds to the number of individual host-ARG pairs. ACMs, ARG-carrying metagenome-assembled genomes. MLS, macrolide–lincosamide–streptogramin. Others, other ARG types. ACM, analysis pipeline based on binning ACMs.

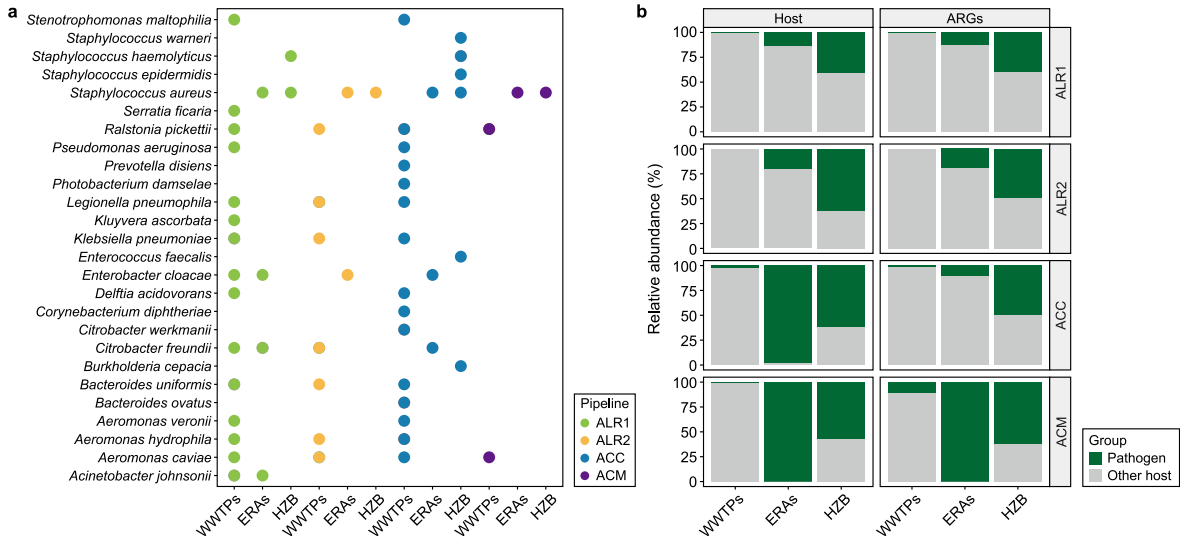

Fig. 6. **a**, Identification of pathogenic antibiotic-resistant bacteria (PARB) based on different analysis pipelines. **b**, The relative abundance of PARB and their carried ARGs in the wastewater and coastal sediment samples. WWTPs, wastewater samples in wastewater treatment plants. ERAs, sediment samples in effluent-receiving areas. HZB, sediment samples in Hangzhou Bay. ALR1, assembly-free analysis pipeline based on prescreening ARG-like reads. ALR2, assembly analysis pipeline based on prescreening ARG-like reads. ACC, analysis pipeline based on assembling ARG-carrying contigs. ACM, analysis pipeline based on binning ARG-carrying metagenome-assembled genomes.

complexity datasets (RH and wastewater samples), saving computational time by 67–96% (Supplementary Material Table S9). The latter two strategies were time-consuming and computationally intensive due to complex analysis steps for many irrelevant sequences [15,40]. In the new ALR-based pipeline, by prescreening the potential ARG reads, more than 99% of irrelevant sequences were removed [41]. This allows for the direct and accurate identification of ARG hosts in a significantly reduced time frame, even

with sequence assembly steps in the ALR2 pipeline. Kraken, utilizing *k*-mers and the LCA algorithm, offers rapid and accurate taxonomic information assignment for metagenomic short reads, surpassing classifiers like Megablast and MetaPhlAn [31]. Furthermore, it enhances speed fivefold and reduces memory usage by 85% [28], enhancing the computational efficiency of taxonomic annotation for ALRs.

