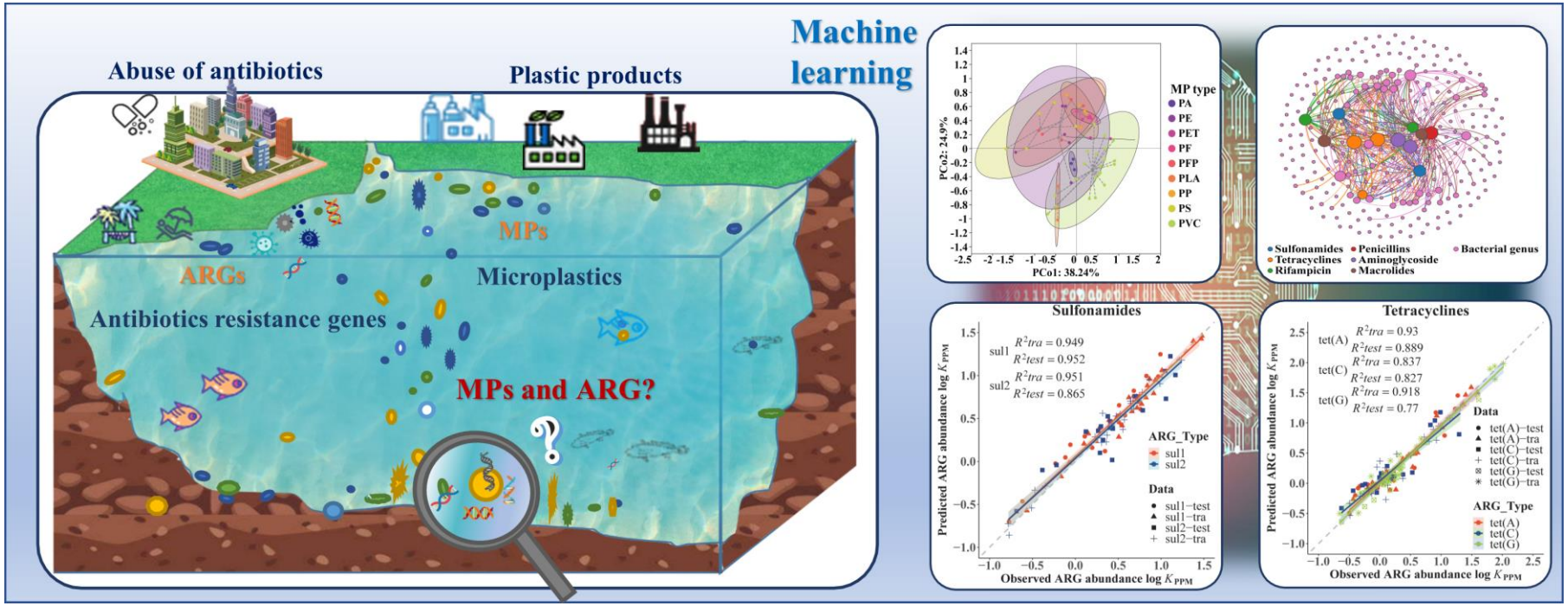
This is the accepted manuscript version of the contribution published as:

Zhu, T., Li, S., Tao, C., **Chen, W.**, Chen, M., Zong, Z., Wang, Y., Li, Y., Yan, B. (2025): Understanding the mechanism of microplastic-associated antibiotic resistance genes in aquatic ecosystems: Insights from metagenomic analyses and machine learning
Water Res. **268, Part A** , art. 122570

The publisher's version is available at:

https://doi.org/10.1016/j.watres.2024.122570

Graphical Abstract



HIGHLIGHTS

(a) MPs in aquatic environments are hotspots for the accumulation of ARGs

(b) The risk of ARGs occurring in degradable microplastics is higher.

(c) Modeling using two independent variables: bacterial taxa and MDs.

(d) Lipophilicity, solubility and surface potential of MPs associations with ARGs

Manuscript Click here to view linked References

Unveiling the Mechanism of Microplastic-Associated Antibiotic Resistance Genes in Aquatic Ecosystems: Insights from Metagenomic Analysis and Machine Learning

Tengyi Zhu [a], Shuyin Li [a], Cuicui Tao [a], Wenxuan Chen [b], Ming Chen [c, d], Zhiyuan Zong [d], Yajun Wang [e], Yi Li [a], Bipeng Yan [a*]

[a] School of Environmental Science and Engineering, Yangzhou University, Yangzhou, 225127, Jiangsu, China

[b] Department of Applied Microbial Ecology, Helmholtz Centre for Environmental Research (UFZ), 04318, Leipzig, Germany

[c] School of Civil Engineering, Southeast University, Nanjing, 210096, China

[d] Department of Engineering Science, University of Oxford, Oxford, OX1 3PJ, UK

[e] School of Civil Engineering, Lanzhou University of Technology, 730050, Lanzhou, China

[*] Corresponding author: Bipeng Yan, Email: bpyan@yzu.edu.cn

Abstract

The pervasive presence of microplastics (MPs) in aquatic systems facilitates the transmission of antibiotic resistance genes (ARGs), posing enduring risks to both ecosystems and human well-being. Nonetheless, research on ARGs associated with plastispheres remains notably scant, particularly regarding whether diverse types of MPs selectively foster ARG enrichment, which remains an enigmatic challenge. This study integrates existing metagenomic datasets to conduct a comprehensive risk assessment and comparative analysis of resistance groups across various MPs. Additionally, we endeavor to elucidate potential associations between ARGs and bacterial taxa as well as MP structural features utilizing machine learning (ML)

methods. The findings underscore the pivotal role of MP type in shaping plastispheres, accounting for 9.56% of biotic variation (Adonis index) and explaining 18.59% of ARG variance. The species evenness and diversity of biodegradable MPs (such as PHA and PLA) are lower compared to conventional MPs, but the risk of ARGs occurrence is higher. These ML approaches effectively forecast ARG abundance utilizing bacterial taxa and molecular structure descriptors (MDs) of MP (average R^2_{tra} = 0.882, R^2_{test} = 0.759). Through feature analysis, it has been observed that MDs associated with lipophilicity, solubility, toxicity, and surface potential can significantly influence the relative abundance of ARGs on plastispheres. The interpretable MLR model, particularly notable, elucidates a linear relationship between bacterial genera and ARGs, offering promise in identifying potential ARG hosts. This investigation offers novel insights into ARGs dynamics and ecological risk within aquatic plastispheres, underscoring the imperative of comprehensive monitoring initiatives for MPs.

Keywords: Microplastics; Antibiotic resistance genes; Bacterial genus; Aquatic environment; Metagenomics; Machine learning.

Abbreviations

MLR, Multiple Linear Regression; MLP, Multilayer Perceptron; GBDT, Gradient Boosting Decision Tree; RF Random Forest; PE, Polyethylene; PS, Polystyrene; PP, Polypropylene; PF, Phenol formaldehyde; PFP, Polyethylene fiber; PVC, Polyvinyl chloride; PLA, Polylactic acid; PHA, Polyhydroxyalkanoates; PET, Polyethylene terephthalate; PA, Polyamide.

1. Introduction

Microplastics (MPs), referring to synthetic polymers with a diameter of less than 5 millimeters, pervade the global ecosystem due to their formidable resistance to

degradation and exceptional mobility (Shahul Hamid et al., 2018). Various organic matter, bacteria, algae, and other substances can adhere to the surface of MPs, gradually forming a microecosystem known as the "plastisphere" (Wang et al., 2021; Rochman et al., 2013; Velzeboer et al., 2014). These plastispheres serve as reservoirs for living organisms and engage in biological exchange with other particles in the environment during long-distance transmission. Intriguingly, plastispheres have been demonstrated to selectively attract antibiotic-resistant bacteria (ARB) and antibiotic resistance genes (ARG), providing channels for the dissemination of antibiotic resistance (AR) (Lin et al., 2021; Shen et al., 2021). The currently burgeoning AR represent one of the most concerning human health issues, having already caused 700,000 deaths annually. It is projected that by 2050, the death toll will reach 10 million, surpassing the number of deaths from cancer each year (Aslam et al., 2018). Authoritative organizations such as the World Health Organization (WHO) and the Centers for Disease Control and Prevention (CDC) have declared AR as a "global public health issue" that requires strict control (Yang et al., 2020).

The ocean was a previously hotspot for the accumulation of MPs, but in recent years, propelled by industrial development and human activities, numerous MP residues have been detected in rivers, lakes, polar regions, and even within organisms (Allen et al., 2019; Wright and Kelly, 2017; Zhu et al., 2022). The widespread presence of MPs in aquatic environments can accelerate the risk of ARGs transmission (Liu et al., 2022). The abundance of ARGs on the plastispheres has been observed to be significantly

higher than that in the environment, as evidenced by researches in the North Pacific circulation, Gulf of Mexico, Huangpu River, and waters near Fujian Island (Lu et al., 2019; Yang et al., 2019; Zhu et al., 2023). Besides, the tendency of ARGs to preferentially adsorb on MPs has also been confirmed in laboratory simulation experiments (Zhang et al., 2022). The prevalence of ARGs in plastisphere surpasses that observed within aqueous domains, accentuating the propensity for plastic substrates to serve as hotspots for ARG accumulation (Basili et al., 2020). Moreover, evidential data underscores the different material and size of MPs exert discernible influences on the abundance of ARGs (Wang et al., 2022). The biological community is the main carrier for the accumulation of ARGs on MPs, however, the complex interaction mechanism between these remains inadequately elucidated.

