Check for updates

Coexistence of Antibiotic Resistance Genes and Virulence Factors Deciphered by Large-Scale Complete Genome Analysis

Yu Pan,[a,b] Jiaxiong Zeng,[a,b] Liguan Li,[d] Jintao Yang,[c] Ziyun Tang,[c] Wenguang Xiong,[a,c] Yafei Li,[e] 🅾 Sheng Chen,[f] Zhenling Zeng[a,b,c]

[a]Guangdong Provincial Key Laboratory of Veterinary Drugs Development and Safety Evaluation, South China Agricultural University, Guangzhou, China
[b]National Risk Assessment Laboratory for Antimicrobial Resistance of Animal Original Bacteria, South China Agricultural University, Guangzhou, China
[c]Guangdong Laboratory for Lingnan Modern Agriculture, Guangzhou, China
[d]Department of Environmental Engineering, Technical University of Denmark, Kongens Lyngby, Denmark
[e]Public Monitoring Center for Agro-product of Guangdong Academy of Agricultural Sciences, Guangzhou, China
[f]Department of Infectious Diseases and Public Health, Jockey Club College of Veterinary Medicine and Life Sciences, City University of Hong Kong, Kowloon, Hong Kong, China

Yu Pan and Jiaxiong Zeng contributed equally to this article. Author order was determined by votes of all the authors.

ABSTRACT Widespread use of antibiotics has enhanced the evolution of highly re-silient pathogens and poses a severe risk to human health via coselection of antibi-otic resistance genes (ARGs) and virulence factors (VFs). In this study, we rigorously evaluate the abundance relationship and physical linkage between ARGs and VFs by performing a comprehensive analysis of 9,070 bacterial genomes isolated from multiple species and hosts. The coexistence of ARGs and VFs was observed in bacteria across distinct phyla, pathogenicities, and habitats, especially among human-associated pathogens. The coexistence patterns of gene elements in different habitats and pathogenicity groups were similar, presumably due to frequent gene transfer. A shorter intergenic distance between mobile genetic elements and ARGs/VFs was detected in human/animal-associated bacteria, indicating a higher transfer potential. Increased accumulation of exogenous ARGs/VFs in human pathogens highlights the importance of gene acquisition in the evolution of human commensal bacteria. Overall, the findings provide insights into the genic features of combina-tions of ARG-VF and expand our understanding of ARG-VF coexistence in bacteria.

IMPORTANCE Antibiotic resistance has become a serious global health concern. De-spite numerous case studies, a comprehensive analysis of ARG and VF coexistence in bacteria is lacking. In this study, we explore the coexistence profiles of ARGs and VFs in diverse categories of bacteria by using a high-resolution bioinformatics approach. We also provide compelling evidence of unique ARG-VF gene pairs coexisting in specific bacterial genomes and reveal the potential risk associated with the coexis-tence of ARGs and VFs in organisms in both clinical settings and environments.

KEYWORDS antibiotic resistance, virulence factors, gene coexistence, bacterial genome, pathogens, natural environment, public health risk

Antibiotic resistance is widely recognized as one of the major crises of public health. It has led to decreased efficacy of antimicrobials in clinical practice and even a lack of effective drugs in treating infections caused by multidrug-resistant bacteria. In Europe, nearly 33,000 people die annually because of antibiotic resistance (1). More worryingly, the emergence of antibiotic-resistant bacteria that contain virulence factors (VFs; which enable a microorganism to establish itself on or within a host and enhance its potential to cause disease, include bacterial toxins, cell surface proteins that mediate bacterial attachment, cell surface carbohydrates and proteins that protect a bacterium,

and hydrolytic enzymes that may contribute to the pathogenicity of the bacterium) has inevitably facilitated clinical outbreaks, causing life-threatening infections and serious public health threats (2, 3).

The virulence mechanisms in bacteria have been modified over millions of years to enhance their adaptation in host defense systems. However, the global emergence and development of antimicrobial resistance has only accelerated over the last 50 years, i.e., since antibiotics were first used. Although antibiotic resistance and bacterial virulence have developed in different timescales, there is likely an interplay (coselection) between antibiotic-resistant genes (ARGs) and VFs under selective pressure (4). In our experience, increased resistance has always been associated with decreased virulence and fitness in that interplay, either directly or indirectly (5). However, this may not be true in all cases. There is increasing evidence indicating that the relationship between resistance and virulence is adjusted in such a way that it confers more benefit to the pathogen (6, 7). Virulence factors are essential for bacteria to overcome host defense systems, and the acquirement of antibiotic resistance contributes to overcome antimicrobial therapies and to adapt to and colonize demanding environments (8). The associations between ARGs and VFs might follow the Darwinian model, in which those characteristics (e.g., antibiotic resistance plus virulence) that confer a specific advantage will be selected and fixed. Simultaneously, gene transfer events and a large genetic library can be utilized by bacteria to compensate for or overcome fitness costs, resulting in the constant emergence of resistant and virulent clones.

Widespread occurrence of ARGs and VFs in single clinical lineages or complexes of multiple pathogenic species has frequently been reported in recent decades, such as in *Staphylococcus aureus* (9), *Klebsiella pneumoniae* (3), *Pseudomonas aeruginosa* (10), *Enterococcus* spp. (11, 12), and *Escherichia coli* ST131 (13). ST11 carbapenem-resistant hypervirulent *Klebsiella pneumoniae* strains cause severe pneumonia with poor response to antibiotic treatment. Dissemination of such pathogenic strains is associated with a high mortality rate in Asia (3). In particular, methicillin-resistant *Staphylococcus aureus* with multidrug-resistant genes and virulence mechanisms continues to be associated with significant morbidity and mortality (>26%) in community infections (2). The coselection between ARGs and VFs has also been reported. Significant coexistence of ARGs and VFs was detected in *Escherichia coli* recovered from wastewater treatment processes (14). The coselection of ARGs and VFs located in mobile gene elements (MGEs) was also detected in *Escherichia coli* (15, 16). However, to our knowledge, previous studies on ARG-VF coexistence have lacked comprehensive data compilation that enabled us to detect common patterns of bacteria in gene selection and evolution processes. Hence, it is essential to systematically investigate and comprehensively understand the ARG-VF coexistence profiles among bacteria from different species and habitats.

In this study, to uncover coexistence profiles of ARGs and VFs in diverse bacterial categories, we used bioinformatic approaches to analyze the large-scale bacterial complete genomes available in the NCBI repository. We investigated the genetic relationships between ARGs and VFs across genome categories (phylogenies, pathogenicities, and habitats) from different perspectives and obtained the abundance profiles of broad-spectrum ARG/VF types and high-resolution evaluation of their genetic relationships. Furthermore, the underlying coexistence structures in different ecologies and the cotransfer potential of ARGs and VFs were also assessed. Comprehensive analysis of the coexistence profiles of ARGs and VFs in a large number of bacterial genomes would not only improve our understanding of the impact of genetic events in shaping the formation of the ARG and VF pool but also provide compelling evidence for the coexistence of unique ARG-VF combinations in specific bacterial genomes.

