Environment International 170 (2022) 107595

Contents lists available at ScienceDirect

Environment International

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

Full length article

Metagenomic and viromic analysis reveal the anthropogenic impacts on the plasmid and phage borne transferable resistome in soil

Hu Liao [a,b], Hu Li [a,b], Chen-Song Duan [a,b], Xin-Yuan Zhou [a,b], Xin-Li An [a,b], Yong-Guan Zhu [a,c], Jian-Qiang Su [a,b,*]

[a] Key Laboratory of Urban Environment and Health, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen 361021, China
[b] University of Chinese Academy of Sciences, Beijing 100049, China
[c] State Key Lab of Urban and Regional Ecology, Research Center for Eco-environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China

ARTICLE INFO

Handling Editor: Hefa Cheng

Keywords:
Soil resistome
Land uses
Clinical ARGs
Health risk of ARGs
Anthropogenic impact

ABSTRACT

Anthropogenic land use changes have been recognized with significant effects on the abundance and diversity of antibiotic resistance genes (ARGs) in soil, but their impacts on ARGs with potential health risk remained poorly understood. In this study, paired metagenomes and viromes were obtained from soils (Anthrosols and Nitisols) with different land uses including urban parks, road verge, forests, vegetable and paddy in a subtropical city, Xiamen, and soils (Anthrosols) with various long-term fertilization treatments in Dezhou located in temperate region, respectively, to explore the influence of anthropogenic activity on soil resistome. The diversity and abundance of antibiotic resistance genes (ARGs) were profiled, and the risk associated factors of ARGs, i.e., genetic location, host, and co-existence with virulence factors (VFs), were systematically investigated at reads and contigs level. We observed that agricultural areas significantly enriched human-related ARGs and viruses, and positively related with clinical ARGs. Most of the ARG-carrying contigs were chromosomes (~85 %), while, human-related ARGs presented a higher odds ratio to locate on plasmids. Soil VFs exhibited land use pattern and distinct distribution between chromosome and plasmids, but less mobile than ARGs. Analysis of 131,014 soil viral genomes indicated that they barely encoded ARGs, nevertheless, transduction of VLPs was implicated in the spread of ARGs. The results can be mutually verified in Xiamen and Dezhou datasets. Overall, the agricultural soils with dry-farming are hotspots for the clinical ARGs, and the transmission of clinical ARGs between human dominated environments and soil is primarily mediated by plasmids, rather than bacterial chromosomes, and the transduction of human-gut related viruses could participate the process. These results highlight the importance of tracking the fate of clinical ARGs for better evaluating the impacts of human activities on soil resistome.

1. Introduction

Human activities are increasingly shaping the microbial world in soil by changes in the land use types, discharge of chemical pollutants, and direct release of human and animal gut microbiota(Zhu and Penuelas 2020). The effects of human activities on health associated microbial functional traits, antibiotic resistance genes (ARGs) and virulence factors (VFs), in soil are of particular concerns, since they are often associated with mobile genetic elements (MGEs) such as plasmids and bacteriophages which can mediate the horizontal gene transfer (HGT) and facilitate the dissemination of antimicrobial resistance and emergence of pathogens (Balasubramanian et al. 2022). It has been shown that fresh produce harbored various transferable ARGs (Blau et al. 2018), and resistant plasmid could spread into plant endophytes via soil bacteria (Xu et al. 2021), suggesting soil play an important role in the transmission of ARGs.

ARGs are ancient and ubiquitous in global environments, and soil is one of the major reservoir (Nesme and Simonet 2015). ARG profiles have been investigated in forest (Hu et al. 2017), agricultural (Forsberg et al. 2014), and urban soil (Yan et al. 2019), Significant higher diversity

Abbreviations: ARG, antibiotic resistance gene; VF, virulence factor; VLP, viral-like particle; HGT, horizontal gene transfer; MGEs, mobile genetic elements; GTA, gene transfer agents; WWTPs, wastewater treatment plants; ICEs, Integrative and conjugative elements.

* Corresponding author at: Key Laboratory of Urban Environment and Health, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen 361021, China.
E-mail address: jqsu@iue.ac.cn (J.-Q. Su).

https://doi.org/10.1016/j.envint.2022.107595
Received 3 August 2022; Received in revised form 18 October 2022; Accepted 19 October 2022
Available online 20 October 2022

and abundance of ARGs was detected in agricultural soil compared to forest soil, partially owing to the application of manure-borne fertilizers (Chen et al. 2021; Chen et al. 2016). The increased level of ARGs could be attributed to taxonomic changes, input of ARGs from manure-borne fertilizers, and on-site selection by antibiotic residues or metals. Identifying the background ARGs and allochthonous ARGs, and how they are introduced into the soil, would be helpful for distinguishing the contribution of these factors to the changes of ARG profiles in soil, which is essential for assigning proper measures to mitigate the emergence and dissemination of ARGs.

Generally, the environmental ARGs were emphasized for the clinical importance if they present resistance to clinical antibiotics, high mobility (association with MGEs), and close evolutionary relationship with those in human pathogens (Forsberg et al. 2012; Martínez et al. 2015). ARG mobility and host pathogenicity are important factors adopted in recent health risk assessment of environmental ARGs (Zhang et al. 2021; Zhang et al. 2022). Plasmid mediated conjugation and transformation, and bacteriophage mediated transduction are the major mechanisms for lateral transfer of ARGs. Plasmids facilitate the global spread of a range of ARGs (Wang et al. 2018), and have been documented as the major ARG vectors in wastewater treatment plants (WWTPs) (Che et al. 2019) and rivers (Lee et al. 2020), while the prevalence of plasmids in soils across various land use types remain elusive. Phages-like particles can mediate HGT through generalized and specialized transduction, which have been well documented in the lab (Chiang et al. 2019). Identification of ARGs in environmental virome implied that phages play important roles in the transfer of ARGs among bacteria (Calero-Cáceres et al. 2019), however, it remains controversial as the abundance of viral ARGs may have been overvalued (Chen et al. 2021; Enault et al. 2017). Characterization of the major vectors of ARGs is thus prominent for profiling the types and abundances of transferable resistome in soil.

Despite diverse and abundant ARGs have been detected in soil, only a few genes were shared between soil and human gut microbiome (Coelho et al. 2021). Recent research exploring the clinical risk of soil resistome and their response to the anthropogenic land use changes at a broad geographic scale detected only twelve pathogen-related ARG sequences (Qian et al. 2021). While, some researches demonstrated that human-related mobile resistome could drive the bloom of antibiotic resistance in human impacted environments such as WWTPs (Karkman et al. 2019) and rivers (Lee et al. 2020). That makes it crucial to understand how the MGEs (plasmids and bacteriophages) influence the dissemination of human-related ARGs in soils. Virulence factors (VFs) including virulence traits and colonization factors that enable pathogenic bacteria to colonize their hosts, establish infections, and confer virulence, thus contributing directly and indirectly to the pathogenecity and processes of infectious diseases (Wu et al. 2008). Plasmids and bacteriophages are the major molecular drivers for the HGT of VFs (Balasubramanian et al. 2022). Previous study uncovered that most of bacteria carrying mobile ARGs and VFs in urban sewage belongs to pathogens (Fresia et al. 2019). Convergence of antibiotic resistance and virulence occurred in *Klebsiella pneumoniae* (Yang et al. 2020), transfer of conjugative virulence-encoding plasmid resulted in augmented virulence in pathogen (Yang et al. 2019). This ARGs-VFs co-existence patterns have been detected in river (Jinsong et al. 2019), while, research on the distribution of mobile VFs and the co-occurrence of ARGs and VFs on MGEs in soil is still lacking.

To address these knowledge gaps, we conducted metagenomic analysis of soils from forest, park, road verge, paddy, and vegetable field in Xiamen to explore the ARG and VF profiles across land use types. Metagenomic assembly and viromic analysis were adopted to identify the major vectors and transferable ARGs and VFs. We further reanalyzed soil metagenomes and viromes previously obtained from farmlands in Dezhou to expand our data and verify the results.

2. Materials and methods

2.1. Sample description

The 49 metagenomes used herein were from our two previous studies (Chen et al. 2021; Liao et al. 2022), and were used to assess the impact of land use and long-term fertilization on soil microbial communities in Xiamen and Dezhou city, respectively. The Xiamen surface soils (0–20 cm depth, 25 samples) (24°5′ N, 118°0′ E) were sampled from urban parks, road verge, forests, vegetable and paddy (five replicates for each) on Jun 2020, their soil types were classified as Nitisols and Anthrosols. The Dezhou Anthrosols were sampled from a field experiment in the long-term experiment station of the Chinese Academy of Agricultural Sciences (CAAS), in Dezhou city of Shandong Province, China (37°20′ N, 116°38′ E) on August 2019, eight triplicate treatments (i.e. CK, 0.5 N, 1 N, 0.5SS, 1SS, 2SS, 4SS, 1CM) were implemented and the detailed information about these samples have been provided in our previous research (Chen et al. 2016). Briefly, the 0.5 N and 1 N were chemical fertilizer treatments with 65.25 and 130.5 kg urea hm^{-2}, 0.5SS, 1SS, 2SS and 4SS, were mixed fertilizer treatments with 65.25 kg urea hm^{-2} plus 4.5, 9, 18, 36 t sewage sludge hm^{-2}, respectively, 1 CM was 65.25 kg urea hm^{-2} and 10 t chicken manure hm^{-2}, and CK represents the non-fertilized control.