In summary, the ALR-based strategy offers benefits such as speed, comprehensiveness (including extremely low-abundance hosts), and enhanced accuracy for ARG-host identification in complex datasets. This may be more suitable for rapid and comprehensive ARG-host screening of environmental samples. If other researchers focus on habitats with low microbial diversity or aim to study the genomic traits of specific major ARG hosts and the genetic background of ARGs, the combination of the ALR- and ACM-based strategies may be a good choice.

4.2. Identification of primary ARG hosts in human-impacted environments

From a "One Health" perspective, WWTP effluent discharge significantly impacts AMR development in coastal environments, potentially leading to the transfer of coastal ARGs or ARB to humans and posing a risk to public health [42]. Effective tools are needed to rapidly and accurately identify small numbers of ARB in these diverse and complex bacterial communities [43]. This study used the new ALR-based metagenomic strategy to analyze ARG-host composition in WWTPs and coastal environments, comparing its efficacy with other metagenomic approaches.

All metagenomic strategies confirmed that ARGs were mainly present in Gammaproteobacteria and Bacilli (Figs. 3 and 5a). This observation can be attributed to (1) the dominance of Gammaproteobacteria and Bacilli in the study area [44], (2) the majority of antibiotic-consuming bacteria within these classes displaying intrinsic resistance to multiple antibiotics [45], and (3) the close relationship between clinical human pathogens within these classes and ARGs (Fig. 6b) [3,9,17]. Comparing the two classes, Gammaproteobacteria contributed a higher diversity of ARB than Bacilli in this study. For Gammaproteobacteria, many diverse genera, such as *Acinetobacter*, *Pseudomonas*, *Citrobacter*, *Escherichia*, *Enterobacter*, and *Thauera*, contributed a high proportion to ARGs, especially in the WWTP samples. These microbes are closely associated with clinical diseases, fecal pollution, and biological wastewater treatment processes [42,46,47]. For Bacilli, *Staphylococcus aureus* emerged as the supercarrier of ARGs in the coastal sediment samples, which have been proven to adapt to high-salt conditions and exhibit heightened antibiotic resistance [48,49]. Regarding the ARG types carried by both classes, multidrug resistance genes, which play important roles in regulating the expression of antibiotic resistance and the transport of cellular metabolites [50,51], were dominant ARGs in both Gammaproteobacteria and Bacilli. Other ARG types like MLS, sulfonamide, aminoglycoside, beta-lactam, and tetracycline resistance genes presented differently in the two classes (Fig. 4). The ACM-based results also revealed that Gammaproteobacteria and Bacilli carried different types of ARGs (Fig. 5b), indicating that the horizontal transfer of ARGs might be impeded between them. This supports a previous hypothesis that phylogeny rather than HGT determines the composition of environmental resistome effectively [13,51]. Nevertheless, the ARG-carrying characteristics of environmental ARB might also have local ecological relevance [52]. Moreover, within the same class or habitat, more potential transfer events of high-risk ARGs between non-pathogenic and pathogenic species should be given great concern (Fig. 5b).

Notably, the relative composition of ARG hosts annotated as Gammaproteobacteria and Bacilli varied significantly in the WWTP, ERA, and HZB samples, suggesting its indicative value for health risks and the ecological impacts of anthropogenic activities (i.e., wastewater discharge). The WWTP samples exhibited higher diversity of both ARGs and ARB compared to the ERA and HZB samples, highlighting WWTPs as significant sources of ARGs and ARB in coastal environments [23,42]. Our previous study revealed that wastewater discharge could drive the spread of ARGs and boost the antibiotic resistome risk in coastal areas [53]. The relative contribution of Gammaproteobacteria to ARGs decreased from WWTPs to ERA and HZB, which could also indicate the influence of wastewater discharge on antibiotic resistome characteristics in coastal environments. More specifically, in the effluent-receiving waterbodies, the presence of certain PARB belonging to Gammaproteobacteria (e.g., *Acinetobacter* sp. And *Pseudomonas* sp.) were regarded as the indicators of wastewater-related pollution [54–56]. The truth behind it was that wastewater-borne PARB normally showed multiple resistance to harsh conditions and persistently spread in effluent-receiving environments [42,57]. Meanwhile, wastewater effluent could decrease the diversity of microbial communities in coastal areas [58], thereby reducing the degree of niche overlap and allowing PARB to achieve competitive advantages [59]. On the other hand, the occurrence and distribution of ARG hosts also varied with natural processes [47]. In less impacted areas like HZB, most ARGs are enriched in specific microbes (*Staphylococcus aureus*), suggesting limitations on ARG spread imposed by phylogeny and natural factors [42,53]. A recent study identified that some human pathogen bacteria, harboring many ARG combinations, were natural supercarriers in a river basin [60]. Therefore, the health risks of ARG hosts in natural environments should not be ignored.