RESULTS AND DISCUSSION

Abundance and enrichment profiles of ARGs and VFs. We randomly collected 9,070 bacterial complete genomes that covered a large diversity of phylogenies (see

Fig. S1 in the supplemental material). The ARGs and VFs were identified in 76% (6,875 genomes in 24 phyla) and 50% (4,563 genomes in 8 phyla) of the 9,070 bacterial complete genomes, respectively. In total, 346 genomes (ca. 3.8%) had high abundances (>50 hits per genome) of both ARGs and VFs, including genomes of *Escherichia*, *Salmonella*, *Pseudomonas*, and *Shigella* strains. The propagation of ARGs in enteric microorganisms was prominent. All of the top 100 genomes that carried abundant ARGs belonged to *Enterobacteriaceae*, which were deemed as a potential ARG repository. The enrichment of ARGs in *Enterobacteria* was consistent with previous studies, which reported the continuous occurrence of typical ARGs in gastrointestinal microbial communities subjected to persistent antibiotic selection pressure (17–19). In contrast, VFs were more abundant in *Legionellaceae*, accounting for nearly 80% of the top 100 VF abundance genomes.

To ascertain whether there was a significant correlation between the abundances of ARGs and VFs in specific phylogenies, the average ARG/VF abundances were counted across different phylogenies. Nearly half (23 genera) of the top 50 genera for ARG and VF abundance in data set were shared. Furthermore, *Enterobacteriaceae* occupied over 25% (the largest proportion) of the 23 intersecting genera. However, no significant linearity ($R^2 < 0.5$) was detected between ARG and VF abundance in *Enterobacteriaceae* (Fig. S2), indicating that there was no obvious joint enrichment between ARGs and VFs in *Enterobacteriaceae*, although the coexistence phenomenon was apparent in some strains of *Enterobacteriaceae*. *Enterobacteriaceae*, which usually act as commensals in various hosts, will be hampered by a short survival time in the host due to lethality caused by the excessive coaccumulation of resistance and virulence genes. Notably, the *mcr*, *bla*$_{NDM}$, and *tet*(X) genes, which encode resistance to last-resort antibiotics (colistin, carbapenem, and tigecycline), could be detected in *Enterobacteriaceae*, as could VFs. Just as the frequent outbreaks of pathogenic and resistant *Klebsiella pneumoniae* (3), *Escherichia coli* (20), and *Salmonella enterica* (21) have been devastating, *Enterobacteriaceae*, which may be elicited by selective pressures, pose a large threat to public health and may cause antibiotic failure during clinical infections. This is contributed by the high abundance of ARGs and VFs and resistance to last-resort antibiotics. The coexistence of a high abundance of ARGs and VFs in *Enterobacteriaceae* indicates that modern anthropogenic activities could drive and shape bacterial genic selection, although virulence and resistance have evolved over very different timescales (22). While large-scale complete genome analyses could demonstrate a comprehensive profile of ARGs and VFs at the forefront of our current knowledge, possible biases may exist in the data due to the uneven sources in public databases and the lack of accessible biological information of largely unculturable environmental microbes (Fig. S3).

In the broad-scale gene detection, an uneven abundance in the distribution of ARG and VF types was revealed (Fig. S4). The top five dominant ARG types (excluding multidrug) were beta-lactam (7.24%), aminoglycoside (6.24%), bacitracin (6.12%), macrolide-lincosamide-streptogramin (MLS) (5.92%), and polymyxin (5.72%). Among the 13 VF types we retrieved, secretion system, adherence, metal uptake, and toxins accounted for more than 80% of the total VF numbers. Antibiotic consumption has been recognized as the main cause of ARG emergence. The ARGs for commonly consumed antimicrobials were more likely to be highly abundant in this study, which was also consistent with previous abundance-based surveys (23, 24).

Next, ARG and VF enrichment in different genomes was evaluated (Fig. S5). Among the four habitats under investigation, most of the ARG types were more likely to be enriched in the human, animal, and soil habitats ($P < 0.01$ and odds ratio [OR] > 1 [Fisher exact test with Benjamini-Hochberg correction]). Across these four major phyla, enrichment was observed in 24 ARG types. We found a higher diversity of ARGs enriched in pathogens than nonpathogens, which was consistent with the results of a previous analysis (25). Among VFs, all 13 types were enriched in different phyla, and more tended to be enriched in the human, animal and soil habitats. The enrichment of "advantageous" genes may protect microorganisms from the detrimental conditions
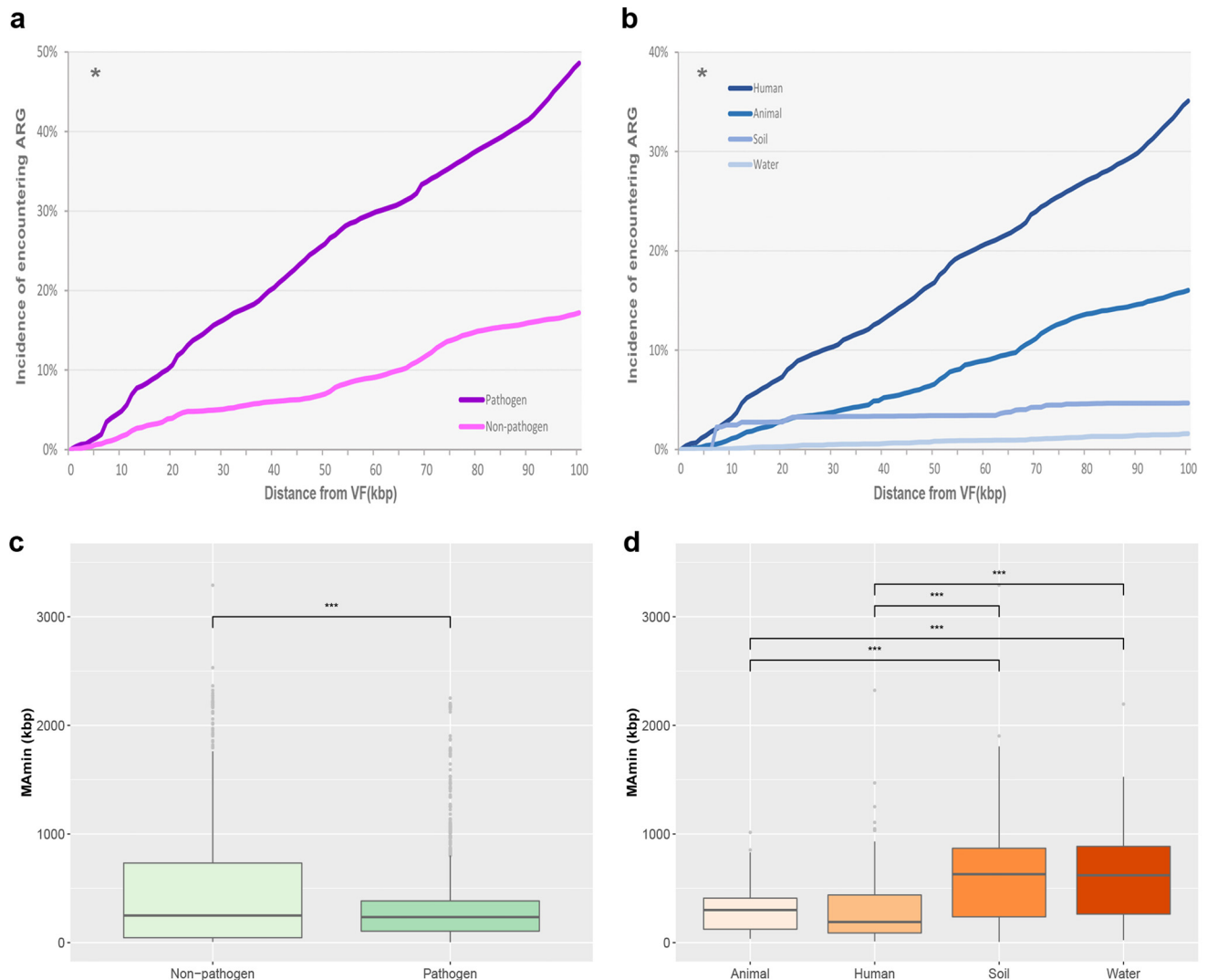
FIG 1 ARG and VF coexistence by pathogenicity (a and b) and habitat (c and d). (a and c) Incidence of encountering ARGs along the distance from VFs (*, $P < 0.05$ [determined by Fisher exact test]); (b and d) MetAmin index (***, $P < 0.001$ [determined by Student t test]).