PowerSoil DNA isolation kit (Mo Bio Laboratories, Inc. Carlsbad, CA) was used to extract microbial genomic DNA from soil samples following the manufacturer's protocols. Sequencing libraries were prepared using the ALFA-SEQ DNA Library Prep kit (mCHIP, China) following the manufacturer's recommendations and the index codes were added. Paired-end sequencing (150 bp) of soil DNA were performed by MagiGene Co. ltd. (Guangzhou in China) on an Illumina Novaseq 6000 platform.

2.2. Identification of ARGs at reads level

Clean reads of metagenomes were obtained by quality filtering, trimming, and adaptor removing using cutadpat 2.11 (Martin 2011) and trimmomatic v0.39 (Bolger et al. 2014). The clean reads were then used to search for ARGs following ARGs-OAP v2.0 pipeline (Yin et al. 2018). The cut-off values for ARGs identification were alignment length > 75 nucleotides, e value of $< 10^{-7}$, and identity of $> 80\%$. The abundances of ARGs were normalized by 16S rRNA gene according the equation proposed by Yang *et al.* (Yang et al. 2016). The core ARGs in Xiamen represents ARGs shared by all types of land uses, and core ARGs in Dezhou agricultural areas represents ARGs shared in all treatments.

2.3. Risk evaluation of ARGs at reads level

For evaluate the health risk of ARGs in sequence level, the output of ARGs-OAP v2.0 were further processed using the package arg_ranker v2.0 (Zhang et al. 2021). The health risks of ARGs were ranked according to the criteria proposed by (Zhang et al. 2021), in which the ARGs were classified into four ranks according to three criteria: (1) associated with human, (2) gene mobility, and (3) host pathogenicity. Rank I, II, III ARGs are human-associated, rank I represent the highest health risk and rank IV is the lowest. In total, 2,759 ARG sequences were assessed, of which 764 were human-related, and only 122 were classified into Rank I (Zhang et al. 2021). The relative abundance of human-associated ARGs were represented by the sum of identified Rank I, II, III ARGs in the soil metagenomes.

2.4. Human fecal indicator in soil metagenome

We downloaded the Gut Phage Database (GPD) (Camarillo-Guerrero et al. 2021), a collection of ~ 142,000 non-redundant human gut viral genomes (>10 kb), as human fecal indicator. A threshold of 50 M reads was selected and the data was randomly subsampled without

replacement across all metagenomes using Seqtk v1.3 (https://github.com/lh3/seqtk). The subsampled metagenomes were mapped to GPD and the number of mapped reads were calculated using tool Samtools v1.11 based on the sam file through bowtie2 v2.4.2 (Langmead and Salzberg 2012). The correlation between mapped reads in metagenomes and the abundance of ARGs were evaluated to explore the potential influence of human feces on soil ARGs. The abundance of each human gut-related virus in soil metagenomes were calculated using the equation:

Abundance = **N***150/**Length**.

The **'N'** represents the number of reads which are mapped to the viral genome in metagenome, the **'Length'** represents the genomic length of human-related virus.

2.5. A comprehensive risk evaluation of intact ARGs and VFs at contigs level

The hosts of ARG were classified according to the description of Qian *et al* (Qian et al. 2021). Briefly, de novo assembly of clean reads of Xiamen metagenomes was done with package MetaSpades v3.13.0 (Nurk et al. 2013) with default parameters, whereas the clean reads of Dezhou metagenomes with eight triplicate treatments were co-assembly using package Megahit v1.2.9 (Li et al. 2016). The contigs<1,500 bp were removed and the protein-coding genes of contigs were predicted by Prodigal(Hyatt et al. 2010). Bacterial ARGs were identified through four tools: (1) the SARG database using blastp with identity ≥ 80% and query coverage (alignment length/reference ARG length) ≥ 80%; (2) the Resistance Gene Identifier (RGI) v.5.2.0 (Alcock *et al.* 2020) with default parameters; (3) the NCBI AMRFinder tool v.3.10.1 (Feldgarden et al. 2019) with default parameters; (4) the DeepARG v2.0 (Arango-Argoty et al. 2018). For those ARGs with different annotation through these tools, the annotations of these genes were assigned following the order of priority, SARG, DeepARG, RGI, AMRfinder. The ARG-carrying contigs were taxonomically classified using CAT (Meijenfeldt et al. 2019).

ARG-carrying plasmid contigs were predicted using plasflow v1.1 (Krawczyk et al. 2018) based on machine learning, and the viral contigs were identified through VIBRANT v1.2.1 (Kieft et al. 2020). Identification of integron gene cassettes was performed using the tool Blastp v2.10.0 to align the INTEGRALL database (Moura et al. 2009), and only the alignments with ≥90% amino acid identity and ≥ 90 % of the query protein length were retained. Similarly, the virulence factors (VFs) were identified through the alignment with the VFDB (set_B) using the tool Blastp v2.10.0 with e-value < 1e-10, only the identity > 80% and query coverage > 90% were retained as VFs. Transposase sequences were collected from UniRef90 (2020_03) (Bateman et al. 2020) database using the following search terms: "transposase", "Transposase". Out of 458,347 collected sequences, 182,202 transposases with insertion sequence (IS) elements based on feature descriptions were used as the reference IS transposase database. Integrative and conjugative elements (ICEs) were surveyed using a local packages ICEfinder (Liu et al. 2018). Transposases were identified through the tool Blastp to align the custom transposases database with ≥ 90% amino acid identity and ≥ 90% query coverage.

To explore the major vectors of the human-related ARGs in soils, we downloaded the Unified Human Gastrointestinal Protein (UHGP) (Almeida et al. 2021) catalog clustered with 90 % identity threshold, the encoded amino acid sequences of identified ARGs on contigs were aligned with the UHGP catalog using tool blastp with parameters -max_target_seqs 1 -evalue 1e-10. Blastp hits were filtered using cut-offs of 90 %, 95 % and 99 % identity across at least 90 % query coverage to search "potentially human-related", "human-related" and "exactly human-related" ARGs, respectively. The correlation degree between plasmids and human-related ARGs in three cut-off values of identity were evaluated using odds ratio (OR) in R.

To explore the occurrence of HGT of ARGs between environmental bacteria and human pathogens in soils, we constructed a protein database of pathogens consist of 12,220,186 proteins from 2,909 completed human pathogen genomes downloaded from NCBI. The pathogens consist of 35 organism groups (Table S1) including *Escherichia coli*, *Shigella*, *Campylobacter jejuni*, *Listeria monocytogenes*, *Klebsiella pneumoniae*, *Streptococcus pneumoniae*, *Staphylococcus aureus*, *Enterococcus faecium* etc. according to the website https://www.ncbi.nlm.nih.gov/pathogens/organisms/ on Oct-2021. All protein sequences of contigs-carried ARGs were aligned with the constructed database of pathogenic proteins using blastp with parament -max_target_seqs 1 -evalue 1e-7. The pathogen-related ARGs were designated with 99 % identity across at least 90 % query coverage. We defined the human- and pathogen-related plasmid borne ARGs as clinical ARGs. The number of mapped reads of VFs and plasmid-borne clinical ARGs in subsampled metagenomes were calculated using bowtie2 v2.4.2 and Samtools v1.11.

2.6. Phage-borne ARGs and VFs

The viromes of Xiamen and Dezhou soil were obtained from the same soil samples, in which Dezhou soil viromes has been published (Chen et al. 2021), and Xiamen viromes has been preprinted recently (Liao et al. 2022). Briefly, the Xiamen viromes (N = 25*2) extracted DNA of extracellular and intracellular virus-like particles (VLPs) respectively, while the DNA from extracellular and intracellular VLPs were together extracted in Dezhou viromes (N = 24). We merged the two viral genomic datasets through a modified protocol. Briefly, the putative viral genomes were identified through VIBRANT[33], the completeness of identified viral genomes were evaluated and flanking host regions on assembled proviruses were removed through CheckV v.0.7.0 (database v.0.6)(Nayfach et al. 2020). The viral genomes were clustered into viral populations based on the criteria at species level (Gregory et al. 2019). We selected genomes with > 50 % estimated completeness for further analysis, resulting in 131,014 viral populations. The translated proteins of viral populations were predicted by Prodigal, viral ARGs and VFs were identified using abovementioned four tools and VFDB (set_B) with the same protocol, respectively. The genome map of viral population of interested were visualized by Easyfig v2.2.5. Transductomics approach was adopted to explore the dissemination of ARGs through transduction—transfer of genetic material by viruses or VLPs (Kleiner et al. 2020) The whole metagenomic and viromic read sets were mapped onto the corresponding contigs carried ARGs using bowtie2 (Langmead and Salzberg 2012), and the resulted bam files from viromes and metagenomes were merged and sorted using samtools (Li 2011). Igvtools (IGV, v. 2.3.67) (Robinson et al. 2020) were used to generate tiled data files (.tdf) from the sorted bam files for data compression and faster access in IGV using the following parameters: count command, zoom levels: 9, using the mean, window size: 25 or 100. Read coverage patterns were displayed and visually assessed in IGV using a log scale.