4.3. Technological limitations and future perspective

We acknowledge a few limitations of the ALR-based strategy. First, short sequences (150 bp) obtained through next-generation sequencing provide limited genetic information, leading to more false positives in species and gene annotation. To overcome this, we introduced the ALR2 pipeline, incorporating sequence assembly for improved accuracy compared to the ALR1 pipeline. Second, the ALR-based strategy, particularly the ALR1 pipeline, struggles with identifying hosts for highly mobile ARGs like aminoglycoside and sulfonamide resistance and faces challenges in detecting multidrug-resistant bacteria with multiple ARGs (Fig. 4) [61]. Third, unlike ACC- and ACM-based strategies, the ALR-based approach encounters difficulties in further analyzing the functional information within ARGs' flanking sequences and identifying the genomic characteristics of ARG–host interactions, hindering comprehensive evaluation of environmental risks posed by ARB. We must acknowledge that the sequence-based phylogenetic assignment to the species level of all the pipelines is often challenging, especially for identifying potential pathogens. Frequent horizontal transfer of ARGs within microbial communities might also lead to some mismatches between ARGs and their hosts based on metagenomic sequencing methods.

Nowadays, the emergence of third-generation sequencing technology allows one to obtain increasingly longer sequence lengths, potentially eliminating the need for assembly. High-throughput chromatin conformation capture (Hi-C) metagenome sequencing has also been suggested to link mobile genes with bacterial genomes more broadly [62]. These protocols provide new avenues for improving the identification of ARG hosts. However, new techniques are still under development, especially for environmental monitoring, and we must recognize that they still have the shortcomings of sequencing errors, operational complexity, and

high costs. Currently, metagenomics based on second-generation sequencing technology continues to dominate relevant studies. Therefore, thoroughly exploring the analysis methods using second-generation sequencing data remains crucial. With the growing genomic data and bioinformatics advancements, studies have delved into the phylogenetic affiliation of ARGs, making understanding clearer and more stable [3,63,64]. Thus, by leveraging the advantages of the ALR-based strategy and increasing relevant empirical knowledge, the wide and large-scale environmental screening of ARG hosts can be achieved to support the implementation of the "One Health" framework.

5. Conclusion

This study proposed a new metagenomic strategy to identify ARG hosts based on prescreening ALRs. We compared the new strategy with two assembly-based metagenomic strategies by synthesizing standardized datasets. The results showed that the new strategy demonstrated advantages in detection accuracy and computation time for identifying ARG hosts in high-complexity datasets and could establish direct connections between the abundance of ARGs and their hosts. All metagenomic strategies confirmed that ARGs were mainly carried by Gammaproteobacteria and Bacilli in the WWTPs and coastal sediments. Their relative composition varied significantly in the different environmental samples, suggesting their indicative value for health risks and the ecological impacts of wastewater discharge. Our proposed strategy could contribute to a more comprehensive understanding of the risk of environmental antibiotic resistome and provide new insights about assessing the impact of anthropogenic activities on environments.

CRediT authorship contribution statement

Zhiguo Su: Writing - Original Draft, Visualization, Software, Methodology, Funding Acquisition, Formal Analysis, Data Curation, Conceptualization. **April Z. Gu:** Writing - Review & Editing, Supervision. **Donghui Wen:** Writing - Review & Editing, Supervision, Funding Acquisition, Conceptualization. **Feifei Li:** Writing - Review & Editing. **Bei Huang:** Resources, Data Curation. **Qinglin Mu:** Resources, Data Curation. **Lyujun Chen:** Writing - Review & Editing, Supervision, Resources, Funding Acquisition.