The intricate relationships between ARGs on plastispheres and their bacterial hosts epitomizes an attractive story. Through redundancy analysis, Zhang et al. (2016) determined that 68.2% of the variations in ARGs can be attributed to the composition of microbial communities in wastewater. Meanwhile, a study conducted on the Huangpu River in China has revealed that MPs have created a unique niche for the genus *Afipia*, enhancing the accumulation of ARGs to rifampicin and vancomycin (Xu et al., 2022). Nevertheless, the network analysis or redundancy analysis methods currently used to study the association between bacterial genera and ARGs may result in spurious correlations or omissions of low-abundance species in small sample sizes, which can increase the difficulty of quantitative analysis (Guo et al., 2017; Rice et al.,

2020). With the rapid progress in gene sequencing technology, accompanied by the decrease in sequencing expenses, and the burgeoning expansion of microbiome exploration, the emergence of high-dimensional and sparse microbiome datasets have further impacted traditional statistical methods (Zhou and Gallins, 2019).

Machine learning (ML) can utilize large datasets for tasks such as recognition, classification, and regression, etc., thereby effectively addressing challenges related to phenotype, feature classification, and intricate associations between biological communities and gene fragments within the domain of microbiology (Li et al., 2019). ML has been used to analyze the interplay between oceanic ARGs and environmental parameters, as well as to predict the relationship between bacterial genomes and minimum inhibitory concentration values of antibiotics in clinical settings (Cuadrat et al., 2020; Dillon et al., 2024). Additionally, ML has been instrumental in monitoring the intricate relationship between ARG abundance in activated sludge and the prevalence of genera, opportunistic pathogens, and nitrifying bacteria (Sun et al., 2021b). ML has demonstrated its efficacy in accurately discerning potential associations between bacterial taxa and ARGs through self-learning mechanisms, cross-validation, and other techniques (Li et al., 2019). Regrettably, owing to the intricate nature of bacterial community structures on plastispheres and the paucity of systematic datasets, there remains a dearth of research utilizing ML methods to investigate the correlation between bacterial community abundance and ARGs on MPs.

To address this gap, this study collected 69 metagenomic datasets (involving 10

types of MPs) from published articles across global aquatic environment, systematically analyzing the composition and distribution of bacteria and ARGs on diverse MPs. In addition, four popular and representative ML algorithms (MLR, MLP, GBDT, RF) are used to explore the complex relationship between the abundance of ARGs and bacterial groups, as well as MPs structural features. We hope that this study will contribute to elucidating the shaping patterns of microbial communities and ARGs on different MPs, as well as tracking potential hosts of ARGs.

2. Material and methods

2.1. Data collection and curation

Starting from May 2023, we embarked on a comprehensive exploration, employing diverse keywords such as plastic, microplastics, MP, water, aqueous, metagenomics, resistance genes, ARG, etc., across renowned databases including Compendex, Biological Science Research, Web of Science, PubMed, and Scopus. This endeavor yielded a rich harvest of 230 peer-reviewed documents. Exclusive inclusion criteria ensured that only articles meeting stringent standards were incorporated: those utilizing metagenomics to scrutinize ARGs on MPs within aqueous environments, and housing metagenomic sequences readily accessible through the National Center for Biotechnology Information (NCBI). Subsequently, 69 metagenomic sequencing files were meticulously selected, including 10 types of MPs. These invaluable datasets constituted the bedrock of our subsequent analyses. Detailed insights regarding the chosen articles and corresponding sequencing file identifiers are documented in Table

S1.

2.2. Bioinformatics analysis

The Aspera tool facilitated the extract of metagenomic sequences from public databases, followed by the trimming of the original sequence readings using Trim_galore (Krueger, 2015). Cutadapt and FastQC were employed to eliminate segments of low-quality reads (Quality Phred score < 20) and remove adapter sequences (Martin, 2011; Utturkar et al., 2020). In the final dataset, the average number of sequences across all plastic samples was 51,689,494, with a range from 1,268,586 to 139,878,101 sequences. The refseq database (2023-05-23) was designated as the reference database (https://kaiju.binf.ku.dk/server), and Kaiju software was employed to annotate the relative abundance of species in metagenomic data, with biological classification refined to the genus level (Menzel et al., 2016). The calculation of ARGs relied on reference data from the Comprehensive Antibiotic Resistance Database (CARD), annotated with RGI-bwt on the trimmed data, requiring a sequence identity of 80% (Alcock et al., 2019; Yang et al., 2019). ARG abundance was measured in PPM (i.e., one ARG-like read per one million metallurgical sequential reads) (Yang et al., 2013). The relative abundance of taxonomic genera in the sample was determined by the ratio of Kaiju-matched readings to the total number of sequencing file readings (Baral et al., 2018). Before modeling, logarithmic transformations were applied to both genera and ARG abundance data.

2.3. Variable preparation

This study investigated two distinct categories of independent variables: the relative abundance of bacterial genera and molecular structure descriptors (MDs) derived from plastic monomers. Through the comparison of metagenomic data with the refseq database, a comprehensive profile of 1935 bacterial genera (excluding unclassified and those unassignable to a non-viral genus) was established. Additionally, employing alvaDesc software (version 1.1.2), 5290 MDs capturing various structural attributes of MPs were generated, followed by the removal of constant values, missing data, and descriptors exhibiting collinearity exceeding 95% (Rácz et al., 2019). Utilizing the CARD, a total of 531 ARGs were identified from the sequenced files. To ensure the representation of modeling data within the samples, screening was conducted for ARGs with an occurrence rate surpassing 50% across all samples. Ultimately, 12 ARGs (*sul1*, *sul2*; *tet (A)*, *tet (C)*, *tet (G)*; *APH(6)-ID*, *APH(3")-Ib*; *EreA*, *ErmF*; *Bado_rpoB_RIF*, *rpoB2*; *VanRO*) corresponding to six prevalent and abundant antibiotic families (namely: sulfonamides; tetracyclines; aminoglycosides; macrolides; rifampicin; penicillin) were chosen for modeling purposes.

In SPSS (version 2019.2.1), MLR analysis was conducted on the independent variables. The optimal variable set was determined based on the ideal combination of R^2_{tra} and Q^2_{tra} with the minimum number of variables, while ensuring $VIF < 5$ (Ghavami and Faham, 2010). Pearson correlation coefficient analysis was employed to ensure the statistical independence of variables, with the specific variables selected for analysis detailed in Table S5 (Rovetta, 2020).

2.4. ML models' development and statistical analysis

Drawing upon metagenomic data sourced from MPs, this study utilized bacterial genera and MDs as pivotal variables, while selecting ARGs as response variables, thus facilitating the development of predictive models. Four widely acclaimed ML algorithms (MLR, MLP, GBDT and RF) were meticulously implemented within the Python software (version 3.9). The dataset underwent random partitioning into a 60% training set and a 40% test set (Sun et al., 2021b). Rigorous methodologies were applied: the training set underwent grid search coupled with 10-fold cross-validation for model selection and hyperparameter optimization, while the test set was designated for impartial evaluation of model performance (Malakouti et al., 2023). This approach yielded a total of 48 distinct models, with their fitting precision assessed using the adjusted coefficient of determination, R^2_{tra}. Model stability was scrutinized via Q^2_{loo}, while predictive accuracy was quantified through the root mean square error ($RMSE_{tra}$). To bolster confidence in the model's extrapolative capabilities, evaluation metrics such as R^2_{test} and Q^2_{test} were employed, with detailed calculation methods elucidated in Text S5.

The distribution and abundance of ARGs on MPs were vividly depicted utilizing the Chiplot software platform (https://www.chiplot.online). Variable correlation patterns were visually elucidated through the ggpairs package in R (version 4.2.2), while the fidelity between observed and predicted values was effectively visualized using the ggplot software package.

3. Results and discussion

3.1. Bacterial community structure on MPs

The comprehensive metagenomic data assembled in this investigation encompasses a spectrum of 10 MP types, with PVC representing the most prevalent type, comprising 37.68% of the total, closely followed by PE and PHA, each accounting for 8.7%. The frequency ranking of other MP types is shown in Fig. 1A. MPs manifest in aquatic environments in various forms including fibers, films, particles, fragments, and foam (Lenaker et al., 2019), with fibers being the predominant form in this study, constituting 44.93% of the samples.