2.7. Statistical analyses

Correlations between ARG abundance and other sample paraments were evaluated by linear model using function lm in R. The significance of differences in the compositions of ARGs and VFs within each of the geographic and land use sample categories was tested using a PERMANOVA test (function ''anosim'' and "adonis") as implemented in the vegan package from R. The beta-diversity (Bray-Curtis dissimilarity) of resistome and VFs were calculated using vegan in R (Oksanen et al. 2015).

2.8. Availability of data and materials

The in-house Python scripts, R scripts and relevant data used to generate figures of this study are provided with this paper and publicly available on GitHub at https://github.com/liaohu1231/health_risk_of_resistomes.git. Raw reads of 25 Xiamen metagenomes from Illumina metagenomes sequencing were submitted to the NCBI under the

project PRJNA746419, and Dezhou metagenomes were deposited at the NCBI under the project PRJNA697905.

3. Results

3.1. ARG composition in different land use type soils in Xiamen at reads level

Approximately 4.3 billion paired clean reads were obtained. A total of 426 ARG subtypes potentially conferring resistance to 24 classes of antibiotics were detected in the Xiamen soils. Multidrug resistance genes were the most abundant (from 0.08 to 0.42 copies per 16S rRNA gene copy) with 71 subtypes observed, followed by vancomycin (from 0.004 to 0.21 copies per 16S rRNA gene copy) resistance genes with 14 subtypes (Fig. 1A). There were 117–171, 119–161, 133–153, 152–284 and 130–210 ARG subtypes consist of 503–847, 543–849, 531–637, 662–1336 and 573–977 gene sequences with resistome abundances of 0.39–0.71, 0.30–0.44, 0.25–0.42, 0.22–1.12 and 0.21–0.24 per 16S rRNA gene detected in forest, park, road verge, vegetable field and paddy soils from Xiamen respectively (Fig. 1B, Fig. 1C and Table S2). ARG abundance significantly positively correlated with ARG subtype diversity in agricultural areas (vegetable and paddy), while such correlation was not observed in non-agricultural areas (forest, urban park and road verge) (Fig. 1D).

3.2. Agricultural areas are hotspot of anthropogenic ARGs

We calculated the proportions of each risk rank using ARG_ranker to evaluate the proportions of human-associated (Rank I, II and III) and Rank I ARGs in soil. The vegetable soil occupied the highest relative abundance of Rank I ARGs (up to 1.1%), and followed by paddy soils (up to 0.13%, Fig. 2A), while the forest, park and road verge barely harbored Rank I ARGs (Fig. 2A). Similarly, the vegetable soil occupied the highest percentage of human-associated ARGs (up to 2.9%), and followed by paddy soil (up to 0.75%), which were significantly higher with those in non-agricultural area (up to 0.24%).

The detected Rank I ARGs in soil could origin from exogenous input such as fertilizer of manure and irrigation. We further assessed the Dezhou samples which have experienced different long-term fertilization treatments (Chen et al. 2016), gradually increase in the percentage of human-associated ARGs were noted with sewage sludge dosages, and soil with chicken manure fertilizer contained the highest proportion of Rank I ARGs (Fig. 2B).

The percent of human-associated ARGs significantly correlated with the total abundance of ARGs in Xiamen soil (Linear model: adjusted $R^2 = 0.69$, $p < 0.01$), and also significantly correlated with the abundance of Rank I ARGs in agricultural area (Linear model: adjusted $R^2 = 0.85$, $p < 0.01$), while such correlations were not observed in non-agricultural areas (Fig. 2C and Fig. 2D). Analysis of Dezhou soil resistome showed consistent results (Figure S3). These results indicated that agricultural areas are hotspots of clinical ARGs compared with non-agricultural area. Furthermore, we observed that the number of mapped reads to GPD showed a significantly positive correlation with the abundance of ARG in agricultural areas (Fig. 2E), implying the potential impact of human feces fertilizers on the bloom of soil ARGs in agricultural areas. Furthermore, we observed that the human gut-related viral profiles based on GPD database presented a significant land use patterns (anosim test: $R = 0.35$, $p < 0.01$) (Figure S4).

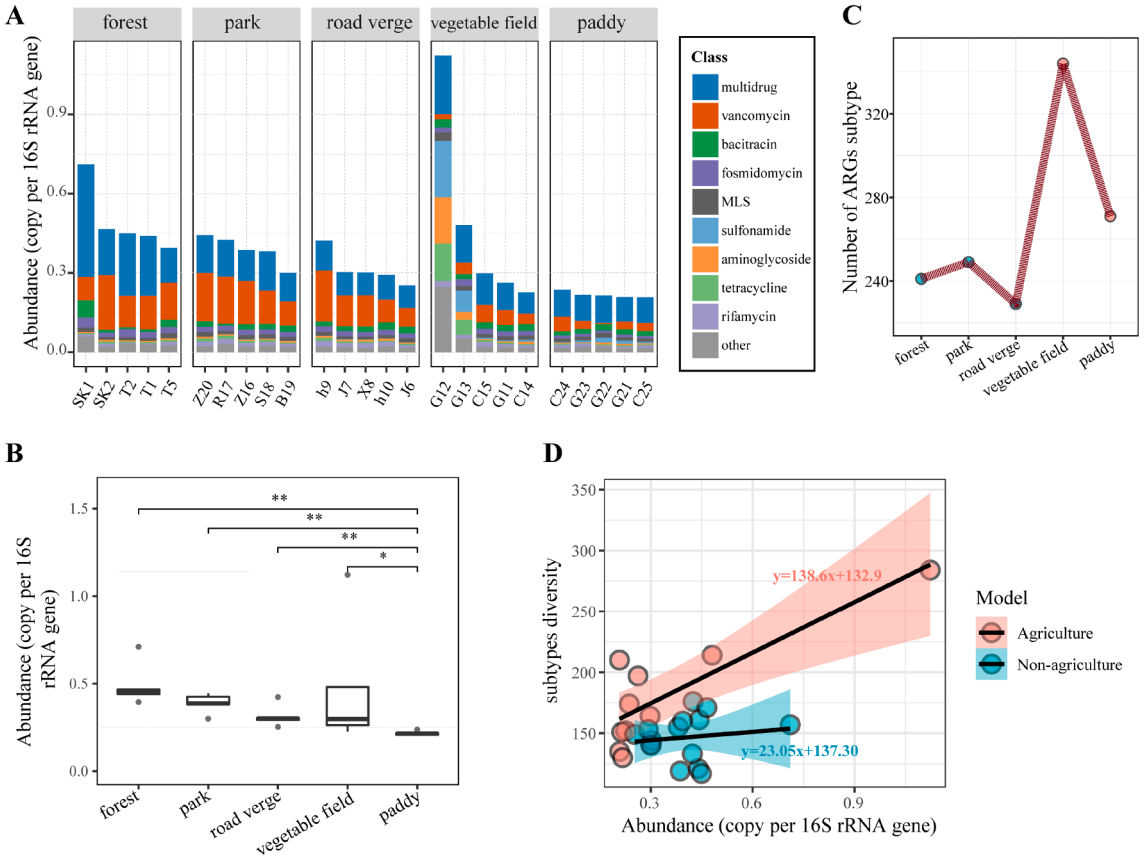

Fig. 1. (A). ARG abundance in soils with different types of land uses. **(B).** The boxplot showed ARGs abundance among different types of land uses. The median and quartiles were shown. The difference between types of land use was tested using the Wilcox.test, ** and * represents $p < 0.01$ and $p < 0.05$, respectively. **(C).** The diversity of ARG subtypes. **(D).** The linear model indicated the agricultural (paddy and vegetable plot) (adjusted $R^2 = 0.68$, $p < 0.01$) and non-agricultural area (park, road verge and forest) ($p > 0.05$) presented different correlation between ARG subtypes diversity and abundance (copy per 16S rRNA gene).