Data availability

Raw data were deposited in the NCBI Sequence Read Archive database with accession numbers PRJNA718650 and PRJNA718910.

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

This study was supported by the National Natural Science Foundation of China (Grant No. 52300243, 52170185, 51938001, and 52070111), the China Postdoctoral Science Foundation (No.2022M721815), the “Shuimu Tsinghua Scholar” Program of Tsinghua University, the National Treasure Ecological Research Synergetic Innovation Center, and High-Performance Computing Platform of Peking University. We appreciate the great help in sample collection by the Zhejiang Marine Ecology and Environment Monitoring Center.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ese.2024.100502.

References

[1] E.C. Pehrsson, P. Tsukayama, S. Patel, M. Mejía-Bautista, G. Sosa-Soto, K.M. Navarrete, M. Calderon, L. Cabrera, W. Hoyos-Arango, M.T. Bertoli, D.E. Berg, R.H. Gilman, G. Dantas, Interconnected microbiomes and resistomes in low-income human habitats, Nature 533 (2016) 212–216, https://doi.org/10.1038/nature17672.
[2] K. Qin, L. Wei, J. Li, B. Lai, F. Zhu, H. Yu, Q. Zhao, K. Wang, A review of ARGs in WWTPs: sources, stressors and elimination, Chin. Chem. Lett. 31 (2020) 2603–2613, https://doi.org/10.1016/j.cclet.2020.04.057.
[3] Y. Hu, X. Yang, J. Li, N. Lv, F. Liu, J. Wu, I.Y.C. Lin, N. Wu, B.C. Weimer, G.F. Gao, Y. Liu, B. Zhu, The bacterial mobile resistome transfer network connecting the animal and human microbiomes, Appl. Environ. Microbiol. 82 (2016) 6672–6681, https://doi.org/10.1128/AEM.01802-16.
[4] C.M. Thomas, K.M. Nielsen, Mechanisms of, and barriers to, horizontal gene transfer between bacteria, Nat. Rev. Microbiol. 3 (2005) 711–721, https://doi.org/10.1038/nrmicro1234.
[5] K.J. Forsberg, A. Reyes, B. Wang, E.M. Selleck, M.O.A. Sommer, G. Dantas, The shared antibiotic resistome of soil bacteria and human pathogens, Science 337 (2012) 1107–1111, https://doi.org/10.1126/science.1220761.
[6] X. Jiang, M.M.H. Ellabaan, P. Charusanti, C. Munck, K. Blin, Y. Tong, T. Weber, M.O.A. Sommer, S.Y. Lee, Dissemination of antibiotic resistance genes from antibiotic producers to pathogens, Nat. Commun. 8 (2017) 15784, https://doi.org/10.1038/ncomms15784.
[7] P.M.C. Huijbers, H. Blaak, M.C.M. de Jong, E.A.M. Graat, C.M.J.E. Vandenbroucke-Grauls, A.M. de Roda Husman, Role of the environment in the transmission of antimicrobial resistance to humans: a review, Environ. Sci. Technol. 49 (2015) 11993–12004, https://doi.org/10.1021/acs.est.5b02566.
[8] J. Hultman, M. Tamminen, K. Pärnänen, J. Cairns, A. Karkman, M. Virta, Host range of antibiotic resistance genes in wastewater treatment plant influent and effluent, FEMS Microbiol. Ecol. 94 (2018), https://doi.org/10.1093/femsec/fiy038.
[9] E.W. Rice, P. Wang, A.L. Smith, L.B. Stadler, Determining hosts of antibiotic resistance genes: a review of methodological advances, Environ. Sci. Technol. Lett. 7 (2020) 282–291, https://doi.org/10.1021/acs.estlett.0c00202.