brought by environment and host (26). To summarize, these results indicate a situation of broad-spectrum antimicrobial resistance and virulence disseminating in human pathogens and animal-associated bacteria.

Coexistence of ARGs and VFs. The coexistence profiles of ARGs and VFs were compared among organisms that belonged to different pathogenicity groups (containing pathogens and nonpathogens; the pathogen group consisted of all the confirmed and potential pathogens) and that resided in different habitats (Fig. 1). The coexistence trend was more significantly confirmed in pathogens than nonpathogens. The encounter incidence of ARGs in pathogen genomes (49% at 100 kbp) was always higher than that in nonpathogen genomes (17% at 100 kbp) as the distance from the VFs increased, indicating that VFs have a close connection with the occurrence of ARGs in pathogens ($P < 0.05$, Fisher exact test). In the MetAmin profile, a significantly closer distance between ARGs and VFs was observed in genomes with potential pathogenicity ($P < 0.001$, Student t test), which further strengthened the previous argument. Among the habitat groups, human-associated bacteria had the highest encounter incidence of ARGs, while animal-associated bacteria had the smallest MetAmin. When antibiotics began to be widely used in clinical treatment after the 1940s, the evolution of pathogens was greatly affected by the indiscriminate use of antibiotics. The develop-

ment of ARGs and VFs in pathogens overlapped considerably, and it was not possible to consider them as separated processes (27). In recent years, bacteria that were resistant to many antimicrobials and carried virulence factors, such as *Enterococcus faecium* (CC17), *Streptococcus pneumoniae* (PMEN1) and *E. coli* ST131 (28–30), had spread globally. These clones were considered highly successful or high-risk strains and could be detected in both humans and animals (31, 32).

To ascertain which pathogen species contributed to the above coexistence phenomenon, a profile of the 31 pathogen species with high coexistence potential was assessed (Fig. S6). Of the 31 high-coexistence-potential species, 14 exhibited a higher incidence of encountering ARGs, along with an increase in the distance from VFs, particularly in *Salmonella bongori*, *Citrobacter freundii*, *Fusobacterium nucleatum*, *Pseudomonas aeruginosa*, and *Escherichia coli*. Moreover, the MetAmin profile revealed substantial coexistence potential among the above 14 pathogen species, since their average MetAmin value (184 kbp) was much shorter than that of nonpathogen genomes (463 kbp). Upon examining the coexistence profiles in different habitats, the coexistence signature was more apparent in genomes derived from human and animal habitats, but it was much less frequent in organisms in soil and water habitats. This observation might be related to a phenomenon whereby pathogens in human and animal habitats are regularly exposed to antimicrobial agents and experience strong selection pressure (33).

Diversity of nearest ARG from VFs. Although the correlation of ARGs and VFs has been sporadically reported (19, 27, 34–36), there is still a lack of systematic knowledge of the general diversity pattern in bacterial genomes. In this study, the coexistence diversity of ARGs was found to be different between genomes of organisms in diverse pathogenicity groups and habitats, but there was no clear difference between animal and soil habitats (Fig. 2a and c). Furthermore, a higher Shannon diversity of ARG was detected in bacteria in the human habitat and pathogen category, suggesting that VFs cooccur with a large-spectrum of ARGs in human-associated pathogens. During the process of bacterial evolution, conserved genes are not uniformly distributed but are organized into clusters. It was suggested in previous studies that gene couples have a close functional relationship if the gene clusters exhibit the same coevolution pattern across different genomes (37). To identify ARG-VF couples that tend to be functionally related, the closest ARGs for all VFs were counted by type (Fig. 2b). The top five abundant ARG types (excluding multidrug), which are most likely to be functionally related to VFs and potentially confer resistance to major classes of antibiotics, encompass beta-lactam, bacitracin, MLS, fosfomycin, and aminoglycoside. In addition, different VF types were commonly present with distinct ARG types as their nearest genetic neighbor (Fig. 2d). For example, enzymes were more likely to correlate with tetracycline; stress proteins with MLS and beta-lactam; and antiphagocytosis with bacitracin, polymyxin, and beta-lactam.