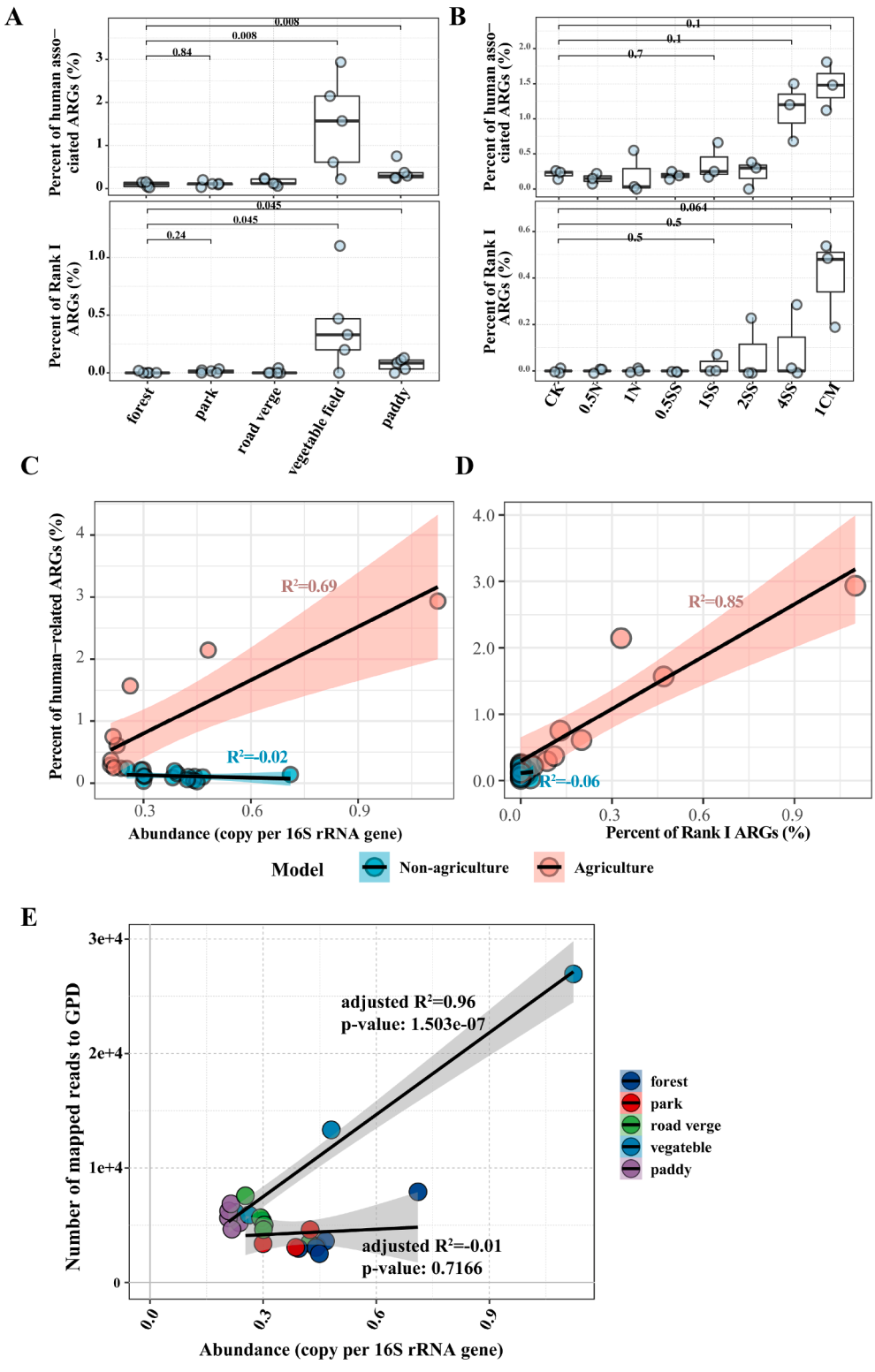

Fig. 2. **(A)** The percent of human-associated ARGs and Rank I risk ARGs in the soil with five land use types in Xiamen. The difference between types of land use was tested using the Wilcox.test. **(B)** The percent of human-associated ARGs and Rank I risk ARGs in Dezhou agricultural field with long-term fertilizer treatments. The difference between fertilized treatments was tested using the Wilcox.test. **(C)** The linear model between the percent of human-associated ARGs and total ARG abundance (copy per 16S rRNA gene) indicated a significant higher goodness of fit in agricultural area (adjusted $R^2 = 0.69$, $p < 0.01$) than non-agricultural area (adjusted $R^2 = 0.02$, $p > 0.05$). **(D)** The linear model between percent of human-associated ARGs and percent of Rank I ARGs indicated a significant higher goodness of fit in agricultural area (adjusted $R^2 = 0.85 p < 0.01$) than non-agricultural area (adjusted $R^2 = 0.06$, $p > 0.05$). **(E)** The linear model between the number of mapped reads in GPD and total ARG abundance (copy per 16S rRNA gene) indicated a significant higher goodness of fit in agricultural area (adjusted $R^2 = 0.96$) than non-agricultural area (adjusted $R^2 = -0.01$).

3.3. *Plasmids are the major vectors of human-related ARGs*

To explore whether the major vectors of human-related ARGs were plasmids, chromosome or bacteriophages in soils due to mobile human-related ARGs could threat human-health in the future. We obtained 1,388,829 contigs longer than 1,500 bp from the 25 metagenomes in Xiamen. The intact ARGs carried on contigs were identified through search against SARG v2 database, only 232 intact genes with identity > 80% and coverage > 80% were detected, of which 63 gene sequences were classified into *vanR* subtype. Only one Rank I ARG sequence (CAI84881, *tetL*) and two Rank III ARG sequences (AAA27431 and AJ249217.gene.p01 for rRNA adenine methyltransferase) were detected, and 65.9% (153 of 232) of them were unassessed in the omics-based framework (Zhang et al. 2021).

We then adopted additional three tools including AMRfinder, RGI and DeepARG to explore more soil ARGs. In total, we detected 3,124

nearly full-length ARG sequences located on 2,868 contigs, they were clustered into 2,426 contig clusters using a 95% sequence identity over 90% coverage cut-off, of which 399 (16.4%) were identified as plasmids (Table S3), 5 harbored prophages and one was identified as viral genome (G13_NODE_1_length_255533_cov_40.326533_73, 255.5kbp), but ARGs were not located in prophage genomes. Only 3.6% (88) of the contig clusters carried 102 "potentially human-related" ARGs (90 % identity over 90% coverage cut-off), of which 39.7 % (35) was identified as plasmids, while plasmids only occupied 15.6% (364 of 2338) of contigs carrying "non-human-related" ARGs, suggesting the "potentially human-related" ARGs (odds ratio (OR) = 3.58; 95% confidence interval (CI) = 2.29–5.55; $P < 0.01$) have a significant higher probability on plasmids compared to the other ARGs. We further observed that 54 contigs carried 66 "human-related" ARGs (95% identity), 53.7% (29) of them were classified as plasmids, and 32 contigs carried 43 "exactly human-related" ARGs (99% identity), where 78.1 % (25) contigs were identified as plasmids (Fig. 3). The "exactly human-related" ARGs were significantly more likely to locate on plasmids compared with "potential human-related" ARGs (the OR value were increased from 3.58 to 18.94 with identity cut-off from 90% to 99%) (Fig. 3). The "human-related" ARGs exhibited habitat preference, with a majority (68.2%, 45 of 66) assembled from agricultural soil metagenome (Fig. 3).

To explore the relationship between "human-related" ARGs and pathogenic bacteria, we constructed a protein database of pathogens consist of 12,220,186 proteins encoded by 2,909 completed genomes of human pathogens. The results of blast alignment indicated that only 0.2% (73 of 33,804) proteins which located on ARG-carrying contigs shared over 99% identity with pathogen proteins, of which 61 (83.6%) located on 26 plasmid contigs, and 37 (60.6%) of them were ARG-coding proteins. Moreover, 86.4% (32 of 37) of these pathogen-related ARGs belongs to the "exactly human-related" ARGs (Table S4), of which 27 (84.3%) were mediated by 22 plasmids, this emphasized most of the "exactly human-related" ARGs (27 of 43) were mobile and pathogen-related, thus representing high human health risk. The 27 plasmid-borne human-related ARGs were defined as clinical ARGs in this study, consisting of genes conferring resistance to tetracycline (*tetA*, *tetR*, *tetL* and *tetG*), aminoglycoside (*aph(3'')-I*, *aph(6)-I*, *aadA*), carbapenem (*OXA-347*), phenicol (chloramphenicol exporter, *cmlA*, *floR*), sulfonamide (*sul1*), fluoroquinolone (*qnrVC1*), and MLS (*msrE* and *mphE*) (Table S4 and Fig. 4A), and the majority was associated with *Pseudomonadota* (Oren and Garrity 2021) pathogens (Table S5). Intriguingly, all of these clinical ARGs were assembled from agricultural soil metagenomes, and the mapped read counts indicated that these genes were significantly enriched in agricultural area, especially vegetable soil (Fig. 4B). Moreover, ARGs were more aggregate in the clinical ARGs-carrying plasmids occupying 46.4% (39 of 84) of their encoded gene contents, while only 10.6% (2,547 of 23,920) genes were annotated as ARGs in the all 2,426 contig clusters.

We observed 53 class 1 integron gene cassettes on 21 ARG-carrying contigs, 83.0% (44 of 53) of them located on 17 plasmids and 61.3 % (27 of 53) carried ARGs, of which 15 ARGs were clinical ARGs (Table S5). All of the 53 gene cassettes were assembled from agricultural soil metagenomes. We also found 22 transposases located on ARG-carrying contigs, of which 12 located on plasmids and 10 of them (83.3 %) carried clinical ARGs. Meanwhile, only 4 ICEs were detected in ARG-carrying contigs, but ARGs was not detected in the four ICEs.

We also analyzed the ARGs in Dezhou farmland using the same workflow (Fig. 3). There were 1,285 ARGs detected in 1,269 contigs clusters, seven "potential human-related" ARGs were obtained, and all were assembled from metagenomes with 1CM and 4SS treatments. Similarly, these "potential human-related" ARGs shared at least 92 % identity with those in pathogens, of which two "exactly human-related" existed in ICM presented 100 % identity with those in pathogens (Table S4).