[10] C.S. Riesenfeld, P.D. Schloss, J. Handelsman, Metagenomics: genomic analysis of microbial communities, Annu. Rev. Genet. 38 (2004) 525–552, https://doi.org/10.1146/annurev.genet.38.072902.091216.
[11] A.Q. Nguyen, H.P. Vu, L.N. Nguyen, Q. Wang, S.P. Djordjevic, E. Donner, H. Yin, L.D. Nghiem, Monitoring antibiotic resistance genes in wastewater treatment: current strategies and future challenges, Sci. Total Environ. 783 (2021) 146964, https://doi.org/10.1016/j.scitotenv.2021.146964.
[12] M. Layeghifard, D.M. Hwang, D.S. Guttman, Disentangling interactions in the microbiome: a network perspective, Trends Microbiol. 25 (2017) 217–228, https://doi.org/10.1016/j.tim.2016.11.008.
[13] K.J. Forsberg, S. Patel, M.K. Gibson, C.L. Lauber, R. Knight, N. Fierer, G. Dantas, Bacterial phylogeny structures soil resistomes across habitats, Nature 509 (2014) 612–616, https://doi.org/10.1038/nature13377.
[14] W. Shi, Y. Liu, J. Li, H. Zhang, R. Shi, J. Chen, H. Li, Distribution pattern of antibiotic resistance genes and bacterial community in agricultural soil samples of Wuliangsuhai watershed, China, Agriculture, Ecosystems & Environment 295 (2020) 106884, https://doi.org/10.1016/j.agee.2020.106884.
[15] M. Boolchandani, A.W. D Souza, G. Dantas, Sequencing-based methods and resources to study antimicrobial resistance, Nat. Rev. Genet. 20 (2019) 356–370, https://doi.org/10.1038/s41576-019-0108-4.
[16] B. Li, F. Ju, L. Cai, T. Zhang, Profile and fate of bacterial pathogens in sewage treatment plants revealed by high-throughput metagenomic approach, Environ. Sci. Technol. 49 (2015) 10492–10502, https://doi.org/10.1021/acs.est.5b02345.
[17] L. Ma, Y. Xia, B. Li, Y. Yang, L. Li, J.M. Tiedje, T. Zhang, Metagenomic assembly reveals hosts of antibiotic resistance genes and the shared resistome in pig, chicken, and human feces, Environ. Sci. Technol. 50 (2016) 420–427, https://doi.org/10.1021/acs.est.5b03522.
[18] R. Zhao, K. Yu, J. Zhang, G. Zhang, J. Huang, L. Ma, C. Deng, X. Li, B. Li, Deciphering the mobility and bacterial hosts of antibiotic resistance genes under antibiotic selection pressure by metagenomic assembly and binning approaches, Water Res. 186 (2020) 116318, https://doi.org/10.1016/j.watres.2020.116318.
[19] C. Quince, A.W. Walker, J.T. Simpson, N.J. Loman, N. Segata, Shotgun metagenomics, from sampling to analysis, Nat. Biotechnol. 35 (2017) 833–844, https://doi.org/10.1038/nbt.3935.
[20] Y. Zhu, Y. Zhao, B. Li, C. Huang, S. Zhang, S. Yu, Y. Chen, T. Zhang, M.R. Gillings, J. Su, Continental-scale pollution of estuaries with antibiotic resistance genes, Nat. Microbiol. 2 (2017), https://doi.org/10.1038/nmicrobiol.2016.270.
[21] P. Fresia, V. Antelo, C. Salazar, M. Giménez, B. D Alessandro, E. Afshinnekoo, C. Mason, G.H. Gonnet, G. Iraola, Urban metagenomics uncover antibiotic resistance reservoirs in coastal beach and sewage waters, Microbiome 7