Clusters of coexistence structures in different genome categories. The composite structure of the ARG-VF pairs was evaluated using principal coordinate analysis (PCoA) to identify whether there was a significant bacterial genome clustering pattern among different genome categories (Fig. S7). Intriguingly, a nonsignificant clustering was detected ($P > 0.001$, analysis of similarities [ANOSIM]) in the genomes from different habitats and pathogenicities. Even if genomes in different ecologies had individual ARG and VF coexistence profiles, we should bear in mind that intense gene dissemination may promote the acquisition and spread of the ARGs and VFs. The nonsignificant clustering suggested that the spread of VF-associated ARGs in bacteria was frequent enough to blur the resistome boundaries in different pathogenicities and habitats. Antibiotic resistance existed before the discovery and invention of currently used antibiotics, and the majority of ARGs may have an environmental origin (38). The high taxonomic and genetic diversity in environments (e.g., soil and sewage) has led to a nonuniform abundance and diversity of ARGs profiles, which are only partially governed by selection pressure. Moreover, the specific bacterial taxa that thrive in
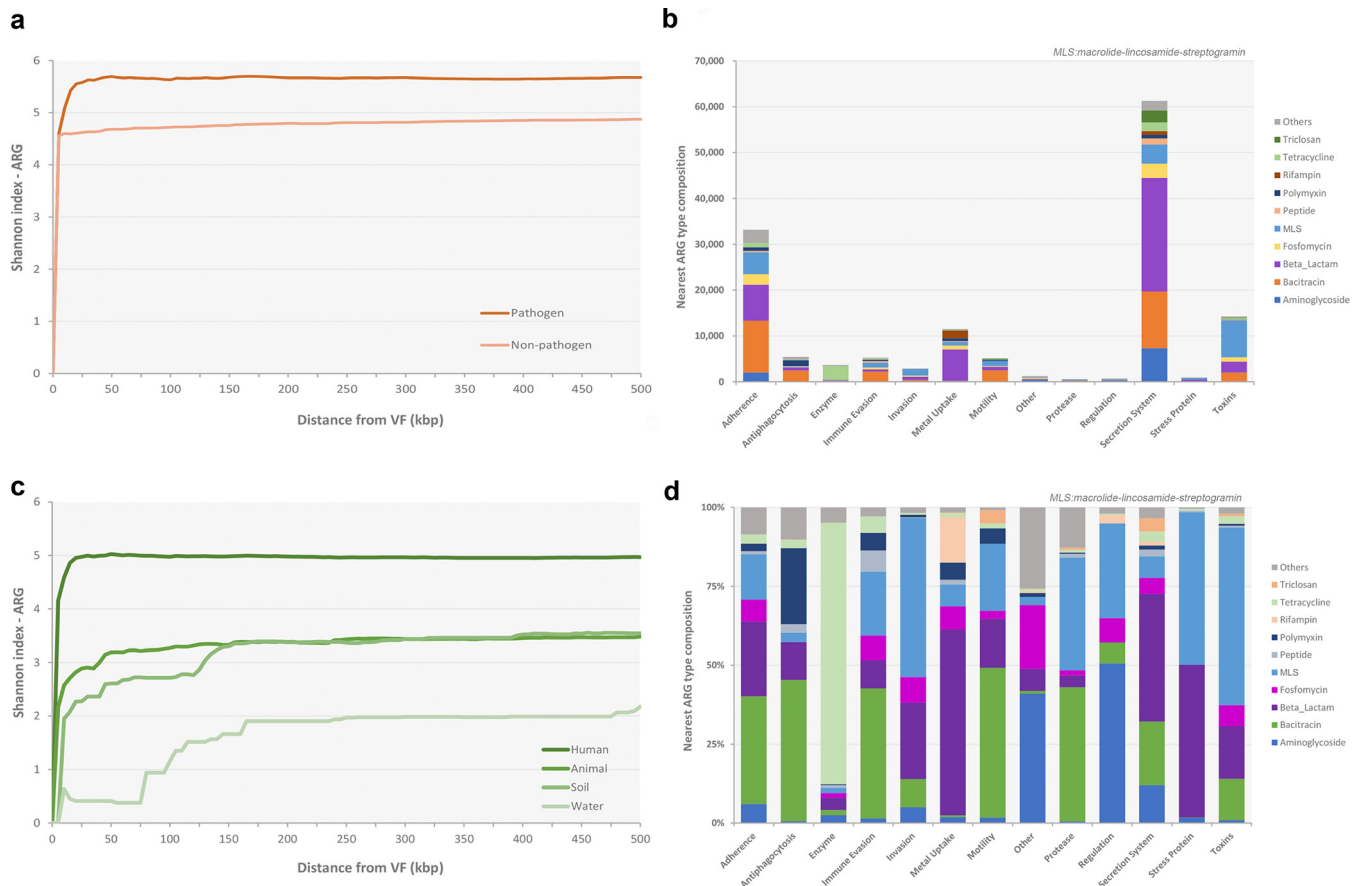
FIG 2 ARG diversity profile along the distance from VFs. Shannon diversity scores of the ARG subtypes along the distance from VFs by pathogenicity (a) and habitat (c), the abundance profile of the nearest ARG from all VF types (excluding multidrug) (b), and the composition of the nearest ARG from all VF types (excluding multidrug) (d) are shown.

several different habitats are likely to play an important role and might serve as mediators for ARGs in crossing ecological dissemination barriers (39–41). Another possible driving force to transmit ARGs across habitats is the increasingly prominent anthropogenic pressure, such as municipal wastewater systems, pharmaceutical manufacturing effluents, aquaculture facilities, and animal husbandry facilities (22). These sites are characterized by extremely high microbial loads coupled with subtherapeutic concentrations of antimicrobials, contributing to the discharge of human/animal fecal microbes and antibiotic-resistant organisms into the environment (42). Globally, the nonpathogenic species and commensals, regardless of their ecological niche, are now less susceptible to antibiotics than before (43). ARG dissemination between pathogens and nonpathogens was also observed. Due to horizontal gene transfer (HGT), pathogens under strong selective pressure may undergo positive selection of favorable genes (e.g., ARGs and VFs) originally belonging to nonpathogens to enhance their fitness (44). In addition, the core ARG-VF coupling types were searched across all the genomes. However, there was no single ARG-VF pair shared among all the genomes. There were five coupling types (multidrug with adherence, secretion system, metal uptake, bacitracin with adherence, and beta-lactam with metal uptake) identified in more than 25% of genomes, which mainly included couples of multidrug ARGs and VFs, highlighting the diversity of the ARG-VF alliance.

In the coexistence network analysis, a comprehensive perspective of the connection between ARG-VF pairs and important pathogen species was demonstrated (Fig. 3). Not surprisingly, the majority of these detected broad-spectrum ARG-VF pairs were shared between different pathogen species. The frequent spread of VF-associated ARGs, which
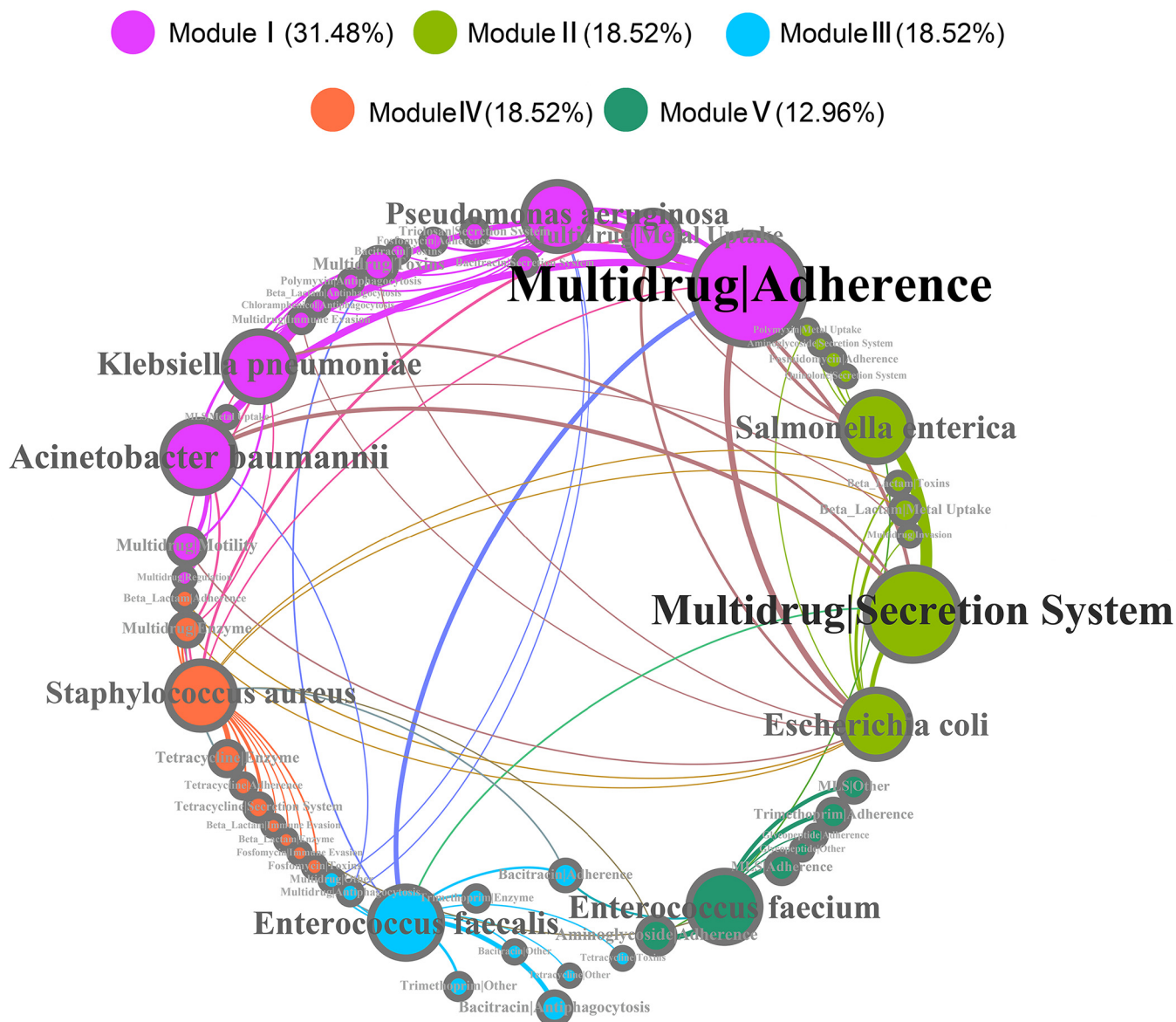
FIG 3 Network (circular layout) revealing the coexistence patterns between ARG-VF pairs and major pathogen species. The nodes are colored according to modularity class. The connections between each node represent that the relative abundance of ARG-VF pairs in that pathogen species exceeds 0.01. The size of each node is proportional to the weighted number of connections, that is, the average weighted degree.