3.4. Soil VFs across different land uses

A total of 2131 VFs were detected on 1552 dereplicated contigs, of which 12.5% (194) contigs were classified as plasmids, the proportion was significant less than that in ARGs (OR = 0.72, p < 0.01) The

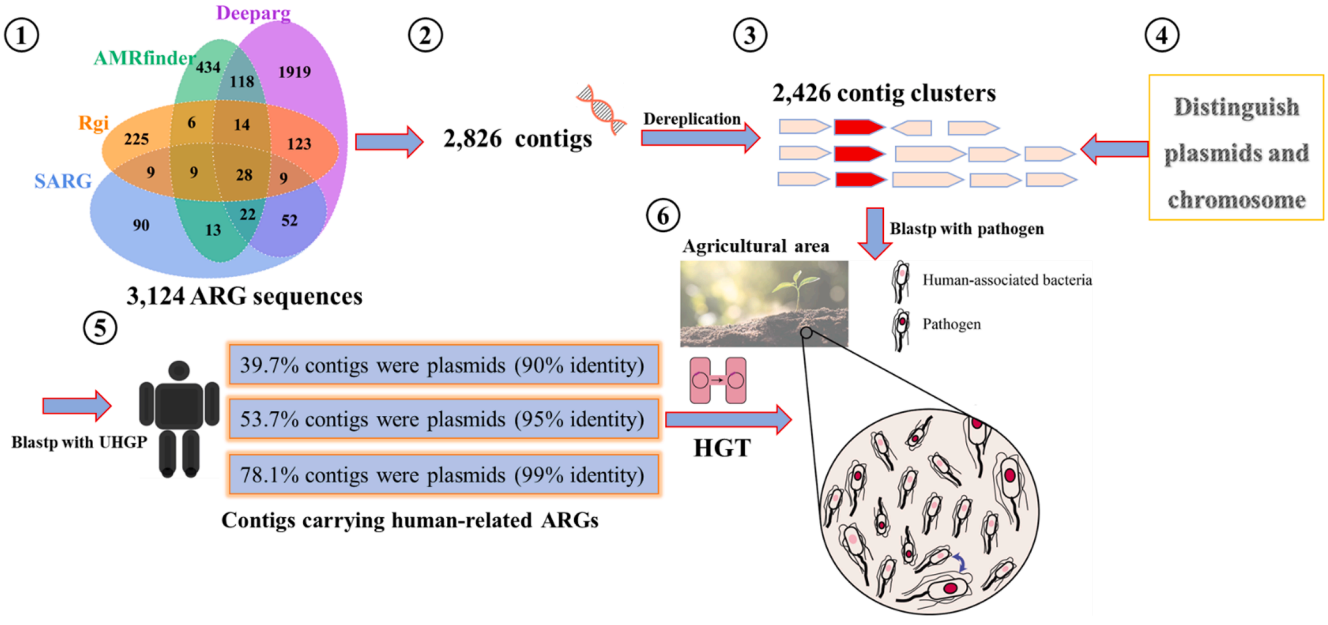

Fig. 3. The workflow for identify clinical ARGs on contigs. (1) Four methods were applied to search the ARGs located on contigs, the Venn diagram illustrated the sharing of ARGs identified by the four methods. (2) The 3,124 ARG sequences were identified on 2,826 contigs, (3) and they were dereplicated into 2,426 contig clusters. (4) plasmid and chromosome contigs were distinguished by plasflow. (5) Furtherly, we identified the "potential human-related", "human-related" and "exactly human-related" ARGs with different identity cutoff over 90% coverage (see materials and methods), and the proportions of plasmids in their carriers showed a gradual increasing from 39.7% to 78.1%. Most of "exactly human-related ARGs" were carried by plasmids. (6) The pathogen-related ARGs were explored, and most of the "exactly human-related" ARGs were pathogen-related in agricultural soil. We defined these human- and pathogen- related plasmid borne ARGs as clinical ARGs in this study.

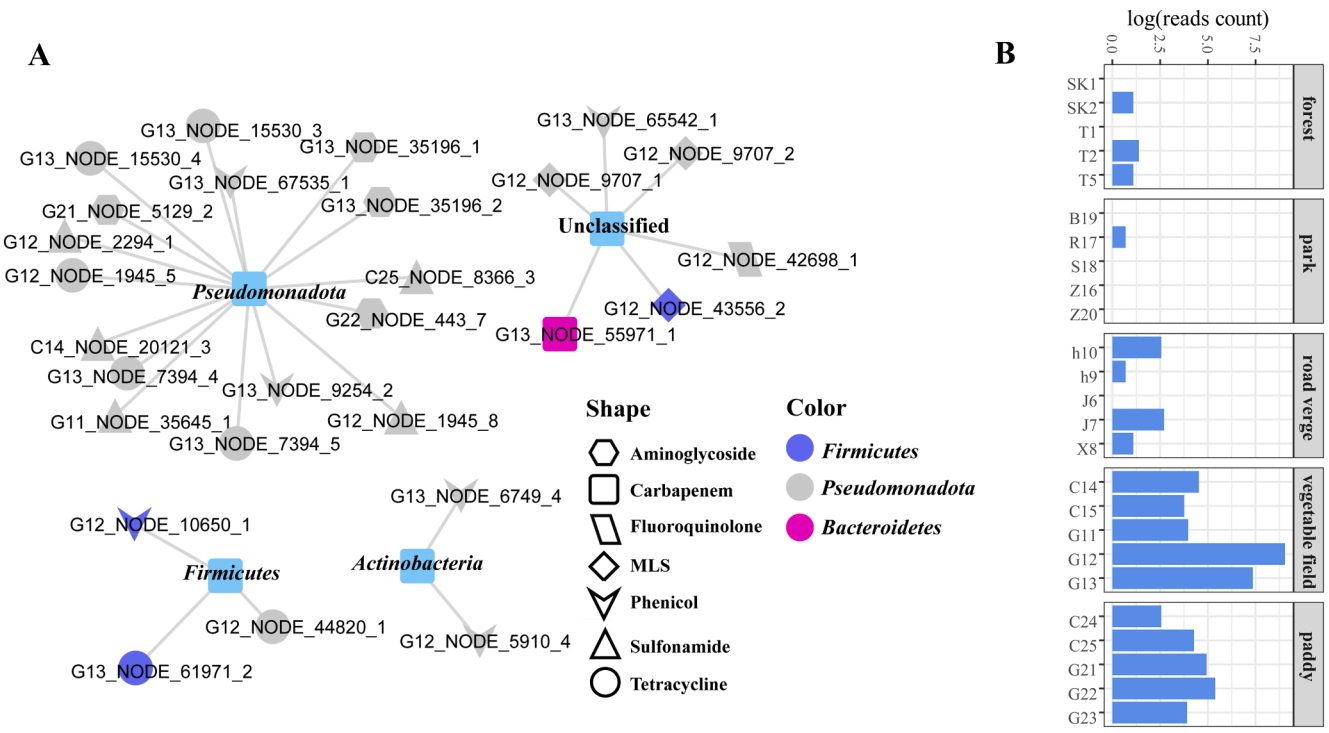

Fig. 4. **(A)** The network illustrated the potential host of the identified 27 clinical ARGs. The shape represents the class of these ARGs. The center nodes of the network represent the phylum of host of plasmid, and the color of peripheral nodes represents the taxonomy of aligned pathogen at phylum-level. **(B)** The bar plot revealed the mapped reads counts of the 22 dereplicated clinical ARGs in subsampled metagenomes (50 M reads) from different land use types.

composition of VFs showed a significant difference across different soils with land use types (Figure S5), but the R value of Anosim test was less than ARGs, suggesting the land use change had a stronger effect on ARGs than VFs. The vegetable and forest soils occupied higher abundance of VFs (Figure S6). We also explored the vectors of human- and pathogen-related VFs. In total, 46 human-related VFs were detected on 45 contigs, of which only 5 (11.1%) contigs were identified as plasmids. Of the 15 identified pathogen-related VFs, only one VFs resided on plasmids. Additionally, 46 ARG-carrying contigs were detected harboring VFs and were classified as chromosome contigs. implied that the convergence of VFs and ARGs in plasmids rarely occurred in soils.

The composition of VFs on plasmids and chromosome had a significant difference (wilcox.test, $p < 0.01$) (Figure S5), the most abundant five VFs on plasmids were LPS (CVF383,17.5%), followed by BvrR-BvrS (VF0368, 7.3%), GPL locus (CVF65, 5.5%), H-T6SS (CVF5353) and pyrimidine biosynthesis (CVF845). While the Flagella (CVF521, 8.3%), EF-Tu (CVF827, 7.7%), type IV pili (AI118, 5.7%), type IV pili biosynthesis (CVF518, 4.2%) and LPS (CVF383, 3.7%) were the dominant VFs on chromosomes (Figure S7). The taxonomy of contigs were assigned using tool CAT, only 302 contigs can be assigned into a genus, most of them were classified as *Pseudomonas* (n = 127) or *Mycobacterium* (Table S6).