Across all samples, Pseudomonadota emerges as the dominant bacterial phylum, averaging 56.34%, trailed by Actinomycetota at an average proportion of 11.28%, securing the second position, and Bacteroidota claiming the third spot at 9.45% (Fig. 1B). Prior investigations have reported the frequently presence of Pseudomonas and Actinobacteria on MPs, aligning with the findings of this study (Lv et al., 2024; Sun et al., 2024). Specific details regarding bacteria's relative abundance are elucidated in Table S2. To gauge microbial community richness and diversity across diverse MP types, species α-diversity indices were computed at the genus level (Finn, 2024). PF plastics exhibit the highest abundance of bacterial genera, tallying 1932, while PLA plastics host the fewest genera at a mere 927. The diversity of bacterial genera is visually depicted in Fig. 1C, with PF plastics boasting an average Shannon index of 6.25 and Simpson index of 0.99, indicative of the highest diversity. Following closely

are PFP plastics, exhibiting Shannon index of 5.20 and Simpson index of 0.94, whereas PHA plastics display the lowest diversity, with a Shannon index value of 2.46 and Simpson index of 0.44. In terms of species evenness (assessed via Pielou's index analysis), microbial communities on PF plastics manifest the greatest dominance compared to others (Pielou: 0.33-0.69), boasting an average Pielou value of 0.83, emblematic of the highest diversity (Wang et al., 2019). The relative abundance analysis of bacterial genera reveals that *Pseudomonas* constitutes a substantial portion in PHA MPs, reaching 59.77%. In addition, the collective proportion of three predominant bacterial genera in PLA, namely *Pseudomonas*, *Brevundimonas*, and *Cloacibacterium*, stands at 39.09%. In contrast, other MPs exhibit a more evenly distributed bacterial genus composition. This observation underscores the rationale for the diminished species diversity and evenness observed in this two MPs. The PHA and PLA are widely used plastics with commendable biodegradability and compatibility (Burzic et al., 2019). The PHA plastics are synthesized via microbial fermentation, with *Pseudomonas* emerging as the principal bacterial genus responsible for generating PHA polymers of medium to long chain lengths (Liu et al., 2021). Previous investigations have elucidated that *Pseudomonas* and *Brevundimonas* possess the capability to cleave lipid bonds within PLA through enzymatic degradation, yielding oligomers and lactic acid monomers. The degraded oligomers can penetrate a semi-permeable membrane as a carbon source (Bubpachat et al., 2018). Additionally, the lactic acid byproduct generated during degradation exerts inhibitory effects on the proliferation of various

bacterial genera, but acid-tolerant bacteria, such as *Cloacibacterium*, may exploit this condition to establish dominance (Panyachanakul et al., 2019).

According to the analysis of biota on these MP samples, it can be observed that there is a close correlation between the biological community and the types of MPs. The Principal Coordinates Analysis (PCoA) findings accentuate pronounced disparities in community composition among distinct MP types (R^2 = 0.49, p = 0.001), as delineated in Fig. 1D, with PS samples diverging from PFP samples (Xu et al., 2022). Adonis analysis further elucidates this divergence, indicating a noteworthy contribution of MP types within the dataset, accounting for 9.56% of the variance in plastisphere biological community structure (p = 0.001). The notable intra-group variance in microbial community composition on plastisphere may stem from the ambiguous origins and transport processes of MPs in aquatic ecosystems, shrouding their historical trajectories (Wang et al., 2023).

<Fig. 1>

3.2. The distribution of ARGs on MPs

Upon precisely scrutiny with reference databases, it is revealed that 63 MP samples within the dataset contain ARGs, with a tally of 532 ARGs identified, exhibiting a relative abundance ranging from 0.021 to 474.93 PPM (Fig. 2). The distribution of ARGs across different MPs evinces significant disparities, with PHA showcasing the highest ARG abundance, summing up to 2497.85 PPM. The relative abundance (sum PPM) of ARGs across other MPs is ranked as follows: PLA (1623.47) >

PE (937.63) > PS (681.31) > PFP (520.52) > PP (506.02) > PF (257.32) > PA (220.59) > PET (81.89), as illustrated in Fig. 3A.

<Fig. 2>

Predominating among all detected ARG types are multi-drug resistance genes, which exhibit the broadest distribution across the sampled specimens, surpassing an occurrence rate of 85% (Zhang et al., 2020). Subsequently, genes encoding resistance against sulfonamides, tetracyclines, and aminoglycosides emerge prominently. Moreover, resistance genes against macrolides, penicillin, and rifampicin are also detected widely across the samples, underpinning the rationale for selecting a specific subset of antibiotics as response variables in this study (González-Pleiter et al., 2021; Zadjelovic et al., 2023). The abundance distribution of 12 representative ARGs is listed in Table S3. Further elaboration on the distribution of these ARGs across different MP types is displayed in Fig. 3B. Significant variations exist in the distribution patterns of ARGs across distinct categories of MPs. ARGs like *tet(A)*, *tet(C)*, *tet(G)*, *sul1*, *sul2*, *APH(6)-Id*, and *APH(3")-Ib* are primarily located on PLA, whereas *Bado_rpoB_RIF* exhibits the highest prevalence on PVC. The *rpoB2* demonstrates minimal presence on PHA and PLA, contrasting with its widespread occurrence on other non-biodegradable plastics. This distribution difference may be related to the survival pressure of ARGs host bacteria on MPs. In the PCoA analysis (Fig. 3C), a notable discrepancy of ARG abundance was evident across 10 MP types along the principal coordinates (PCo1: 38.24%). Adonis analysis revealed that the type of MPs explained 18.59% (R^2) of the

relative abundance differences in ARGs among different samples (p = 0.001). Much like the bacterial genera, substantial discrepancies were also observed in the distribution of ARGs across different MPs. These differences may be due to the distinct surface ecological niches of various MP types, which can create selective pressure on the expression of specific ARGs (Li et al., 2024). In summary, compared to non-degradable plastics, PLA and PHA have the highest risk of ARGs occurrence. This heightened risk can be attributed to the favorable biocompatibility of PLA and PHA, which fosters microbial activity on their surfaces, consequently facilitating the release and dissemination of ARGs (Song et al., 2022).

Beyond considerations of external pressures and raw data origins, these variations in ARG distribution are intricately linked to the traits of biological communities. Fig. 3D visually portrays the intricate symbiotic interplay between biological communities and ARGs. A primary challenge in unraveling resistance groups within MPs lies in pinpointing the hosts of ARGs (Piergiacomo et al., 2022), a task riddled with complexity and subtlety. To address this challenge, we employ ML methods adept at processing high-dimensional, low-sparsity data, thereby accurately uncovering latent trends and intricate dynamics within biological datasets.

<Fig. 3>

3.3. Development and evaluation of ML models

The MLR method was employed to conduct dimensionality reduction and screening for 12 groups of bacterial genera abundances corresponding to ARGs, along

with MDs. A total of 66 response variables were obtained, comprising 62 bacterial genera and 4 MDs. The correlation coefficients of variables were all less than 0.8, as depicted in the correlation analysis shown in Figure S4. Using four modeling methods (MLR, MLP, GBDT, and RF), a total of 48 ML models were developed, and the hyperparameters used for modeling and validation are listed in Table S7. The predictive performance of the developed models is illustrated in Fig. 4, showing that all models could explain over 80% of the variability in ARG abundance training data set. All modeling methods effectively accounted for the variability in the ARG training set, with average R^2_{tra} values of 0.912 for MLR, 0.875 for MLP, 0.871 for GBDT, and 0.887 for RF, respectively. These ML algorithms exhibited strong predictive accuracy and robustness, with average $RMSE_{tra}$ ranging from 0.231 to 0.288 and average Q^2_{loo} from 0.633 to 0.817. When forecasting ARG abundance within unfamiliar datasets, the ML models also exhibited commendable performance, elucidating more than 65% of the variance in abundance. The MLR model achieved an average R^2_{test} of 0.881, while the MLP model attained 0.837. As for GBDT and RF, their respective R^2_{test} values stood at 0.656 and 0.664. For the majority of ARGs datasets, all four ML algorithms adeptly fitted and forecasted the association between ARGs abundance and variables (Table S8-S11). However, when employing the GBDT and RF modeling methods on the *sul2*, *EreA*, *ErmF*, and *vanRO* datasets, the models' performance in terms of robustness and predictability fell short of ideal, with R^2_{test} ranging from 0.283 to 0.686. The predictive efficacy of RF and GBDT declined when applied to external datasets, a phenomenon

that could be attributed to a mismatch between the model's operational logic and data features, or a lack of generalization ability (Lang et al., 2021). Additionally, factors such as the specificity of variables in the original dataset (such as *APH(3'')-Ib*, *EreA*, *ErmF*, *rpoB2*, *sul2*, these five datasets contain MD variables) or the small size of the dataset itself should also be considered (Haffiez et al., 2022).

<Fig. 4>

Overall, all four modeling approaches in this investigation achieved respectable performance metrics; however, the MLR model emerged as the standout performer across all evaluated aspects. The establishment of a linear relationship between independent variables (bacterial genera and MDs) and dependent variables (ARGs) represents a promising outcome. The interpretative capability of MLR coefficients facilitates a clear understanding of each variable's influence on ARGs abundance, circumventing the complexities associated with black-box models (Panov and Varaksin, 2010).