(2019), https://doi.org/10.1186/s40168-019-0648-z.
[22] J. Chen, Z. Su, T. Dai, B. Huang, Q. Mu, Y. Zhang, D. Wen, Occurrence and distribution of antibiotic resistance genes in the sediments of the East China Sea bays, J. Environ. Sci. 81 (2019) 156–167, https://doi.org/10.1016/j.jes.2019.01.016.
[23] Z. Su, A. Li, J. Chen, B. Huang, Q. Mu, L. Chen, D. Wen, Wastewater discharge drives ARGs spread in the coastal area: a case study in Hangzhou Bay, China, Mar. Pollut. Bull. 151 (2020) 110856, https://doi.org/10.1016/j.marpolbul.2019.110856.
[24] Y. Liu, Y. Qin, T. Chen, M. Lu, X. Qian, X. Guo, Y. Bai, A practical guide to amplicon and metagenomic analysis of microbiome data, Protein & Cell 12 (2020) 315–330, https://doi.org/10.1007/s13238-020-00724-8.
[25] X. Yin, X. Jiang, B. Chai, L. Li, Y. Yang, J.R. Cole, J.M. Tiedje, T. Zhang, ARGs-OAP v2.0 with an expanded SARG database and Hidden Markov Models for enhancement characterization and quantification of antibiotic resistance genes in environmental metagenomes, Bioinformatics 34 (2018) 2263–2270, https://doi.org/10.1093/bioinformatics/bty053.
[26] R.C. Edgar, Search and clustering orders of magnitude faster than BLAST, Bioinformatics 26 (2010) 2460–2461, https://doi.org/10.1093/bioinformatics/btq461.
[27] Y. Yang, X. Jiang, T. Zhang, C.A. Ouzounis, Evaluation of a hybrid approach using UBLAST and BLASTX for metagenomic sequences annotation of specific functional genes, PLoS One 9 (2014) e110947, https://doi.org/10.1371/journal.pone.0110947.
[28] D.E. Wood, J. Lu, B. Langmead, Improved metagenomic analysis with Kraken 2, Genome Biol. 20 (2019), https://doi.org/10.1186/s13059-019-1891-0.
[29] D. Li, C.M. Liu, R. Luo, K. Sadakane, T.W. Lam, MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph, Bioinformatics 31 (2015) 1674–1676, https://doi.org/10.1093/bioinformatics/btv033.
[30] D. Hyatt, G.L. Chen, P.F. Locascio, M.L. Land, F.W. Larimer, L.J. Hauser, Prodigal: prokaryotic gene recognition and translation initiation site identification, BMC Bioinf. 11 (2010) 119, https://doi.org/10.1186/1471-2105-11-119.
[31] D.E. Wood, S.L. Salzberg, Kraken: ultrafast metagenomic sequence classification using exact alignments, Genome Biol. 15 (2014) R46, https://doi.org/10.1186/gb-2014-15-3-r46.
[32] G.V. Uritskiy, J. DiRuggiero, J. Taylor, MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis, Microbiome 6 (2018), https://doi.org/10.1186/s40168-018-0541-1.
[33] M.R. Olm, C.T. Brown, B. Brooks, J.F. Banfield, dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication, ISME J. 11 (2017) 2864–2868, https://doi.org/10.1038/ismej.2017.126.
[34] D.H. Parks, M. Chuvochina, D.W. Waite, C. Rinke, A. Skarshewski, P. Chaumeil, P. Hugenholtz, A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life, Nat. Biotechnol. 36 (2018) 996–1004, https://doi.org/10.1038/nbt.4229.
[35] H. Gourle, O. Karlsson-Lindsjo, J. Hayer, E. Bongcam-Rudloff, Simulating Illumina metagenomic data with InSilicoSeq, Bioinformatics 35 (2019) 521–522, https://doi.org/10.1093/bioinformatics/bty630.
[36] D.M. Emms, S. Kelly, OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy, Genome Biol. 16 (2015), https://doi.org/10.1186/s13059-015-0721-2.
[37] N. Lam-Tung, H.A. Schmidt, V.H. Arndt, M.B. Quang, IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies, Mol. Biol. Evol. 32 (2015) 268–274, https://doi.org/10.1093/molbev/msu300.