3.5. Virus-like particles mediated ARG through transduction

To explore the potential ARGs carried by viral populations, we re-analyzed 50 viromes in Xiamen soil from the parallel research and obtained 89,583 viral populations. To expanded the soil viral database, another dataset from Dezhou soil were merged into the database, resulting in 131,014 viral populations with 4,622,088 protein-coding genes (longer than 20 amino acids). Only 17 (0.13%) viral populations borne ARGs were detected using the same criteria for the identification of ARGs in metagenome. All of them were annotated as the dihydrofolate reductase, and shared low homogeneity with those from corresponding metagenome, suggesting that viral genomes in soils rarely encoded ARGs. In addition, only 14 VFs were detected in 9 viral populations (Figure S8), most of which (13 of 14) were carried by lysogenic phages, and most of their encoded proteins were homologs of LPS glucosylation (CVF473) of *Shigella flexneri* 5 str. 8401.

In the generalized transduction and gene transfer agent (GTA)-like model, DNA fragments of host can be encapsulated into capsids of virus-like particles (VLPs), rather than integrated into the viral genomes. To assess potential transduction of ARGs via VLPs in complex soil microbiomes, we applied a modified transductomic approach (see materials and methods) proposed by Kleiner *et al* (Kleiner et al. 2020). Of the 2,426 ARG-carrying contigs, 257 contigs showed a read coverage pattern (mapped > 200 reads) in at least one virome, indicating potential mobilization of ARGs mediated by VLPs, in which 98 contigs were longer than 40 kbp, suggesting that potential GTA-like or generalized transduction occurred on these contigs. ARGs were detected in these transducted DNA fragments (Figure S9A and Figure S9B). For example, one of the contigs (SK1_NODE_79_length_62337_cov_12.376802, *Burkholderiaceae*) showed several sharp coverages drops in the viromic reads. Another shorter contigs (<40kbp, SK1_NODE_971_length_15515_cov_9.889521, *Burkholderiaceae*) showed two high coverage peaks in regions encoding "human-related" ARGs *adeF* in the viromes corresponded with metagenomes (Figure S9C). The clinical ARGs *tetG* and *sul1* can be transducted by virus-like particles as well (Figure S9D).

4. Discussions

In this study, a comprehensive metagenomic analysis was conducted to explore the soil ARG profiles across different land use types, including forest, paddy, vegetable, road verge, park in a subtropical city, Xiamen. Soil metagenomes from temperate agricultural soil (Dezhou) was used as complementary data. The relative abundance of ARGs in paddy soil was significantly lower than that in non-agricultural soil (Fig. 1B), which is contrasting with many previous studies (Xiang et al. 2019; Zhu et al. 2021). The detected number of ARGs were significantly higher in

agricultural soil than non-agricultural soil (Fig. 1B), and the core ARGs in agricultural area occupied a significant lower proportion than non-agricultural area (Fig. S2b), indicating input of exogenous ARGs in agricultural soils caused by frequent agricultural practice. However, the abundance of human-associated (Rank I, II, III) and Rank I risk ARGs classified by an omics-based framework showed a significantly higher abundance in agricultural dry-farming soils from Xiamen (Fig. 2A), and elevated abundance of these ARGs in Dezhou soil with sewage sludge and manure fertilizers (Fig. 2B). Significantly positive correlations between the abundance of these ARGs and total soil ARG diversity and abundance were only observed in agricultural soils (Fig. 1c and Fig. 3c). These results suggest that human activities, such as manure and sludge fertilization, are responsible for the enrichment of human-associated ARGs which partially contribute to the variations of total resistome in agricultural soils (Fang et al. 2018; Qian et al. 2021).

To fully capture the diversity of ARGs and explore the vectors of ARGs, the reads were assembled and four tools were adopted to identify ARGs on contigs as many as possible (Fig. 3). Unlike previous study showing that plasmids were the dominant vectors of ARGs in WWTPs (Che et al. 2019), most of the identified ARGs resided on chromosome contigs (75.8%) in this study, suggesting a different pattern between soil and WWTPs. However, our workflow can find more intact human-related ARGs in soils compared to omics-based framework proposed by Zhang *et al.* (Zhang et al. 2021), thus the enrichment of human-related ARGs on plasmids was observed in agricultural soils (Fig. 3). Most of these plasmids carried integrase or transposase, implying that plasmids serve as a dominant shuttle in the transmission of anthropogenic ARGs to the soil, which was in consistence with previous reports showed that human-related ARGs in aquatic ecosystem were more frequently found in the MGE context (Lee et al. 2020). Moreover, we found that most of the human-related ARGs in soils were related to pathogens across three phyla including *Pseudomonadota*, *Firmicutes* and *Bacteridetes*, and most of them were mobile (Fig. 3 and Table S3). We defined these human- and pathogen-related plasmid borne ARGs as clinical ARGs with "current threat" to human health in accordance to the criteria proposed by Zhang *et al.* (Zhang et al. 2021) for omics-based rating of the health risk of environmental ARGs. These clinical ARGs were enriched in the agricultural areas and all the clinical ARG-carried contigs were assembled from agricultural soils, indicating that the agricultural areas are hotspot of clinical ARGs compared to forest and urban soil (Fig. 4). Overall, these results suggest increased HGT of ARGs events mediated by plasmids between human pathogens and edaphic bacteria upon agricultural activities.

To explore the role of phages in the spread of ARGs, we applied same criteria for metagenome-assembled contig analysis to identify ARGs located on viral genomes. Only 17 phage-borne ARGs were identified and the percent of ARG-carrying phages were extremely low (0.013%), further supporting that the plasmids are the major MGEs carring human-related ARGs in soil. The results implying the All of phage-borne ARGs encode the dihydrofolate reductase (DHFR) pertaining to the folate biosynthetic pathway (van Hoek et al. 2011), which could lead to an increased trimethoprim resistance in host (Peters et al. 2019), and could also provide additional functions such as machinery for replication (Moon et al. 2020). Nevertheless, transductomics analysis indicated transfer of ARGs through transduction of VLPs (Figure 5). A recent study reported that transduction as an important ARG propagation mechanism occurred under complex hydrodynamic conditions with changed frequency (Sun et al. 2021). There were debated results that whether bacteriophages serve as reservoirs and transfer agents of ARGs (Chevallereau et al. 2021; Enault et al. 2017). Despite a substantial portion of ARG-carrying bacteriophages was identified in the human gut phage database MGV (Nayfach et al. 2021), the frequency of ARG-carrying bacteriophages in agricultural slurry viromes was low (Cook et al. 2021). The discrepancy could be attributed to the different threshold when identifying ARGs from phage genomes, or to the ARG levels of host since human gut microbiome was well documented as rich reservoirs of ARGs Our results, based on strict ARG identification in 131,014 obtained viral genomes, provide strong evidence that soil bacteriophages rarely encode ARGs, but generalized transductions of VLPs is implicated in the spread of ARGs (Enault et al. 2017). Furthermore, the mapped read counts of GPD presented a significant positive relationship with abundance of ARGs in agricultural areas (Fig. 2E), and the profile of human-related viral communities in soils illustrated a significant land use pattern (Figure S4), implying that the transduction of human gut viruses participated in the transmission of human-related ARGs from human-gut to soils. However, the contribution of transduction mediated by specific virus to the spread of ARG remain difficult to quantify, the development of single-virus technology could resolve the issue in the future.

We further conducted similar analysis of VFs for better understanding of anthropogenic effects on human health associated microbial functional traits in soils. The composition of VFs showed a significant difference across different soils with land use types (Figure S5), but the R value of ANOSIM test was less than ARGs, suggesting land use change had a stronger effect on ARGs than VFs. The proportion of VF-carried contigs was significantly lower than that of ARG-carried contigs, suggesting a weaker mobile potential of VFs than ARGs. In addition, human-pathogen-related VFs was less prevalent in plasmids. Significantly different compositions between plasmid-borne and chromosome-borne VFs (Figure S6) was observed, of which LPS and two-component system BvrR/BvrS were enriched in plasmids, while the Flagella and Type IV pili related to the adherence were enriched in chromosome, suggesting the complementary role of plasmid-borne VFs in affecting the virulence of host bacteria (López-Goñi et al. 2003). The vegetable soils harbored the highest abundance of VFs, suggesting that dry farming also contribute to increased abundance of VFs in consistent with clinical ARGs in agricultural area. Previous research pinpointed the convergence of VFs and ARGs in a single plasmid potentially promoting the emergence of hypervirulent and multidrug-resistant *Klebsiella pneumoniae* strains (Yang et al. 2020). While, co-occurrence of ARGs and VFs on plasmids was not detected in this study, implying that the convergence of VFs and ARGs on plasmids rarely occurred in soils. Nevertheless, most of the assembled plasmids were incomplete, thus further investigation with extended dataset is warranted to verify this finding.