3.4. Feature analysis of the ML models

The 62 identified bacterial genus variables were primarily categorized into 16 bacterial phyla, with Pseudomonas emerging as the predominant phylum (38.7%), followed by Actinomycetota (16.13%) and Bacteroidota (16.13%), consistent with the distribution of abundant bacterial phyla observed in the comprehensive sample analysis presented in section 3.1. Previous studies on resistance groups in fermentation beds of aquaculture farms have indicated that these three bacterial phyla are putative hosts of

ARG, and metagenomic box tracing has confirmed that the main multidrug-resistant phenotypes in Actinomycetes are rifampicin and vancomycin (Pan et al., 2024). Hernández et al. (2023) elucidated, via isotope labeling, the dominance of Pseudomonas as a bacterium in environments subjected to antibiotic exposure. The presence of *APH(3'')* was discerned within heavy fractions, signifying robust resistance to aminoglycoside antibiotics. Previous studies on soil resistance groups have revealed a correlation between the prevalence of sulfonamide and tetracycline ARGs and the abundance of Bacteroidota and Proteobacteria in environmental matrices. In addition, Bacteroidota has been identified as a potential host for diverse sulfonamide and tetracycline ARGs (Razavi et al., 2017; Saha et al., 2023). Bacteria exhibiting resistance to multiple antibiotics can swiftly occupy ecological niches and gain survival advantages when external stresses diminish, thereby facilitating the heightened accumulation of ARGs (Hayward et al., 2010; Wang et al., 2024).

The predictive performance of the 12 MLR models is shown in Fig. 5 and Figure S5. Tables S5 and S6 present the selected independent variables along with the coefficients of the linear regression equation. Taking the *rpoB2* dataset as a case study, which encompasses two categories of independent variables, the linear regression equation is as follows: PPM_{rpoB2} = 0.159 *Ornithininimicrobium* - 0.074 *Pontibacter* - 0.071 *P_VSA_s_6* - *0.028 Pantoea* - 0.032 *Rhodanobacter* + 0.022 *Marinobacter* + 1.396. According to the regression equation, the selected bacterial genera and MDs can explain 95.2% of the changes in the *rpoB2* dataset, and the prediction accuracy for test

datasets reaches 96.8%. *Ornithininimicrobium* and *Marinobacter* are frequently found in oceans and salt lakes, with high salt tolerance and diverse metabolic abilities, playing important roles in ecosystems (Bonin et al., 2015; Manoharan et al., 2022). The elevation of water salinity may precipitate a reduction in biological activity within plastispheres and a decline in the abundance of ARGs, while bacteria endowed with high salt tolerance may exhibit formidable resistance to such antagonistic influences. These bacterial genera may be stimulated by antibiotics to secrete metabolites that promote the expression of ARGs or have certain symbiotic relationships with specific resistant bacteria. *Pontibacter*, *Pantoea*, and *Rhodanobacter*, all Gram-negative bacteria, are common bacteria in soil and aquatic environments. *Pontibacter* exhibits excellent substrate degradation capabilities, and some strains even possess the ability to degrade environmental DNA to maintain survival (Zhou et al., 2016). Previous research has identified *Pantoea* as a potential host for ARGs, particularly associated with the rifampicin resistance gene, and most *Pantoea* strains display certain pathogenicity (Agrios, 2005; Baral et al., 2018). *Rhodanobacter* plays a crucial role in sulfur metabolism and is involved in the sulfur cycle (Aburto-Medina et al., 2021). This negative correlation may indicate that these bacterial genera may inhibit the expression of *rpoB2* through resource competition or the production of inhibitory factors, leading to a decrease in gene expression levels. The *P_VSA_s_6* is a MD representing the van der Waals surface area (VSA), which is indirectly related to the molecular size and lipophilicity (Zhang et al., 2024). Lipophilic substances are more likely to penetrate the

cell membrane, making bacteria more sensitive to external inhibitors. Previous studies on the toxicity of organic chemicals have indicated a positive correlation between *P_VSA_s_6* and toxicity, which may affect the regulatory mechanisms of resistance genes (Sun et al., 2021a). These findings help to further explore the potential expression relationship of ARG in different environments and reveal the complex interaction between microbial community structure and resistance gene transmission.

<Fig. 5>

Furthermore, the MDs encompass additional parameters: 0.071 *J_Dz(v)*, -0.071 *ChiA_B(s)*, 0.259 *SpMax_B(s)*, and -0.051 *X0Av*. *J_Dz(v)* is a Balaban like index related to van der Waals volume, which can reflect the size and shape information of molecules (Chavan et al., 2014). *ChiA_B(s)* and *SpMax_B(s)* are indices related to the polarity, solubility, and surface activity of molecules (Kubik and Wiczling, 2016). *X0Av* can reflect the affinity of electrons and is related to chemical reactivity (Kowalewski and Ray, 2020). Previous studies have demonstrated that PS MPs increase the contact potential between donor and recipient cells, thereby enhancing the frequency of ARG transfer (Cheng et al., 2022). Moreover, PS particles can dissolve into the hydrophobic core of lipid bilayers, thereby changing the membrane structure and dynamic characteristics of bacteria, promoting the permeability of receptor membranes, and increasing the transfer frequency of ARGs (Hu et al., 2022). Additionally, organic compounds released during the aging of MPs such as PVC and PA can further fuel the enrichment of ARGs (Yuan et al., 2022). Collectively, these MDs furnish insights into

the molecular topology and electronic interactions of MPs, which can affect the physicochemical properties and surface features of MPs, thereby differentially enriching ARGs by modulating the dynamic attributes of the biological community.

3.5. Limitations and perspective

In summary, plastic material significantly influence the distribution of ARGs, with the bacterial genera and structural characteristics of MPs serving as crucial features for predicting ARG abundance. ML models effectively uncover response relationships among variables. However, it's important to acknowledge that our study is meta-analytical in nature, and conclusions may be influenced in part by sample strategies, environmental factors, and other conditions. Given the limited availability of metagenomic research on MPs, the modest number of datasets utilized in this study may introduce some regression bias when developing models (Zhang and Lu, 2012). Furthermore, mobile genetic elements (including horizontal transfer of ARGs) and pollutants adsorbed onto MPs were not considered. ML has found diverse applications in biology, including investigating the correlation between resistance genes in urban economies and environments, shaping bacterial community resistance profiles based on environmental factors and antibiotic concentrations, and predicting the impact of water treatment processes on ARG removal (Aburto-Medina et al., 2021; Haffiez et al., 2022; Sun et al., 2018). As next-generation sequencing technologies continue to advance, the volume of microbial information is expanding rapidly, and ML will play an increasingly vital role in processing future biological data. While our work underscores the superior

reliability and accuracy of ML in biological data processing, further experimental endeavors are warranted to develop more robust ML algorithms, for precise recognition of classification and resistance groups with unique associations to MP features.

4、Conclusions

In this study, we conducted an in-depth analysis of variances in bacterial community structure and the occurrence of ARGs across diverse MP types through metagenomics. Additionally, this study explored, for the first time, the relationship between ARG abundance and the bacterial genera as well as MDs of MPs using ML methods, thereby offering a novel perspective on plastispheres in aquatic environments. The research results indicated that a notable presence of abundant ARGs in the metagenomic data of 10 MP types, underscoring MPs as pivotal hotspots for the accumulation and dissemination of ARGs. The occurrence rate of multidrug-resistant genes in all samples exceeded 85%, which has been identified as the main characteristic of ARGs in MPs. Biodegradable microplastics, such as PLA and PHA, demonstrate reduced species diversity and evenness, along with elevated relative abundance of ARGs, in contrast to conventional MPs. This may be attributed to the easier degradation of PLA and PHA, as their generated water-soluble oligomers can serve as carbon sources, promoting microbial growth and reproduction, thereby stimulating the production or transfer of ARGs. Through ML techniques, we identified that molecular features pertaining to lipophilicity, solubility, toxicity, and surface potential of MP significantly influence ARGs on plastispheres. The four modeling methods of MLP,

GBDT, RF, and MLR all effectively explained the intricate relationship between the abundance of ARGs and bacterial taxonomic genera and MDs, with an average R^2_{tra} of 0.882 and R^2_{test} of 0.759. Particularly noteworthy is the interpretable MLR model, which allows for a clear understanding of positive or negative dependencies between variables and responses. These functional relationships can aid in speculating on potential bacterial hosts or inhibitors of ARGs. This study provides a useful starting point for formulating strategies to accurately characterize the mechanisms of microplastic transmission resistance groups and assess related ecological risks.

Microplastics, colonized by diverse and metabolically complex bacterial communities, have emerged as sanctuaries for pathogens, ARGs, and various independent factors, posing a significant threat as a novel pollutant (Pham et al., 2021). Given the widespread distribution of MPs in global aquatic environments, the ecological niches formed on their surfaces may exacerbate the risks posed by ARGs to both ecology and human health (Ren et al., 2023). Hence, it is imperative to undertake further investigations into source identification, pollution control, and ecological risk assessment management of MPs and antibiotics in water bodies. In the future, the integration of metagenomics with ML research will be crucial for uncovering the prevalence and expression of broad-spectrum ARGs, as well as other molecular processes associated with biofilm formation. We trust that this study can offer valuable insights for the classification and risk management of resistance groups in plastispheres.

Acknowledgements

The present work was funded by National Natural Science Foundation of China

(Grant No. 42077331 and 52100027). It was also sponsored by Qing Lan Project of Jiangsu Province, and by the High-end Talent Support Program of Yangzhou University, China.

Appendix A. Supplementary Material

Supplementary information: 5 Texts, 11 Tables, 5 Figures.

Reference

Aburto-Medina, A., Shahsavari, E., Taha, M., Bates, B., R., Van Ieperen, L. and Ball, A.S., 2021. The impacts of different biological treatments on the transformation of explosives waste contaminated sludge. Molecules 26 (16), 4814.

Agrios, G.N., 2005. Plant Pathology (Fifth Edition). Academic Press, San Diego.