might cross ecological dissemination barriers, was further confirmed. Based on the modular classification, there was a significant modular structure (i.e., clusters of nodes that interact among themselves more than other nodes in a random association) observed in the network (the modularity index was 0.465) (45). Every module had a distinct coexistence constitution and a more frequent exchange of ARG-VF pairs between species in the module. Module I, containing *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, and *Acinetobacter baumannii*, occupied a large number of vertices. These pathogen species contained a higher diversity of ARG-VF pairs and had a closer correlation in ARG-VF coexistence. Moreover, an apparent luxuriance of multidrug-adherence and multidrug-metal uptake gene type pairs was detected in module I. Module II, containing *Salmonella enterica* and *Escherichia coli*, was more likely to carry gene type pairs that contained multidrug and beta-lactam resistance genes, such as multidrug-secretion system and beta-lactam-metal uptake. Module III (containing *Enterococcus faecalis*), module IV (containing *Staphylococcus aureus*), and module V

(containing *Enterococcus faecium*) frequently carried the ARG-VF pairs containing bacitracin, multidrug and tetracycline, and MLS and aminoglycoside resistance genes, respectively. Pathogen species in the same module might reside in similar environmental niches and under common selection pressures. Positive selection in pathogenic species to different ARG-VF pairs could provide a functional support for their survival or development in specific niches.

Transfer potential of ARGs and VFs. The cotransfer ability of ARGs and VFs in different pathogenicities and habitats was evaluated based on the nearest distance from MGEs (Fig. 4). The distances of MGEs from the nearest ARGs and VFs were much shorter in the genomes from pathogen species and human and animal habitats ($P <$ 0.001, Student t test). The shorter distances suggested a possible role for MGEs in transferring antibiotic resistance and virulence in human pathogens and animal-associated bacteria (46–48). MGEs (mainly as plasmids and as integrative and conjugative elements) play a vital role in HGT and the coselection of ARGs and VFs in bacteria (30, 49, 50). ARGs and VFs can be disseminated extensively when equipped on the proper transfer machinery and efficient vehicles in gene shuffling, which could promote the evolution of microbe groups under environmental stress (51, 52). Sequence and functional genome analysis in previous studies revealed that many resistance and virulence genes in major pathogens have been obtained by horizontal transfer during evolution (53–55). This has caused serious outbreaks of superbugs in hospital and communal settings. Recently, an outbreak of carbapenem-resistant *Enterobacteriaceae* was attributed to the horizontal transfer of *Klebsiella pneumoniae* carbapenemase-encoding genes between strains and species by a conjugative plasmid (56). Lessons learned from transmissible ARGs and VFs in pathogens may lead to the development of antibiotic management strategies to control resistance and clinical outbreaks.

The cotransfer vehicles of all nearest ARG-VF couples were investigated for the strength of transfer preference for transposons, integrons, and plasmids (Fig. 4e). We found that a broad spectrum of ARG-VF couples could be cotransferred by the three genic vehicles. There could be a possible mechanism to help transposons and plasmids cotransfer certain gene couples. Transposons were more likely to transfer gene couples, such as bacitracin-toxins, mupirocin-antiphagocytosis, and aminoglycoside-adherence, while plasmids were much more likely to transfer gene couples of multidrug-secretion systems (57). No marked preference was observed in integrons containing a variety of functional gene cassettes. Of note, the ARG-VF couples transferred by high mobility potential MGEs, such as broad-host-range plasmids and class 1 integrons, could readily spread across species or even phylum boundaries (58).