5. Conclusions

The soil ARG profiles exhibits significant land use patterns of which agricultural and non-agricultural areas presents distinct modes as a consequence of different forms and intensities of human activity, such as application of human and animal feces associated fertilizers. Agricultural area, particularly with dry-farming, occupy higher percentage of human-related and clinical ARGs, demonstrating that agricultural soils are hotspot of clinical ARGs compared with non-agricultural soils. Plasmid, rather than phages, are the major vectors for the transferable resistome in soil and primarily mediated the transfer of clinical ARGs between the human-dominated environments and soils. Apparently, more attention should be paid to clinical ARGs in agricultural areas since these ARGs could be easily transferred into human pathogens. Nevertheless, those plasmid-borne but not pathogen-related ARGs are non-negligible given that these transferable ARGs may serve as sources for the emerging of new antibiotic resistance determinants in pathogens.

Declarations

Author Information.

Hu Liao wrote the main manuscript and executed all analysis and experiments, **Hu Li** provided constructive ideas, **Chen-Song Duan, Xin-Yuan Zhou, Xin-Li An** provided some suggestions, **Yong-Guan Zhu** provided resource, **Jian-Qiang Su** provided resource, writing & editing, supervision, funding acquisition.

Funding Sources.

This study was financially supported by the National Key Research and Development Program of China (2020YFC1806902) and the National Natural Science Foundation of China (42021005).

Ethics approval and consent to participate.
Not applicable.
Consent for publication.
Not applicable.
Competing interests.
The authors declare no conflict of interest.

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.

Data availability

Data will be made available on request.

Acknowledgements

Not applicable

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envint.2022.107595.

References

Alcock, B.P., Raphenya, A.R., Lau, T., Tsang, K.K., Mc Arthur, A.G.C.A.R.D., 2020. antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucl Acids Res 2019, 48.

Almeida, A., Nayfach, S., Boland, M., Strozzi, F., Beracochea, M., Shi, Z.J., Pollard, K.S., Sakharova, E., Parks, D.H., Hugenholtz, P., Segata, N., Kyrpides, N.C., Finn, R.D., 2021. A unified catalog of 204,938 reference genomes from the human gut microbiome. Nat Biotech 39, 105–114.

Arango-Argoty, G., Garner, E., Pruden, A., Heath, L.S., Vikesland, P., Zhang, L., 2018. DeepARG: a deep learning approach for predicting antibiotic resistance genes from metagenomic data. Microbiome 6, 23.

Balasubramanian, D., López-Pérez, M., Grant, T.A., Ogbunugafor, C.B., Almagro-Moreno, S., 2022. Molecular mechanisms and drivers of pathogen emergence. Trends Microbiol.

Bateman, A., Martin, M.-J., Orchard, S., Magrane, M., Agivetova, R., Ahmad, S., Alpi, E., Bowler-Barnett, E., Britto, R., Bursteinas, B., Bye-A-Jee, H., Coetzee, R., Cukura, A., Silva, A., Denny, P., Dogan, T., Ebenezer, T., Fan, J., Castro, L., Teodoro, D., 2020. UniProt: the universal protein knowledgebase in 2021. Nucl Acids Res 49.

Blau, K., Bettermann, A., Jechalke, S., Fornefeld, E., Vanrobaeys, Y., Stalder, T., Top, E. M., Smalla, K., 2018. The Transferable Resistome of Produce. mBio 9.

Bolger, A.M., Marc, L., Bjoern, U., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2114–2120.

Calero-Cáceres, W., Ye, M., Balcázar, J., 2019. Bacteriophages as Environmental Reservoirs of Antibiotic Resistance. Trends Microbiol.

Camarillo-Guerrero, L.F., Almeida, A., Rangel-Pineros, G., Finn, R.D., Lawley, T.D., 2021. Massive expansion of human gut bacteriophage diversity. Cell 184 (1098–1109), e1099.

Che, Y., Xia, Y., Liu, L., Li, A., Yang, Y., Zhang, T., 2019. Mobile antibiotic resistome in wastewater treatment plants revealed by Nanopore metagenomic sequencing. Microbiome 7.

Chen, Q., An, X., Li, H., Su, J., Ma, Y., Zhu, Y.G., 2016. Long-term field application of sewage sludge increases the abundance of antibiotic resistance genes in soil. Environ Int 92–93, 1–10.

Chen, M.-L., An, X.-L., Liao, H., Yang, K., Su, J.-Q., Zhu, Y.-G., 2021. Viral Community and Virus-Associated Antibiotic Resistance Genes in Soils Amended with Organic Fertilizers. Environ Sci Technol.

Chevallereau, A., Pons, B.J., van Houte, S., Westra, E.R., 2021. Interactions between bacterial and phage communities in natural environments. Nat Rev Microbiol.

Chiang, Y.N., Penadés, J.R., Chen, J., 2019. Genetic transduction by phages and chromosomal islands: The new and noncanonical. PLoS Pathogens 15, e1007878.

Coelho, L.P., Alves, R., Del Rio, A.R., Myers, P.N., Cantalapiedra, C.P., Giner-Lamia, J., Schmidt, T.S., Mende, D.R., Orakov, A., Letunic, I., Hildebrand, F., Van Rossum, T., Forslund, S.K., Khedkar, S., Maistrenko, O.M., Pan, S., Jia, L., Ferretti, P., Sunagawa, S., Zhao, X.M., Nielsen, H.B., Huerta-Cepas, J., Bork, P., 2021. Towards the biogeography of prokaryotic genes. Nature.

Cook, R., Hooton, S., Trivedi, U., King, L., Dodd, C.E.R., Hobman, J.L., Stekel, D.J., Jones, M.A., Millard, A.D., 2021. Hybrid assembly of an agricultural slurry virome reveals a diverse and stable community with the potential to alter the metabolism and virulence of veterinary pathogens. Microbiome 9, 65.

Enault, F., Briet, A., Bouteille, L., Roux, S., Sullivan, M.B., Petit, M.A., 2017. Phages rarely encode antibiotic resistance genes: a cautionary tale for virome analyses. ISME J 11, 237–247.

Fang; Wang; Min; Xu; Robert; D; Stedtfeld; Hongjie; Sheng; Jianbo. Long-term Effect of Different Fertilization and Cropping Systems on the Soil Antibiotic Resistome. Environ Sci Technol 2018.

Feldgarden, M., Brover, V., Haft, D.H., Prasad, A.B., Klimke, W., 2019. Validating the NCBI AMRFinder Tool and Resistance Gene Database Using Antimicrobial Resistance Genotype-Phenotype Correlations in a Collection of NARMS Isolates. Antimicrobial Agents and Chemotherapy 63.

Forsberg, K., Patel, S., Gibson, M., Lauber, C., Knight, R., Fierer, N., Dantas, G., 2014. Bacterial phylogeny structures soil resistomes across habitats. Nature 509.

Forsberg, K.J., Reyes, A., Wang, B., Selleck, E.M., Sommer, M.O., Dantas, G., 2012. The shared antibiotic resistome of soil bacteria and human pathogens. Science 337, 1107–1111.

Fresia, P., Antelo, V., Salazar, C., Gimenez, M., D'Alessandro, B., Afshinnekoo, E., Mason, C., Gonnet, G.H., Iraola, G., 2019. Urban metagenomics uncover antibiotic resistance reservoirs in coastal beach and sewage waters. Microbiome 7, 35.

Gregory, A.C., Zayed, A.A., Conceicao-Neto, N., Temperton, B., Bolduc, B., Alberti, A., Ardyna, M., Arkhipova, K., Carmichael, M., Cruaud, C., Dimier, C., Dominguez-Huerta, G., Ferland, J., Kandels, S., Liu, Y., Marec, C., Pesant, S., Picheral, M., Pisarev, S., Poulain, J., Tremblay, J.E., Vik, D., Tara Oceans, C., Babin, M., Bowler, C., Culley, A.I., de Vargas, C., Dutilh, B.E., Iudicone, D., Karp-Boss, L., Roux, S., Sunagawa, S., Wincker, P., Sullivan, M.B., 2019. Marine DNA Viral Macro- and Microdiversity from Pole to Pole. Cell 177 (1109–1123), e1114.

Hu, H., Wang, J., Singh, B.K., Liu, Y., He, J., 2017. Diversity of herbaceous plants and bacterial communities regulates soil resistome across forest biomes. Environ Microbiol 20.

Hyatt, D., Chen, G., Locascio, P.F., Land, M., Larimer, F.W., Hauser, L., 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11.

Jinsong, L., Mao, G., Yin, X., Ma, L., Liu, L., Bai, Y., Zhang, T., Qu, J., 2019. Identification and quantification of bacterial genomes carrying antibiotic resistance genes and virulence factor genes for aquatic microbiological risk assessment. Water Research 168, 115160.

Karkman, A., Parnanen, K., Larsson, D.G.J., 2019. Fecal pollution can explain antibiotic resistance gene abundances in anthropogenically impacted environments. Nat Commun 10, 80.

Kieft, K., Zhou, Z., Anantharaman, K., 2020. VIBRANT: automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences. Microbiome 8, 90.

Kleiner, M., Bushnell, B., Sanderson, K.E., Hooper, L.V., Duerkop, B.A., 2020. Transductomics: sequencing-based detection and analysis of transduced DNA in pure cultures and microbial communities. Microbiome 8, 158.