Alcock, B.P., Raphenya, A.R., Lau, T.T.Y., Tsang, K.K., Bouchard, M., Edalatmand, A., Huynh, W., Nguyen, A.-L.V., Cheng, A.A., Liu, S., Min, S.Y., Miroshnichenko, A.S., Tran, H.-K., Werfalli, R.E., Nasir, J.A., Oloni, M., Speicher, D.J., Florescu, A., Singh, B., Faltyn, M., Hernández-Koutoucheva, A., Sharma, A.N., Bordeleau, E., Pawlowski, A.C., Zubyk, H.L., Dooley, D.M., Griffiths, E.J., Maguire, F., Winsor, G.L., Beiko, R.G., Brinkman, F.S.L., Hsiao, W.W.L., Domselaar, G.H.V. and McArthur, A.G., 2019. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res. 48, 517-525.

Allen, S., Allen, D., Phoenix, V.R., Le Roux, G., Durántez Jiménez, P., Simonneau, A., Binet, S. and Galop, D., 2019. Atmospheric transport and deposition of microplastics in a remote mountain catchment. Nat. Geosci. 12 (5), 339-344.

Aslam, B., Wang, W., Arshad, M.I., Khurshid, M., Muzammil, S., Rasool, M.H., Nisar, M.A., Alvi, R.F., Aslam, M.A., Qamar, M.U., Salamat, M.K.F. and Baloch, Z., 2018. Antibiotic resistance: a rundown of a global crisis. Infect. Drug Resist. 11, 1645-1658.

Baral, D., Dvorak, B.I., Admiraal, D., Jia, S., Zhang, C. and Li, X., 2018. Tracking the sources of antibiotic resistance genes in an urban stream during wet weather using shotgun metagenomic analyses. Environ. Sci. Technol. 52 (16), 9033-9044.

Basili, M., Quero, G., Giovannelli, D., Manini, E., Vignaroli, C., Avio, C.G., De Marco, R. and Luna, G., 2020. Major role of surrounding environment in shaping biofilm community composition on marine plastic debris. Front. Mar. Sci. 7, 262.

Bonin, P., Vieira, C., Grimaud, R., Militon, C., Cuny, P., Lima, O., Guasco, S., Brussaard, C.P.D. and Michotey, V., 2015. Substrates specialization in lipid compounds and hydrocarbons of Marinobacter genus. Environ. Sci. Pollut. Res. 22 (20), 15347-15359.

Bubpachat, T., Sombatsompop, N. and Prapagdee, B., 2018. Isolation and role of polylactic acid-degrading bacteria on degrading enzymes productions and PLA biodegradability at mesophilic conditions. Polym. Degrad. Stab. 152, 75-85.

Burzic, I., Pretschuh, C., Kaineder, D., Eder, G., Smilek, J., Másilko, J. and Kateryna, W., 2019. Impact modification of PLA using biobased biodegradable PHA biopolymers. Eur. Polym. J. 114, 32-38.

Chavan, S., Nicholls, I.A., Karlsson, B.C.G., Rosengren, A.M., Ballabio, D., Consonni, V. and Todeschini, R., 2014. Towards global QSAR model building for acute toxicity: Munro database case study. Int. J. Mol. Sci. 15 (10), 18162-18174.

Cheng, Y., Lu, J., Fu, S., Wang, S., Senehi, N. and Yuan, Q., 2022. Enhanced propagation of intracellular and extracellular antibiotic resistance genes in municipal wastewater by microplastics. Environ. Pollut. 292, 118284.

Cuadrat, R.R.C., Sorokina, M., Andrade, B.G., Goris, T. and Dávila, A.M.R., 2020. Global ocean resistome revealed: Exploring antibiotic resistance gene abundance and distribution in TARA Oceans samples. GigaScience 9 (5), 46.

Dillon, L., Dimonaco, N. and Creevey, C.J., 2024. Accessory genes define species-specific routes to antibiotic resistance. Life Sci. Alliance 7 (4), 202302420.

Finn, D.R., 2024. A metagenomic alpha-diversity index for microbial functional biodiversity. FEMS Microbiol. Ecol. 100 (3), 19.

Ghavami, R. and Faham, S., 2010. QSRR models for Kováts' retention indices of a variety of volatile organic compounds on polar and apolar GC stationary phases using molecular connectivity indexes. Chromatographia 72 (9), 893-903.

González-Pleiter, M., Pedrouzo-Rodríguez, A., Verdú, I., Leganés, F., Marco, E., Rosal, R. and Fernández-Piñas, F., 2021. Microplastics as vectors of the antibiotics azithromycin and clarithromycin: Effects towards freshwater microalgae. Chemosphere 268, 128824.

Guo, J., Li, J., Chen, H., Bond, P.L. and Yuan, Z., 2017. Metagenomic analysis reveals wastewater treatment plants as hotspots of antibiotic resistance genes and mobile genetic elements. Water Res. 123, 468-478.

Haffiez, N., Chung, T.H., Zakaria, B.S., Shahidi, M., Mezbahuddin, S., Maal-Bared, R. and Dhar, B.R., 2022. Exploration of machine learning algorithms for predicting the changes in abundance of antibiotic resistance genes in anaerobic digestion. Sci. Total Environ. 839, 156211.

Hayward, A.C., Fegan, N., Fegan, M. and Stirling, G.R., 2010. Stenotrophomonas and Lysobacter: ubiquitous plant-associated gamma-proteobacteria of developing significance in applied microbiology. J. Appl. Microbiol. 108 (3), 756-770.

Hernández, M., Roy, S., Keevil, C.W. and Dumont, M.G., 2023. Identification of diverse antibiotic resistant bacteria in agricultural soil with $H_2{}^{18}O$ stable isotope probing combined with high-throughput sequencing. Environ. Microbiome 18 (1), 34.

Hu, X., Waigi, M.G., Yang, B. and Gao, Y., 2022. Impact of plastic particles on the horizontal transfer of antibiotic resistance genes to bacterium: Dependent on particle sizes and antibiotic resistance gene vector replication capacities. Environ. Sci. Technol. 56 (21), 14948-14959.

Kowalewski, J. and Ray, A., 2020. Predicting novel drugs for SARS-CoV-2 using

machine learning from a >10 million chemical space. Heliyon 6 (8), 04639.

Krueger, F., 2015. Trim Galore!: A wrapper around Cutadapt and FastQC to consistently apply adapter and quality trimming to FastQ files, with extra functionality for RRBS data. URL https://github.com/FelixKrueger/TrimGalore. (Date of access: 22/02/2023).

Kubik, Ł. and Wiczling, P., 2016. Quantitative structure-(chromatographic) retention relationship models for dissociating compounds. J. Pharm. Biomed. Anal. 127, 176-183.

Lang, L., Tiancai, L., Shan, A. and Xiangyan, T., 2021. An improved random forest algorithm and its application to wind pressure prediction. Int. J. Intell. Syst. 36 (8), 4016-4032.

Lenaker, P.L., Baldwin, A.K., Corsi, S.R., Mason, S.A., Reneau, P.C. and Scott, J.W., 2019. Vertical distribution of microplastics in the water column and surficial sediment from the Milwaukee River Basin to Lake Michigan. Environ. Sci. Technol. 53 (21), 12227-12237.

Li, K., Xu, L., Bai, X., Zhang, G., Zhang, M. and Huang, Y., 2024. Potential environmental risks of field bio/non-degradable microplastic from mulching residues in farmland: Evidence from metagenomic analysis of plastisphere. J. Hazard. Mater. 465, 133428.

Li, Y., Huang, C., Ding, L., Li, Z., Pan, Y. and Gao, X., 2019. Deep learning in bioinformatics: Introduction, application, and perspective in the big data era. Methods 166, 4-21.

Lin, Z., Yuan, T., Zhou, L., Cheng, S., Qu, X., Lu, P. and Feng, Q., 2021. Impact factors of the accumulation, migration and spread of antibiotic resistance in the environment. Environ. Geochem. Health 43 (5), 1741-1758.

Liu, C.-H., Chen, H.-Y., Chen, Y.-L.L. and Sheu, D.-S., 2021. The polyhydroxyalkanoate (PHA) synthase 1 of Pseudomonas sp. H9 synthesized a 3-hydroxybutyrate-dominant hybrid of short- and medium-chain-length PHA. Enzyme Microb. Technol. 143, 109719.

Liu, X., Wang, H., Li, L., Deng, C., Chen, Y., Ding, H. and Yu, Z., 2022. Do microplastic biofilms promote the evolution and co-selection of antibiotic and metal resistance genes and their associations with bacterial communities under antibiotic and metal pressures? J. Hazard. Mater. 424 (Pt A), 127285.

Lu, J., Zhang, Y., Wu, J. and Luo, Y., 2019. Effects of microplastics on distribution of antibiotic resistance genes in recirculating aquaculture system. Ecotoxicol. Environ. Saf. 184, 109631.

Lv, S., Li, Y., Zhao, S. and Shao, Z., 2024. Biodegradation of typical plastics: From microbial diversity to metabolic mechanisms. Int. J. Mol. Sci. 25, 593.

Malakouti, S.M., Menhaj, M.B. and Suratgar, A.A., 2023. The usage of 10-fold cross-validation and grid search to enhance ML methods performance in solar farm power generation prediction. Clean. Eng. Technol. 15, 100664.