To further evaluate the striking mobility of ARGs and VFs, the exogenously originated ARGs and VFs were identified across the large-scale genome collection (especially in human pathogens) (Fig. 5). Not surprisingly, compared to nonpathogens, the pairwise genomic comparisons showed more ARG/VF-carrying windows in human pathogens, in which the tetranucleotide frequency (TNF) signature substantially deviated from the genome-wide TNF pattern (correlation coefficient of <0.5). This indicated a higher incidence of exogenous acquisition of ARGs and VFs in human pathogens. The exogenous acquisition of VFs was also apparent in nonpathogens (Fig. 5c), which revealed the strong transfer ability of VFs. More interestingly, a clear bias in the distribution of exogenous fragments was detected in ARG-carrying windows (Fig. 5b). More than 70% of exogenous fragments were composed of genes conferring resistance against aminoglycoside, tetracycline, sulfonamide, and beta-lactam. This pattern corresponded with the mass consumption and long-period use of these common antibiotics in humans and animals (59, 60). Notably, a relatively higher incidence of exogenous fragments carrying glycopeptide resistance genes was observed in human pathogens. In view of the fact that glycopeptide (including vancomycin, teicoplanin, and telavancin) resistance genes are likely endemic components in nonpathogens, recent acquisition of these genes was regarded as being derived from environmental sources. Vancomycin, as the last line drug to defend against Gram-positive bacteria,
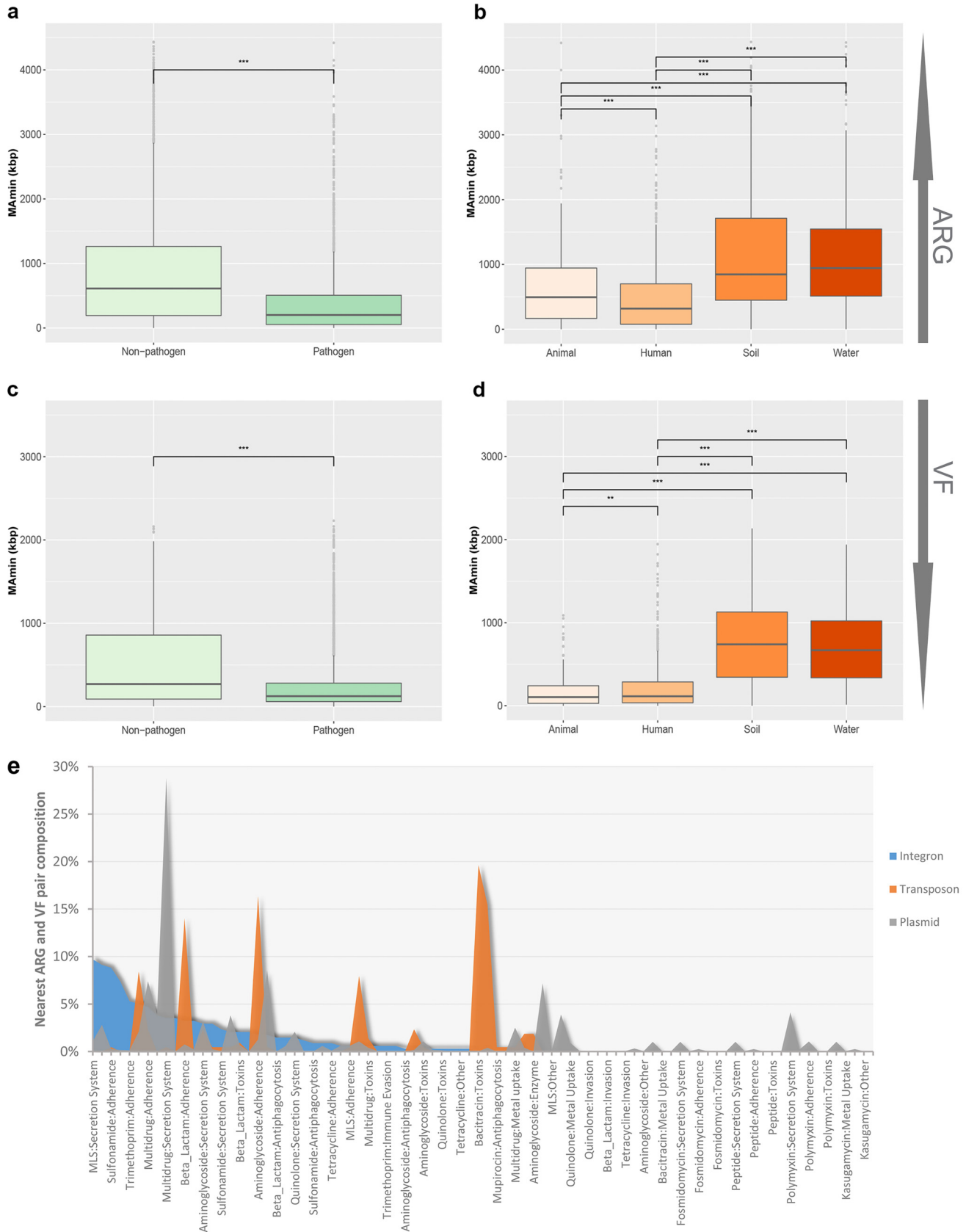
FIG 4 ARG and VF cotransfer potential. (a to d) The MetAmin index boxplot of the nearest ARG (a and b) and VF (c and d) from each MGE in each genome across the different genome categories (***, $P < 0.001$; **, $P < 0.01$ [determined by Student t test]); (e) cotransfer preference of the three MGE types (transposon, integron, and plasmid) for the nearest ARG and VF pair across the data set.
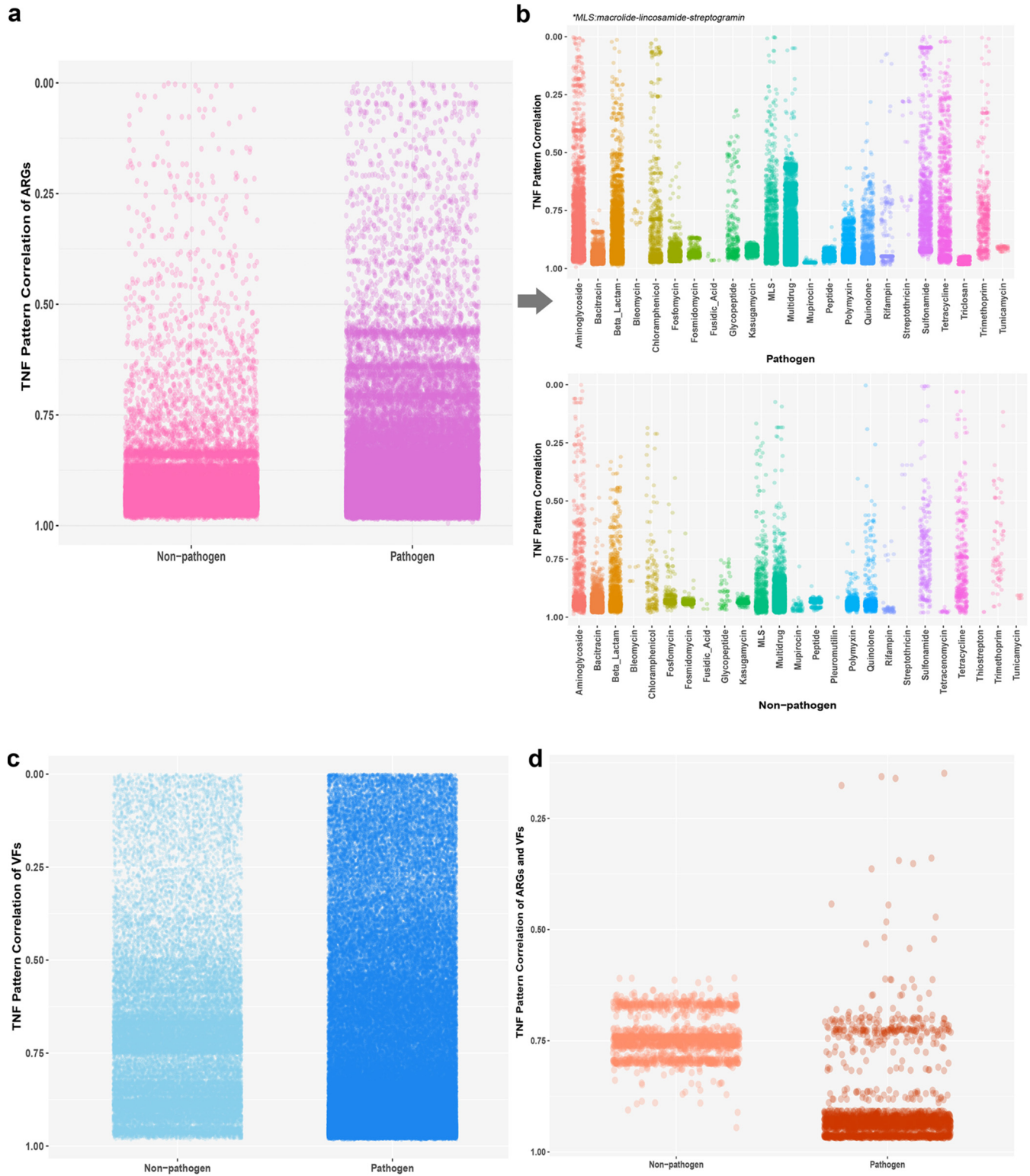
FIG 5 TNF pattern correlation between an ARG/VF-carrying fragment and its global genome sequence in pathogen and nonpathogen groups. (a and b) TNF pattern correlation profiles of fragments with ARGs; (c) TNF correlation profile of fragments with VFs; (d) TNF correlation profile of fragments with both ARGs and VFs.