[38] S. Raza, H. Shin, H. Hur, T. Unno, Higher abundance of core antimicrobial resistant genes in effluent from wastewater treatment plants, Water Res. 208 (2022) 117882, https://doi.org/10.1016/j.watres.2021.117882.
[39] S. Hernando-Amado, T.M. Coque, F. Baquero, J.L. Martínez, Defining and combating antibiotic resistance from one health and global health perspectives, Nat. Microbiol. 4 (2019) 1432–1442, https://doi.org/10.1038/s41564-019-0503-9.
[40] J. Henson, G. Tischler, Z. Ning, Next-generation sequencing and large genome assemblies, Pharmacogenomics 13 (2012) 901–915, https://doi.org/10.2217/pgs.12.72.
[41] Y. Yang, X. Jiang, B. Chai, L. Ma, B. Li, A. Zhang, J.R. Cole, J.M. Tiedje, T. Zhang, ARGs-OAP: online analysis pipeline for antibiotic resistance genes detection from metagenomic data using an integrated structured ARG-database, Bioinformatics 32 (2016) 2346–2351, https://doi.org/10.1093/bioinformatics/btw136.
[42] Z. Su, L. Chen, D. Wen, Impact of wastewater treatment plant effluent discharge on the antibiotic resistome in downstream aquatic environments: a mini review, Front. Env. Sci. Eng. 18 (2024), https://doi.org/10.1007/s11783-024-1796-3.
[43] Z. Zhang, G. Zhang, F. Ju, Using culture-enriched phenotypic metagenomics for targeted high-throughput monitoring of the clinically important fraction of the β-lactam resistome, Environ. Sci. Technol. 56 (2022) 11429–11439, https://doi.org/10.1021/acs.est.2c03627.
[44] Z. Su, T. Dai, Y. Tang, Y. Tao, B. Huang, Q. Mu, D. Wen, Sediment bacterial community structures and their predicted functions implied the impacts from natural processes and anthropogenic activities in coastal area, Mar. Pollut. Bull. 131 (2018) 481–495, https://doi.org/10.1016/j.marpolbul.2018.04.052.
[45] G. Dantas, M.O.A. Sommer, R.D. Oluwasegun, G.M. Church, Bacteria subsisting on antibiotics, Science 320 (2008) 100–103, https://doi.org/10.1126/science.1155157.
[46] Z. Su, B. Huang, Q. Mu, D. Wen, Evaluating the potential antibiotic resistance status in environment based on the trait of microbial community, Front. Microbiol. 11 (2020), https://doi.org/10.3389/fmicb.2020.575707.
[47] Z. Su, D. Wen, Characterization of antibiotic resistance across Earth's microbial genomes, Sci. Total Environ. 816 (2022) 151613, https://doi.org/10.1016/j.scitotenv.2021.151613.
[48] Y. Zhang, H. Hu, H. Yan, J. Wang, S.K. Lam, Q. Chen, D. Chen, J. He, Salinity as a predominant factor modulating the distribution patterns of antibiotic resistance genes in ocean and river beach soils, Sci. Total Environ. 668 (2019) 193–203, https://doi.org/10.1016/j.scitotenv.2019.02.454.
[49] M.J. Bruins, P. Juffer, M.J.H.M. Wolfhagen, G.J.H.M. Ruijs, Salt tolerance of methicillin-resistant and methicillin-SusceptibleStaphylococcus aureus, J. Clin. Microbiol. 45 (2007) 682–683, https://doi.org/10.1128/JCM.02417-06.
[50] D. Du, X. Wang-Kan, A. Neuberger, H.W. van Veen, K.M. Pos, L.J.V. Piddock, B.F. Luisi, Multidrug efflux pumps: structure, function and regulation, Nat. Rev. Microbiol. (2018), https://doi.org/10.1038/s41579-018-0048-6.
[51] M.W. Van Goethem, R. Pierneef, O.K.I. Bezuidt, Y. Van De Peer, D.A. Cowan, T.P. Makhalanyane, A reservoir of 'historical' antibiotic resistance genes in remote pristine Antarctic soils, Microbiome 6 (2018), https://doi.org/10.1186/s40168-018-0424-5.