Manoharan, H., Kvl, A., Mariappan, S., Sekar, U. and Venkataramana, G., 2022.

Molecular characterization of high-level aminoglycoside resistance among Enterococcus species. J. Lab. Physicians 14 (3), 290-294.

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. 17, 10-12.

Menzel, P., Ng, K.L. and Krogh, A., 2016. Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nat. Commun. 7, 11257.

Pan, Z., Wang, W., Chen, J., Chen, Z., Avellán-Llaguno, R.D., Xu, W., Duan, Y., Liu, B. and Huang, Q., 2024. Temporal dynamics of microbial composition and antibiotic resistome in fermentation bed culture pig farms across various ages. Sci. Total Environ. 912, 168728.

Panov, V.G. and Varaksin, A.N., 2010. Relation between the coefficients of simple and multiple regression models. Siberian Math. J+. 51 (1), 162-167.

Panyachanakul, T., Sorachart, B., Lumyong, S., Lorliam, W., Kitpreechavanich, V. and Krajangsang, S., 2019. Development of biodegradation process for Poly(DL-lactic acid) degradation by crude enzyme produced by Actinomadura keratinilytica strain T16-1. Electron. J. Biotechnol. 40, 52-57.

Pham, D.N., Clark, L. and Li, M., 2021. Microplastics as hubs enriching antibiotic-resistant bacteria and pathogens in municipal activated sludge. J. Hazard. Mater. Lett. 2 (21), 100014.

Piergiacomo, F., Brusetti, L. and Pagani, L., 2022. Understanding the interplay between antimicrobial resistance, microplastics and xenobiotic contaminants: A leap towards one health? Int. J. Environ. Res. Public Health 20 (1), 42.

Rácz, A., Bajusz, D. and Héberger, K., 2019. Intercorrelation limits in molecular descriptor preselection for QSAR/QSPR. Mol. Inf. 38 (8-9), 1800154.

Razavi, M., Marathe, N.P., Gillings, M.R., Flach, C.-F., Kristiansson, E. and Joakim Larsson, D.G., 2017. Discovery of the fourth mobile sulfonamide resistance gene. Microbiome 5 (1), 160.

Ren, J., Shi, H., Liu, J., Zheng, C., Lu, G., Hao, S., Jin, Y. and He, C., 2023. Occurrence, source apportionment and ecological risk assessment of thirty antibiotics in farmland system. J. Environ. Manage. 335, 117546.

Rice, E.W., Wang, P., Smith, A.L. and Stadler, L.B., 2020. Determining hosts of antibiotic resistance genes: A review of methodological advances. Environ. Sci. Technol. Lett. 7 (5), 282-291.

Rochman, C.M., Manzano, C.A., Hentschel, B.T., Simonich, S.L.M. and Hoh, E., 2013. Polystyrene plastic: a source and sink for polycyclic aromatic hydrocarbons in the marine environment. Environ. Sci. Technol. 47 (24), 13976-13984.

Rovetta, A., 2020. Raiders of the lost correlation: A guide on using Pearson and Spearman coefficients to detect hidden correlations in medical sciences. Cureus. 12(11), 11794.

Saha, S., Xiong, J.-Q., Patil, S., Ha, G.-S., Hoh, J.-K., Park, H.-K., Chung, W., Chang, S.-W., Khan, M., Park, H. and Jeon, B.-H., 2023. Dissemination of sulfonamide resistance genes in digester microbiome during anaerobic digestion of food waste

leachate. J. Hazard. Mater. 452, 131200.

Shahul Hamid, F., Bhatti, M.S., Anuar, N., Anuar, N., Mohan, P. and Periathamby, A., 2018. Worldwide distribution and abundance of microplastic: How dire is the situation? Waste Manage. Res. 36 (10), 873-897.

Shen, M., Zeng, Z., Li, L., Song, B., Zhou, C., Zeng, G., Zhang, Y. and Xiao, R., 2021. Microplastics act as an important protective umbrella for bacteria during water/wastewater disinfection. J. Cleaner Prod. 315, 128188.

Song, R., Sun, Y., Li, X., Ding, C., Huang, Y., Du, X. and Wang, J., 2022. Biodegradable microplastics induced the dissemination of antibiotic resistance genes and virulence factors in soil: A metagenomic perspective. Sci. Total Environ. 828, 154596.

Sun, G., Zhang, Y., Pei, L., Lou, Y., Mu, Y., Yun, J., Li, F., Wang, Y., Hao, Z., Xi, S., Li, C., Chen, C., Zhao, L., Zhang, N., Zhong, R. and Peng, Y., 2021a. Chemometric QSAR modeling of acute oral toxicity of Polycyclic Aromatic Hydrocarbons (PAHs) to rat using simple 2D descriptors and interspecies toxicity modeling with mouse. Ecotoxicol. Environ. Saf. 222, 112525.

Sun, W., Xiao, E., Häggblom, M., Krumins, V., Dong, Y., Sun, X., Li, F., Wang, Q., Li, B. and Yan, B., 2018. Bacterial survival strategies in an alkaline tailing site and the physiological mechanisms of dominant phylotypes as revealed by metagenomic analyses. Environ. Sci. Technol. 52 (22), 13370-13380.

Sun, Y., Clarke, B., Clarke, J. and Li, X., 2021b. Predicting antibiotic resistance gene abundance in activated sludge using shotgun metagenomics and machine learning. Water Res. 202, 117384.

Sun, Y., Wu, M., Xie, S., Zang, J., Wang, X., Yang, Y., Li, C. and Wang, J., 2024. Homogenization of bacterial plastisphere community in soil: a continental-scale microcosm study. ISME Commun. 4 (1), 12.

Utturkar, S., Dassanayake, A., Nagaraju, S. and Brown, S.D., 2020. Bacterial Differential Expression Analysis Methods. Methods Mol. Biol. 2096, 89-112.

Velzeboer, I., Kwadijk, C. and Koelmans, A.A., 2014. Strong sorption of PCBs to nanoplastics, microplastics, carbon nanotubes, and fullerenes. Environ. Sci. Technol. 48 (9), 4869-4876.

Wang, J., Guo, X. and Xue, J., 2021. Biofilm-developed microplastics as vectors of pollutants in aquatic environments. Environ. Sci. Technol. 55 (19), 12780-12790.

Wang, J., Peng, C., Dai, Y., Li, Y., Jiao, S., Ma, X., Liu, X. and Wang, L., 2022. Slower antibiotics degradation and higher resistance genes enrichment in plastisphere. Water Res. 222, 118920.

Wang, L., Zhang, T., Cai, T., Xiang, Q., Liu, X. and Zhu, D., 2024. The pH-specific response of soil resistome to triclocarban and arsenic co-contamination. J. Hazard. Mater. 464, 132952.

Wang, S., Shi, Y., Wang, H., Li, Z. and Zhao, M., 2023. Succession of bacteria attached to microplastics after transferring from a mariculture area to a seagrass meadow. Bull. Environ. Contam. Toxicol. 110 (4), 69.

Wang, X., Xu, Q., Jiang, M., Liu, P. and Wang, Z., 2019. Zooplankton distribution and influencing factors in the South Yellow Sea in spring. Mar. Pollut. Bull. 146, 145-154.

Wright, S.L. and Kelly, F.J., 2017. Plastic and human health: A micro issue? Environ. Sci. Technol. 51 (12), 6634-6647.

Xu, C., Lu, J., Shen, C., Wang, J. and Li, F., 2022. Deciphering the mechanisms shaping the plastisphere antibiotic resistome on riverine microplastics. Water Res. 225, 119192.

Yang, K., Chen, Q.-L., Chen, M.-L., Li, H.-Z., Liao, H., Pu, Q., Zhu, Y.-G. and Cui, L., 2020. Temporal dynamics of antibiotic resistome in the plastisphere during microbial colonization. Environ. Sci. Technol. 54 (18), 11322-11332.

Yang, Y., Li, B., Ju, F. and Zhang, T., 2013. Exploring variation of antibiotic resistance genes in activated sludge over a four-year period through a metagenomic approach. Environ. Sci. Technol. 47 (18), 10197-10205.

Yang, Y., Liu, G., Song, W., Ye, C., Lin, H., Li, Z. and Liu, W., 2019. Plastics in the marine environment are reservoirs for antibiotic and metal resistance genes. Environ. Int. 123, 79-86.

Yuan, Q., Sun, R., Yu, P., Cheng, Y., Wu, W., Bao, J. and Alvarez, P.J.J., 2022. UV-aging of microplastics increases proximal ARG donor-recipient adsorption and leaching of chemicals that synergistically enhance antibiotic resistance propagation. J. Hazard. Mater. 427, 127895.

Zadjelovic, V., Wright, R.J., Borsetto, C., Quartey, J., Cairns, T.N., Langille, M.G.I., Wellington, E.M.H. and Christie-Oleza, J.A., 2023. Microbial hitchhikers harbouring antimicrobial-resistance genes in the riverine plastisphere. Microbiome 11 (1), 225.

Zhang, G., Chen, J. and Li, W., 2022. Conjugative antibiotic-resistant plasmids promote bacterial colonization of microplastics in water environments. J. Hazard. Mater. 430, 128443.