such as *Staphylococcus aureus*, *Streptococcus pneumoniae*, and *Enterococcus*, has been prudently prescribed to treat infections caused by pan-drug-resistant pathogens over the past several decades (61, 62). Previous studies verified that a number of vancomycin resistance genes could be detected in swine manure samples (17), human feces

samples (63), and even permafrost sediments that were segregated from anthropogenic activities (64), indicating that human pathogens have rapidly adapted to newly introduced antibiotics under infectious treatment. The majority of exogenous ARG/VF-carrying fragments (>60%) were identified from clinically important pathogen species of *Acinetobacter baumannii*, *Escherichia coli*, *Klebsiella pneumoniae*, and *Salmonella enterica*, in which HGT plays an important role in the development of resistance (28, 65, 66).

Conclusions. In this study, we first bonded a large-scale complete genome data set with powerful genetic analysis to investigate the ARG-VF coexistence in the background of microbes from multiple species and hosts. Our results revealed the diversity of ARG-VF coexistence in genomes from distinct pathogenicities and habitats, in which ARGs and VFs were found to have a strong correlation and an intimate genetic linkage in pathogen species and human/animal habitats. ARGs and VFs with high cotransfer potential were speculated to blur ecological boundaries. Moreover, we revealed the significant role of HGT in shaping resistome and pathogenicities of human commensal bacteria. In summary, the relationship between ARG and VF evaluated in this study may depend on four main factors: the bacterial species involved, potential pathogenicity, the habitats and hosts, and genetic transmission. Although this allowed us to obtain a comprehensive view of ARG-VF coexistence, we appreciate that the data set is likely affected by anthropogenic sampling biases, which are overrepresented in samples from regions with active surveillance and research foundations. Nevertheless, we believe our study highlights the potential risk of ARG-VF coexistence at both clinical and environmental levels. We provide a road map and a bioinformatics method for future gene correlation researches. We also hope that the coexistence relationships revealed by this study will be further validated using other approaches.

MATERIALS AND METHODS

Complete genome collection and data source retrieval. To provide a comprehensive description of ARG-VF coexistence in bacteria, we randomly collected a large data set of bacterial complete genomes ($n = 9,070$, covering a large diversity of bacteria) from the NCBI RefSeq genome database (ftp://ftp.ncbi.nlm.nih.gov/genomes/refseq/bacteria) in March 2018 according to the proportions of different species in the database (considering the nonuniformity of data collection, the following analyses have eliminated the impact of sample size as much as possible). Taxonomic information of the downloaded data set was retrieved from the NCBI taxonomy database. The complete list of genomes is summarized in Table S1 in the supplemental material. The coding sequence (CDS) and their genetic location information were obtained to generate a downstream analysis. The habitat information was searched from the IMG database (https://img.jgi.doe.gov/) using the corresponding metadata (March 2018) by merging the fields of "habitat," "ecology," and "isolation" (Table S2). The potential pathogenicity of genome collections was obtained by mapping their taxonomic information with a published database covering all confirmed and potential species of human pathogens (67, 68).

ARG, VF, and MGE retrieval. The DeepARG database (merging the nonredundant data of the three major databases: CARD, ARDB, and UniProt) and SARG database (containing a hierarchical structure of type–subtype–reference sequence) were used to identify and quantify a broad-spectrum of ARGs (69, 70). The VFDB (containing cumulative information of VFs for the most important bacterial pathogens, virulence-associated genes, protein structural features, functions, mechanisms, and important literatures) and Victors database (including virulence factors for many different microbes that are pathogenic to animals and humans [http://www.phidias.us/victors/]) were used to detect VFs (71). All the CDS of downloaded genomes were extracted to identify ARGs and VFs by searching against the above database. The CDS were assigned an E value cutoff of 1e–5 against the database using BLASTP. CDS were annotated as ARGs and VFs when their best hit had ≥90% identity and ≥80% query coverage of the reference sequences (72). To explore the moving ability and heritable stability of DNA fragments, MGEs were identified by string matches to one of the following keywords in the gene description: transposon, integron, conjugal, mobilization, recombination, and plasmid.

Enrichment analysis. After we filtered the low-quality BLASTP results, deduplication of ARG- and VF-like sequences was performed and then categorized into antibiotic and VF types, respectively (VFs were classified into 13 types according to the function and mechanism description in the databases). The enrichment of ARGs and VFs across the phylogenies, habitats, and pathogenicities were tested by the Fisher exact test, and the P value was further adjusted by the Benjamini-Hochberg correction for multiple comparisons. Significant enrichment was defined as a P value of <0.01 and an odds ratio of >1.

ARG and VF coexistence analysis. Two indexes, including the average minimum distance and the incidence of encountering, were used to evaluate the coexistence relationship of ARGs and VFs on circular genomes (including plasmid sequences). The genomes carrying both ARGs and VFs were extracted for the coexistence analysis. Next, the average minimum distances [MetAmin(bp)] were

calculated on each genome as the sum of distances of the closest ARG from each VF divided by the number of VFs (equation 1 in Text S1 in the supplemental material). To calculate the incidence of gene encountering, we counted the number of ARGs within the assigned distance (1 to 100 kbp, by 1 kbp) from every VF in each genome, and the raw counts were then averaged in a given genome category (equation 2 in Text S1 in the supplemental material). The pathogen species, which were filtered by a standard of total ARG and VF numbers of ≥5 per genome and genome numbers of ≥5 per species, were considered to have high ARG and VF coexistence potential. To assess the diversity of coexistence patterns, the Shannon index was calculated by the number of unique ARG subtypes that were adjacent to VFs along a given distance (1 to 500 kbp) in each genome. The nearest ARG types of each VF type were summed up. Next, the MetAmin between ARGs/VFs and MGEs were calculated to estimate the gene cotransfer potential on the genome. The nearest ARG and VF from each MGE were deemed as an ARG-VF gene type pair (e.g., bacitracin-toxins) to analyze the cotransfer preference.

Cluster analysis. A count matrix of the nearest ARG-VF pairs in each ARG-VF type (e.g., bacitracin-toxins) was created for each genome to compare the underlying coexistence structures among the different genome categories. The count matrix was used to calculate the Bray-Curtis distance matrix, and a PCoA was then performed using R with package vegan. The coexistence structure in different categories of genome clusters was based on a P value of <0.001 calculated by using ANOSIM. The biplot position was calculated by the weighted average of the coordinate position of all genomes in the PCoA, where the weight was the abundance of the ARG-VF type in every genome and was plotted in a dimensional space using the ggplot2 package in R.