[52] J.L. Martínez, T.M. Coque, F. Baquero, What is a resistance gene? Ranking risk in resistomes, Nat. Rev. Microbiol. 13 (2015) 116–123, https://doi.org/10.1038/nrmicro3399.
[53] Z. Su, D. Wen, A.Z. Gu, Y. Zheng, Y. Tang, L. Chen, Industrial effluents boosted antibiotic resistome risk in coastal environments, Environ. Int. 171 (2023) 107714, https://doi.org/10.1016/j.envint.2022.107714.
[54] M. Kvesić, H. Kalinić, M. Dželalija, I. Šamanić, R. Andričević, A. Maravić, Microbiome and antibiotic resistance profiling in submarine effluent-receiving coastal waters in Croatia, Environ. Pollut. 292 (2022) 118282, https://doi.org/10.1016/j.envpol.2021.118282.
[55] P. Lorenzo, A. Adriana, S. Jessica, B. Carles, F. Marinella, L. Marta, B.J. Luis, S. Pierre, Antibiotic resistance in urban and hospital wastewaters and their impact on a receiving freshwater ecosystem, Chemosphere 206 (2018) 70–82, https://doi.org/10.1016/j.chemosphere.2018.04.163.
[56] A. Luczkiewicz, E. Kotlarska, W. Artichowicz, K. Tarasewicz, S. Fudala-Ksiazek, Antimicrobial resistance of Pseudomonas spp. isolated from wastewater and wastewater-impacted marine coastal zone, Environ. Sci. Pollut. Res. 22 (2015) 19823–19834, https://doi.org/10.1007/s11356-015-5098-y.
[57] F. Ju, K. Beck, X. Yin, A. Maccagnan, C.S. McArdell, H.P. Singer, D.R. Johnson, T. Zhang, H. Bürgmann, Wastewater treatment plant resistomes are shaped by bacterial composition, genetic exchange, and upregulated expression in the effluent microbiomes, ISME J. 13 (2019) 346–360, https://doi.org/10.1038/s41396-018-0277-8.
[58] T. Dai, Y. Zhao, D. Ning, B. Huang, Q. Mu, Y. Yang, D. Wen, Dynamics of coastal bacterial community average ribosomal RNA operon copy number reflect its response and sensitivity to ammonium and phosphate, Environ. Pollut. 260 (2020) 113971, https://doi.org/10.1016/j.envpol.2020.113971.
[59] Q. Chen, X. An, B. Zheng, M. Gillings, J. Peñuelas, L. Cui, J. Su, Y. Zhu, Loss of soil microbial diversity exacerbates spread of antibiotic resistance, Soil Ecol. Lett. 1 (2019) 3–13, https://doi.org/10.1007/s42832-019-0011-0.
[60] J. Wang, R. Pan, P. Dong, S. Liu, Q. Chen, A.G.L. Borthwick, L. Sun, N. Xu, J. Ni, Supercarriers of antibiotic resistome in a world's large river, Microbiome 10 (2022), https://doi.org/10.1186/s40168-022-01294-z.
[61] Y. Che, Y. Xia, L. Liu, A. Li, Y. Yang, T. Zhang, Mobile antibiotic resistome in wastewater treatment plants revealed by Nanopore metagenomic sequencing, Microbiome 7 (2019), https://doi.org/10.1186/s40168-019-0663-0.
[62] R. Wu, M.R. Davison, W.C. Nelson, M.L. Smith, M.S. Lipton, J.K. Jansson, R.S. MacClure, J.E. McDermott, K.S. Hofmockel, Hi-C metagenome sequencing reveals soil phage–host interactions, Nat. Commun. 14 (2023) 7666, https://doi.org/10.1038/s41467-023-42967-z.
[63] A.N. Zhang, C. Hou, M. Negi, L. Li, T. Zhang, Online searching platform for the antibiotic resistome in bacterial tree of life and global habitats, FEMS Microbiol. Ecol. 96 (2020), https://doi.org/10.1093/femsec/fiaa107.
[64] L. Li, Y. Xia, T. Zhang, Co-occurrence of antibiotic and metal resistance genes revealed in complete genome collection, ISME J. 11 (2017) 651–662, https://doi.org/10.1038/ismej.2016.155.