Zhang, G. and Lu, Y., 2012. Bias-corrected random forests in regression. J. Appl. Stat. 39(1), 151-160.

Zhang, J., Chen, M., Sui, Q., Tong, J., Jiang, C., Lu, X., Zhang, Y. and Wei, Y., 2016. Impacts of addition of natural zeolite or a nitrification inhibitor on antibiotic resistance genes during sludge composting. Water Res. 91, 339-349.

Zhang, Y., Lu, J., Wu, J., Wang, J. and Luo, Y., 2020. Potential risks of microplastics combined with superbugs: Enrichment of antibiotic resistant bacteria on the surface of microplastics in mariculture system. Ecotoxicol. Environ. Saf. 187, 109852.

Zhang, Y., Zheng, Z., Chen, F., Li, M., Ding, P. and Lu, W., 2024. Accelerating the discovery of N-annulated perylene organic sensitizers via an interpretable machine learning model. J. Mol. Struct. 1296, 136855.

Zhou, Y.-X., Xie, Z.-H., Zhao, J.-X., Du, Z.-J. and Chen, G.-J., 2016. Pontibacter locisalis Sy30T sp. nov. isolated from soil collected from an abandoned saltern.

Antonie van Leeuwenhoek 109 (3), 415-420.

Zhou, Y.H. and Gallins, P.J., 2019. A review and tutorial of machine learning methods for microbiome host trait prediction. Front. Genet. 10, 579.

Zhu, D., Ma, J., Li, G., Rillig, M.C. and Zhu, Y.-G., 2022. Soil plastispheres as hotspots of antibiotic resistance genes and potential pathogens. ISME J. 16 (2), 521-532.

Zhu, L., Li, R., Yang, K., Xu, F., Lin, C., Chen, Q., Zhu, D., Sun, Q., Zhu, Y.-G. and Cui, L., 2023. Quantifying health risks of plastisphere antibiotic resistome and deciphering driving mechanisms in an urbanizing watershed. Water Res. 245, 120574.

Figure

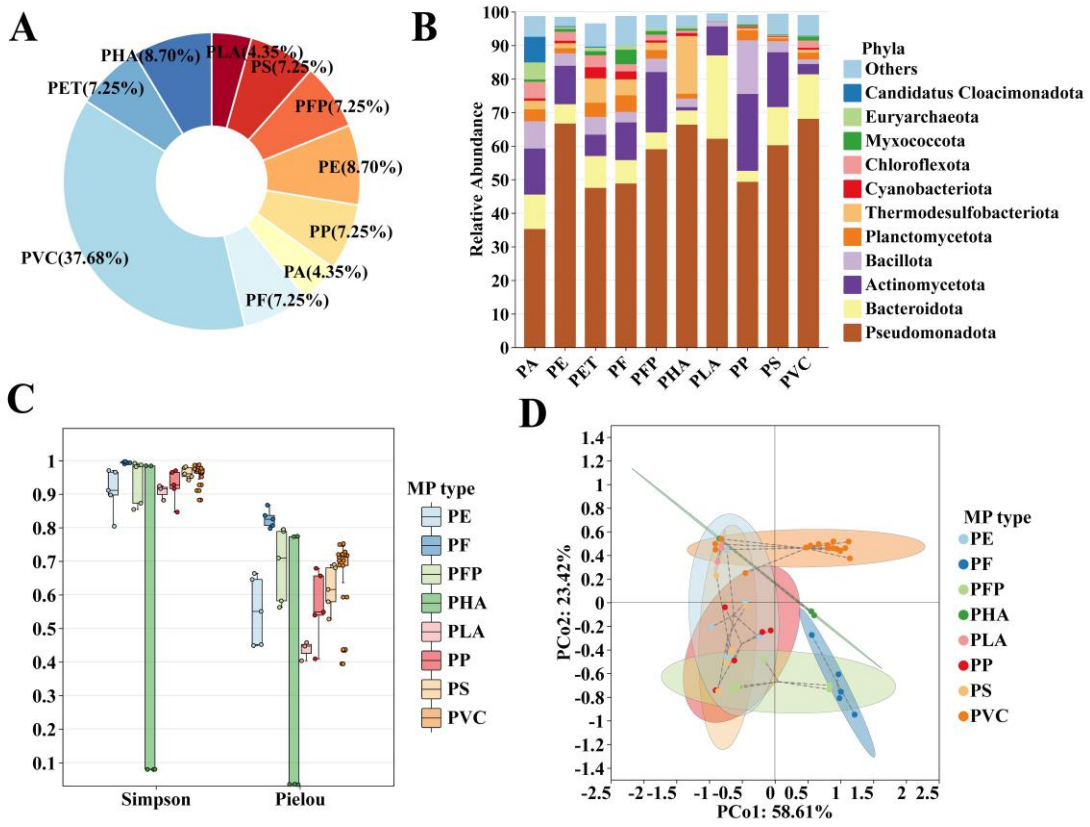


Fig. 1. (A) The distribution of MP types in the dataset. (B) The relative abundance of different phyla in MPs. (C) Diversity and evenness analysis of biological communities on different MPs. (D) The PCoA plot of microbial community differences among various MPs.

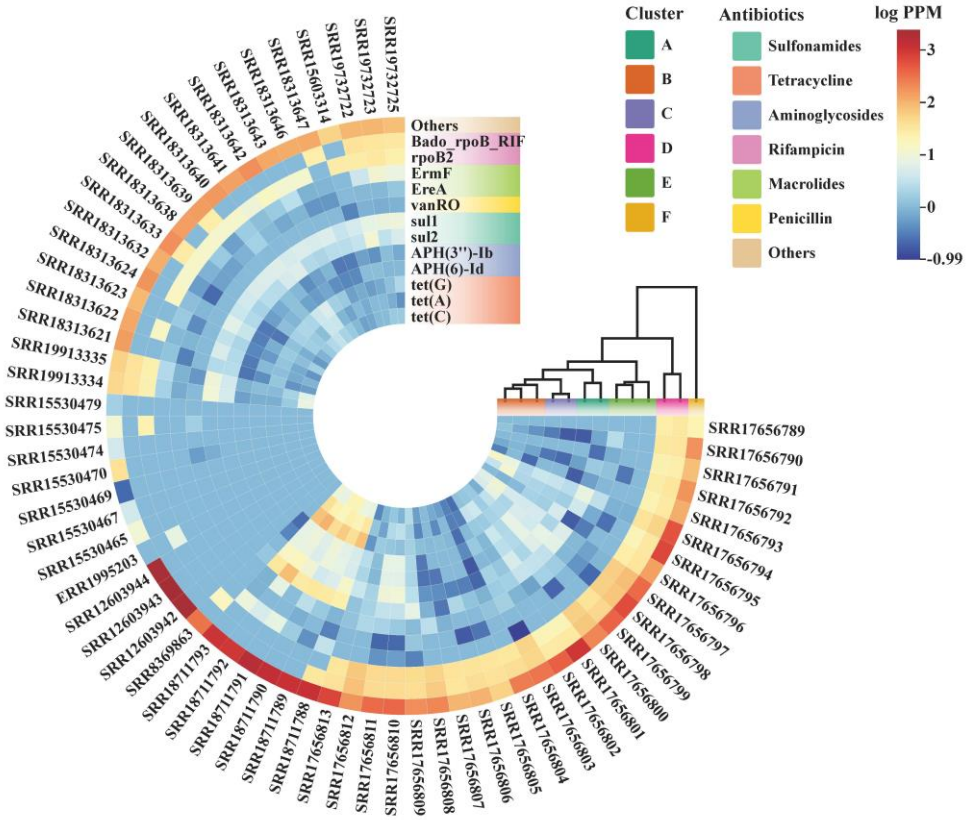


Fig. 2. Relative abundance of 12 selected ARGs in all samples.