Krawczyk; Pawel; S.; Lipinski; Leszek; Dziembowski; Andrzej. PlasFlow: predicting plasmid sequences in metagenomic data using genome signatures. Nucl Acids Res 2018.

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9, 357–359.

Lee, K., Kim, D.W., Lee, D.H., Kim, Y.S., Bu, J.H., Cha, J.H., Thawng, C.N., Hwang, E.M., Seong, H.J., Sul, W.J., Wellington, E.M.H., Quince, C., Cha, C.J., 2020. Mobile resistome of human gut and pathogen drives anthropogenic bloom of antibiotic resistance. Microbiome 8, 2.

Li, H., 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27, 2987–2993.

Li, D., Luo, R., Liu, C.-M., Leung, C.-M., Ting, H.-F., Sadakane, K., Yamashita, H., Lam, T.-W., 2016. MEGAHIT v1.0: A Fast and Scalable Metagenome Assembler driven by Advanced Methodologies and Community Practices. Methods 102.

Liao, H., Li, H., Chen-Song, D., Xin-Yuan, Z., Qiu-Ping, L., Xinli, A., Yong-guan, Z., Jian-Qiang, S., 2022. Response of soil viral communities to land use changes. Nature Communications 13, 6027.

Liu, M., Li, X., Xie, Y., Bi, D., Sun, J., Li, J., Tai, C., Deng, Z., Ou, H.-Y., 2018. ICEberg 2.0: an updated database of bacterial integrative and conjugative elements. Nucl Acids Res 47.

López-Goñi, I., Guzmán-Verri, C., Manterola, L., Sola, A., Moriyón, I., Moreno, E., 2003. Regulation of Brucella virulence by the two-component system BvrR/BvrS. Veterinary microbiology 90, 329–339.

Martin, M. CUTADAPT removes adapter sequences from high-throughput sequencing reads. EMBnetjournal 2011;17.

Martínez, J., Coque, T.M., Baquero, F., 2015. What is a resistance gene? Ranking risk in resistomes. Nat Rev Microbiol 13, 116–123.

Meijenfeldt, F.A.B.V.; Arkhipova, K.; Cambuy, D.D.; Coutinho, F.H.; Dutilh, B.E. Robust taxonomic classification of uncharted microbial sequences and bins with CAT and BAT. Genome Biol 2019;20:217.

Moon, K., Jeon, J.H., Kang, I., Park, K.S., Lee, K., Cha, C.J., Lee, S.H., Cho, J.C., 2020. Freshwater viral metagenome reveals novel and functional phage-borne antibiotic resistance genes. Microbiome 8, 75.

Moura, A., Soares, M., Pereira, C., Leitão, N., Henriques, I., Correia, A., 2009. INTEGRALL: a database and search engine for integrons, integrases and gene cassettes. Bioinformatics 25, 1096–1098.

Nayfach, S., Camargo, A.P., Schulz, F., Eloe-Fadrosh, E., Roux, S., Kyrpides, N.C., 2020. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. Nat. Biotech.

Nayfach, S., Paez-Espino, D., Call, L., Low, S.J., Sberro, H., Ivanova, N.N., Proal, A.D., Fischbach, M.A., Bhatt, A.S., Hugenholtz, P., Kyrpides, N.C., 2021. Metagenomic

compendium of 189,680 DNA viruses from the human gut microbiome. Nat Microbiol 6, 960–970.
Nesme; Simonet. The soil resistome: a critical review on antibiotic resistance origins, ecology and dissemination potential in telluric bacteria. Environ Microbiol 2015.
Nurk, S., Bankevich, A., Antipov, D., Gurevich, A., Korobeynikov, A., Lapidus, A., Prjibelsky, A., Pyshkin, A., Sirotkin, A., Sirotkin, Y., Stepanauskas, R., McLean, J., Lasken, R., Clingenpeel, S.R., Woyke, T., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2013. Assembling Genomes and Mini-metagenomes from Highly Chimeric Reads. Springer, Berlin Heidelberg, Berlin, Heidelberg.
Oksanen, J.; Blanchet, F.G.; Kindt, R.; Legendre, P.; Minchin, P.; O'Hara, R.B.; Simpson, G.; Solymos, P.; Stevens, M.H.H.; Wagner, H. vegan: Community Ecology Package. CRAN R package. 2015.
Oren, A., Garrity, G.M., 2021. Valid publication of the names of forty-two phyla of prokaryotes. Int J Syst Evol Microbiol 71.
Peters, D.L., Mccutcheon, J.G., Stothard, P., Dennis, J.J., 2019. Novel Stenotrophomonas maltophilia temperate phage DLP4 is capable of lysogenic conversion. BMC Genomics 20.
Qian, X., Gunturu, S., Guo, J., Chai, B., Cole, J.R., Gu, J., Tiedje, J.M., 2021. Metagenomic analysis reveals the shared and distinct features of the soil resistome across tundra, temperate prairie, and tropical ecosystems. Microbiome 9, 108.
Robinson, J.T.; Thorvaldsdóttir, H.; Turner, D.; Mesirov, J.P. igv.js: an embeddable JavaScript implementation of the Integrative Genomics Viewer (IGV). bioRxiv 2020: 2020.2005.2003.075499.
Sun, R.; Yu, P.; Zuo, P.; Alvarez, P. Bacterial Concentrations and Water Turbulence Influence the Importance of Conjugation Versus Phage-Mediated Antibiotic Resistance Gene Transfer in Suspended Growth Systems. ACS Environmental Au 2021;XXXX.
van Hoek, A.; Mevius, D.; Guerra, B.; Mullany, P.; Roberts, A.; Aarts, H. Acquired Antibiotic Resistance Genes: An Overview. Frontiers in Microbiology 2011;2.
Wang, R., van Dorp, L., Shaw, L.P., Bradley, P., Wang, Q., Wang, X., Jin, L., Zhang, Q., Liu, Y., Rieux, A., Dorai-Schneiders, T., Weinert, L.A., Iqbal, Z., Didelot, X., Wang, H., Balloux, F., 2018. The global distribution and spread of the mobilized colistin resistance gene mcr-1. Nat Commun 9, 1179.
Wu, H.J., Wang, H.J., Jennings, M.P., 2008. Discovery of virulence factors of pathogenic bacteria. Current Opinion in Chemical Biology 12, 93–101.
Xiang, Q., Dong, Z., Giles, M., Neilson, R., Yang, X.-R., Qiao, M., Chen, Q., 2019. Agricultural activities affect the pattern of the resistome within the phyllosphere microbiome in peri-urban environments. J Hazard Materi 382, 121068.
Xu, H., Chen, Z., Huang, R., Cui, Y., Li, Q., Zhao, Y., Wang, X., Mao, D., Luo, Y., Ren, H., 2021. Antibiotic Resistance Gene-Carrying Plasmid Spreads into the Plant Endophytic Bacteria using Soil Bacteria as Carriers. Environ Sci Technol 55.
Yan, Z.Z., Chen, Q.L., Zhang, Y.J., He, J.Z., Hu, H.W., 2019. Antibiotic resistance in urban green spaces mirrors the pattern of industrial distribution - ScienceDirect. Environment International 132.
Yang, X., Chan, W.C., Zhang, R., Chen, S., 2019. A conjugative plasmid that augments virulence in Klebsiella pneumoniae. Nat Microbiol 4, 1–5.
Yang, X., Dong, N., Chan, W.C., Zhang, R., Chen, S., 2020. Carbapenem Resistance-Encoding and Virulence-Encoding Conjugative Plasmids in Klebsiella pneumoniae. Trends Microbiol 29.
Yang, Y., Jiang, X., Chai, B., Ma, L., Zhang, T., 2016. ARGs-OAP: Online analysis pipeline for antibiotic resistance genes detection from metagenomic data using an integrated structured ARG-database. Bioinformatics 32.
Yin, X., Xiao-Tao, J., Chai, B., Li, L., Yang, Y., Cole, J.R., Tiedje, J.M., Zhang, T., 2018. 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 13.
Zhang, A.N., Gaston, J.M., Dai, C.L., Zhao, S., Poyet, M., Groussin, M., Yin, X., Li, L.G., van Loosdrecht, M.C.M., Topp, E., Gillings, M.R., Hanage, W.P., Tiedje, J.M., Moniz, K., Alm, E.J., Zhang, T., 2021. An omics-based framework for assessing the health risk of antimicrobial resistance genes. Nat Commun 12, 4765.
Zhang, Z., Zhang, Q., Wang, T., Xu, N., Lu, T., Hong, W., Penuelas, J., Gillings, M., Wang, M., Gao, W., Qian, H., 2022. Assessment of global health risk of antibiotic resistance genes. Nat Commun 13, 1553.
Zhu, D., Delgado-Baquerizo, M., Su, J.-Q., Ding, J., Li, H., Gillings, M.R., Penuelas, J., Zhu, Y.-G., 2021. Deciphering Potential Roles of Earthworms in Mitigation of Antibiotic Resistance in the Soils from Diverse Ecosystems. Environ Sci Technol 55, 7445–7455.
Zhu, Y.G., Penuelas, J., 2020. Changes in the environmental microbiome in the Anthropocene. Glob Chang Biol.