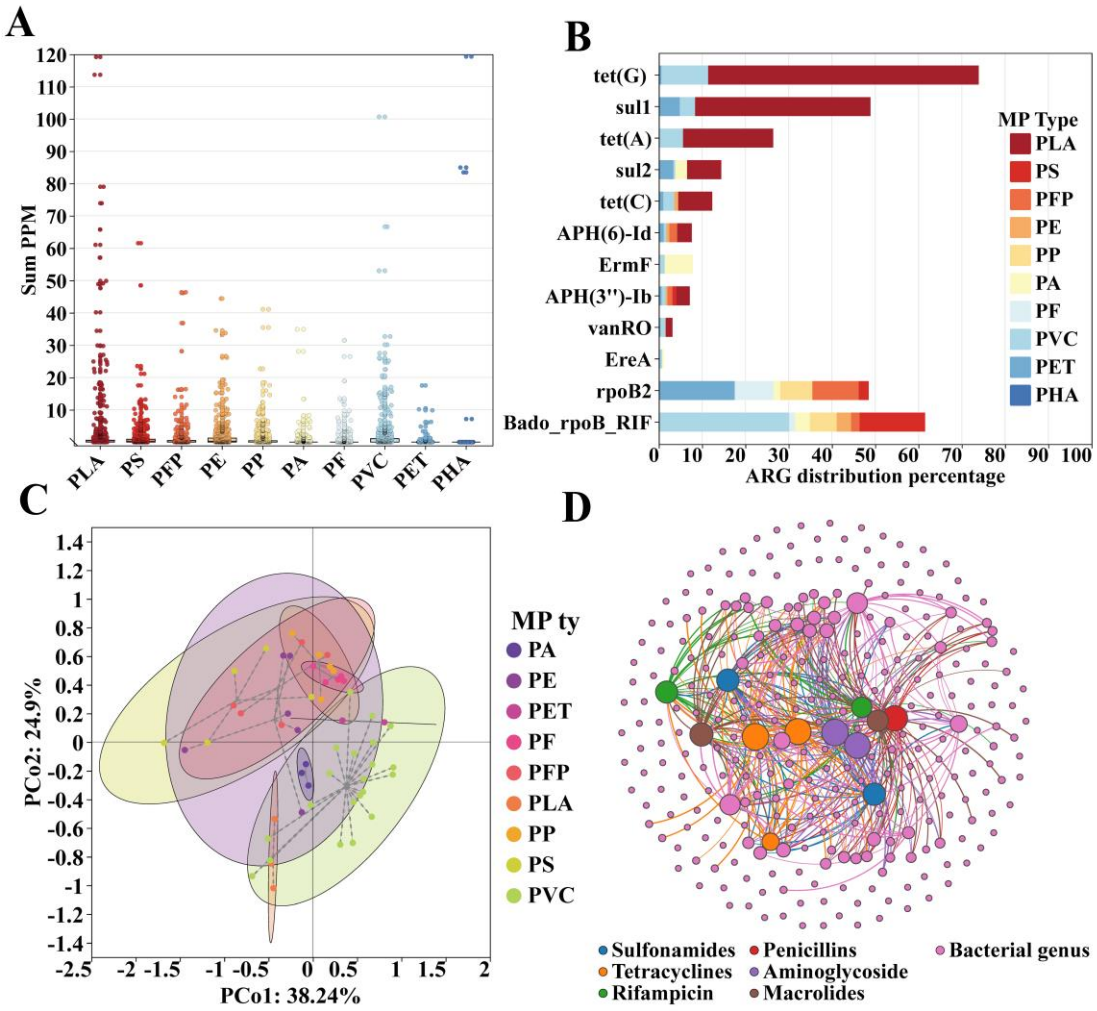


Fig. 3. (A)The sum PPM of all ARGs on different MPs. (B) The distribution trend of 12 selected ARGs on MPs. (C) The PCoA plot of the differences in ARGs on various MPs. (D) Network analysis diagram describing the symbiosis between ARGs and bacterial genera.

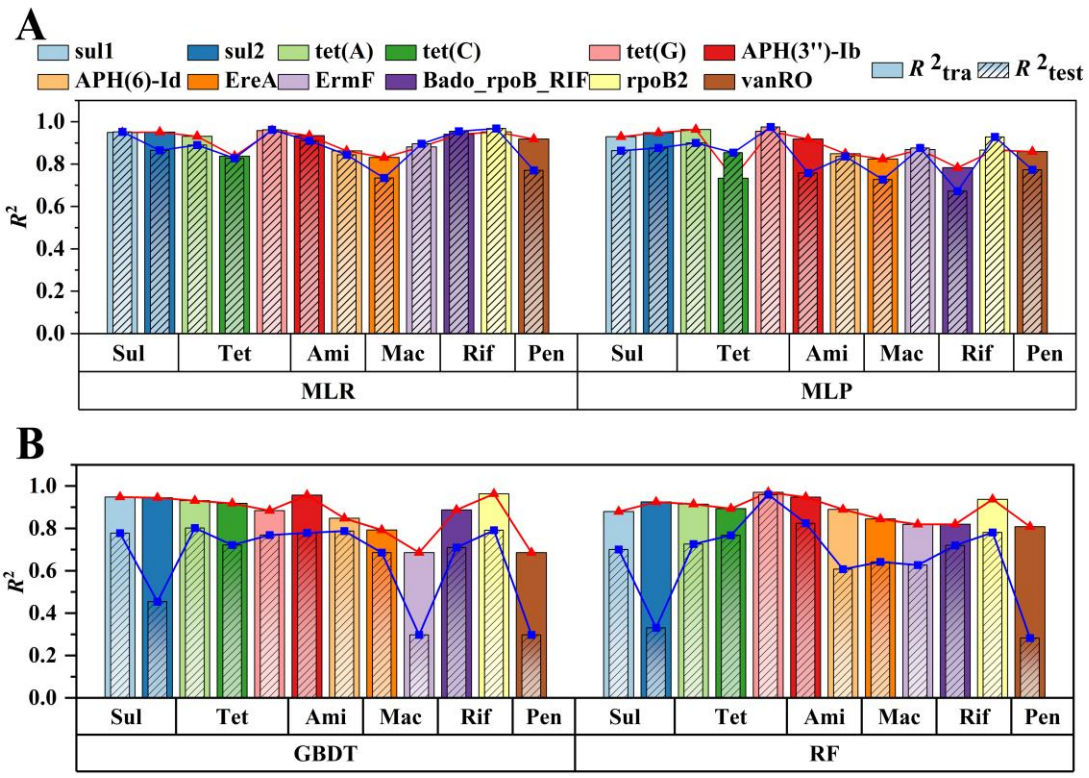


Fig. 4. Fit (R^2_{tra}) and predictive performance (R^2_{test}) of the models. The first horizontal axis is the abbreviation for drugs, and the second horizontal axis is the modeling algorithm.

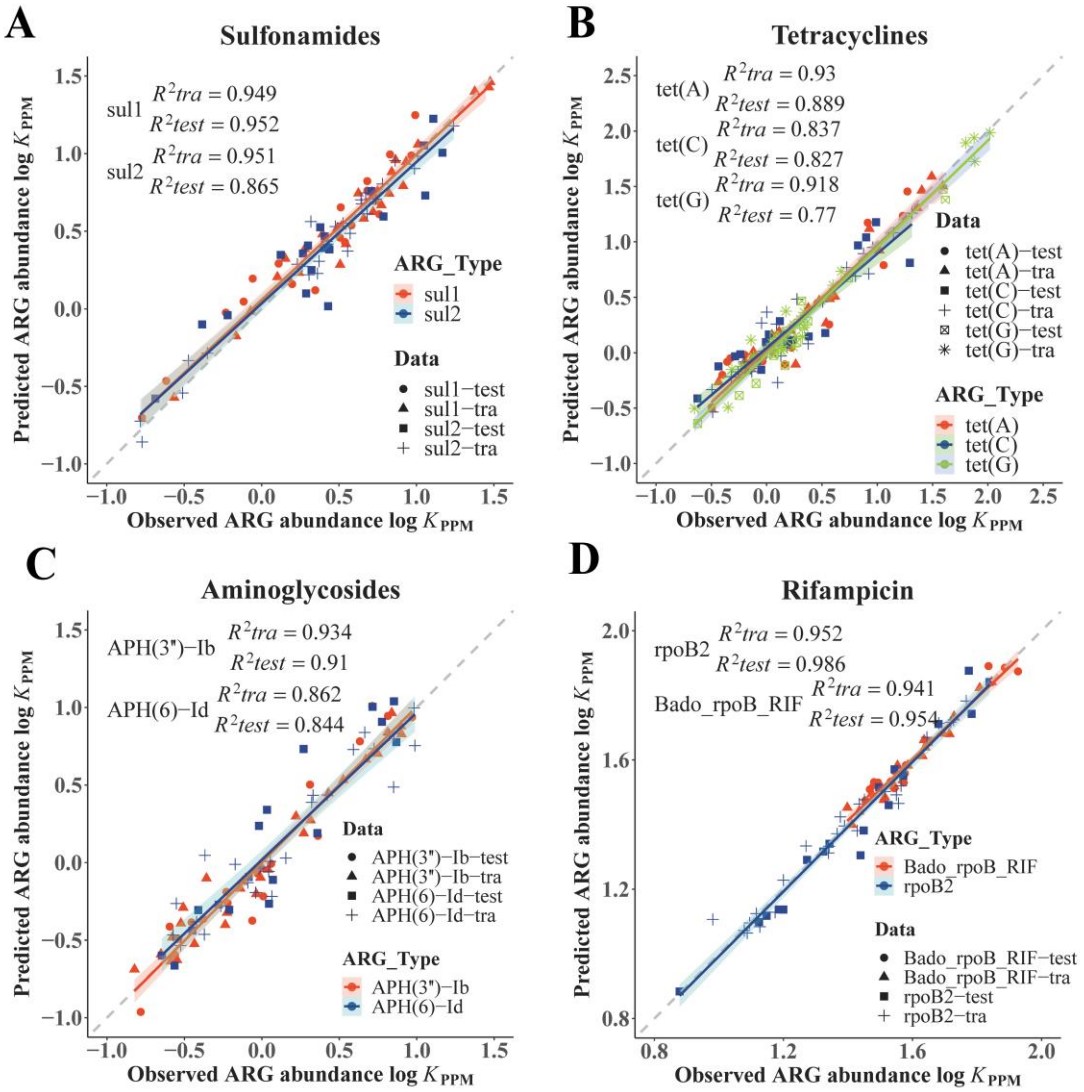


Fig. 5. Scatter plots of the predicted versus observed for log K_{PPM} based on MLR models.

Click here to access/download
Electronic Supplementary Material (for online publication only)
Supplementary.pdf

Declaration of interests

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

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