To visualize the coexistence structure in important clinical pathogens, eight major pathogen species (*Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Staphylococcus aureus*, *Pseudomonas aeruginosa*, *Escherichia coli*, *Salmonella enterica*, *Enterococcus faecalis*, and *Enterococcus faecium*) were analyzed through a network interface by counting their nearest ARG-VF types. A preliminary filter was implemented to remove those poorly represented ARG-VF pairs and reduce the artificial bias (those ARG-VF pairs were filtered, if the relative abundances were <0.01). The robust pairwise correlation of the ARG-VF pairs and the pathogen species mentioned above formed their coexistence network. The network and modularity classes were explored and visualized using the interactive platform of Gephi (73).

Tetranucleotide signature analysis. Tetranucleotide frequency (TNF) signatures contain a species-specific symbol (74) and have been shown to be a good compromise between computational calculation power and a pronounced phylogenetic signal (75). TNF signatures are widely used as an alignment-free genomic similarity index to depict the degree of sequence homology (76–78). To further evaluate the gene mobility and retrieve the exogenous ARGs and VFs, genome-wide TNF signatures were calculated (in forward and reverse strand) for each genome using the Biostrings package in R. By setting sliding window of 5 kbp (step = 1 kbp), TNF signatures were calculated across the whole length of each genome. Based on the TNF content comprising the whole genome and all the windows of each genome, the Pearson's correlation coefficient (r^2) was calculated between TNF counts [in terms of log(1+counts)] in each window and that of the corresponding whole genome to obtain the deviation of TNF in each window from the global pattern (78). Pearson correlation coefficients of the windows containing ARG were outlined based on the ARG type they carried.

Data availability. The data set supporting the conclusions of this article is included within the article and in the supplemental material.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

TEXT S1, DOC file, 0.03 MB.
FIG S1, TIF file, 2.6 MB.
FIG S2, TIF file, 0.9 MB.
FIG S3, TIF file, 2.5 MB.
FIG S4, TIF file, 0.3 MB.
FIG S5, TIF file, 0.4 MB.
FIG S6, TIF file, 1.4 MB.
FIG S7, TIF file, 2.5 MB.
TABLE S1, XLS file, 3.6 MB.
TABLE S2, XLSX file, 0.01 MB.

REFERENCES

1. Gasser M, Zingg W, Cassini A, Kronenberg A. 2019. Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in Switzerland. Lancet Infect Dis 19:17–18. https://doi.org/10.1016/S1473-3099(18)30708-4.

2. Kalil AC, Van Schooneveld TC, Fey PD, Rupp ME. 2014. Association between vancomycin minimum inhibitory concentration and mortality among patients with Staphylococcus aureus bloodstream infections: a systematic review and meta-analysis. JAMA 312:1552–1564. https://doi.org/10.1001/jama.2014.6364.

3. Gu D, Dong N, Zheng Z, Lin D, Huang M, Wang L, Chan E-C, Shu L, Yu J, Zhang R, Chen S. 2018. A fatal outbreak of ST11 carbapenem-resistant hypervirulent Klebsiella pneumoniae in a Chinese hospital: a molecular epidemiological study. Lancet Infect Dis 18:37–46. https://doi.org/10.1016/S1473-3099(17)30489-9.

4. Geisinger E, Isberg RR. 2017. Interplay between antibiotic resistance and virulence during disease promoted by multidrug-resistant bacteria. J Infect Dis 215(Suppl 1):S9–S17. https://doi.org/10.1093/infdis/jiw402.

5. Zhang Y, Liao K, Gao H, Wang Q, Wang XJ, Li HN, Wang RB, Wang H. 2017. Decreased fitness and virulence in ST10 Escherichia coli harboring bla_{NDM-5} and mcr-1 against a ST4981 strain with bla_{NDM-5}. Front Cell Infect Microbiol 7:242. https://doi.org/10.3389/fcimb.2017.00242.

6. Moya B, Juan C, Alberti S, Perez JL, Oliver A. 2008. Benefit of having multiple ampD genes for acquiring beta-lactam resistance without losing fitness and virulence in Pseudomonas aeruginosa. Antimicrob Agents Chemother 52:3694–3700. https://doi.org/10.1128/AAC.00172-08.

7. Renzoni A, Huggler E, Kelley WL, Lew D, Vaudaux P. 2009. Increased uptake and improved intracellular survival of a teicoplanin-resistant mutant of methicillin-resistant Staphylococcus aureus in non-professional phagocytes. Chemotherapy 55:183–188. https://doi.org/10.1159/000215304.

8. Livermore DM. 2003. Bacterial resistance: origins, epidemiology, and impact. Clin Infect Dis 36(Suppl 1):S11–S23. https://doi.org/10.1086/344654.

9. Aanensen DM, Feil EJ, Holden MT, Dordel J, Yeats CA, Fedosejev A, Goater R, Castillo-Ramirez S, Corander J, Colijn C, Chlebowicz MA, Schouls L, Heck M, Pluister G, Ruimy R, Kahlmeter G, Ahman J, Matuschek E, Friedrich AW, Parkhill J, Bentley SD, Spratt BG, Grundmann H, European SRL Working Group. 2016. Whole-genome sequencing for routine pathogen surveillance in public health: a population snapshot of invasive Staphylococcus aureus in Europe. mBio 7:e00444-16. https://doi.org/10.1128/mBio.00444-16.

10. Oliver A, Mulet X, López-Causapé C, Juan C. 2015. The increasing threat of Pseudomonas aeruginosa high-risk clones. Drug Resist Updat 21–22:41–59. https://doi.org/10.1016/j.drup.2015.08.002.

11. Raven KE, Reuter S, Gouliouris T, Reynolds R, Russell JE, Brown NM, Torok ME, Parkhill J, Peacock SJ. 2016. Genome-based characterization of hospital-adapted Enterococcus faecalis lineages. Nat Microbiol 1:15033. https://doi.org/10.1038/nmicrobiol.2015.33.

12. Raven KE, Reuter S, Reynolds R, Brodrick HJ, Russell JE, Torok ME, Parkhill J, Peacock SJ. 2016. A decade of genomic history for healthcare-associated Enterococcus faecium in the United Kingdom and Ireland. Genome Res 26:1388–1396. https://doi.org/10.1101/gr.204024.116.

13. Ben Zakour NL, Alsheikh-Hussain AS, Ashcroft MM, Nhu NTK, Roberts LW, Stanton-Cook M, Schembri MA, Beatson SA. 2016. Sequential acquisition of virulence and fluoroquinolone resistance has shaped the evolution of Escherichia coli ST131 (vol 7, e00347-16, 2016). mBio 7:e00958-16. https://doi.org/10.1128/mBio.00958-16.

14. Biswal BK, Mazza A, Masson L, Gehr R, Frigon D. 2014. Impact of wastewater treatment processes on antimicrobial resistance genes and their co-occurrence with virulence genes in Escherichia coli. Water Res 50:245–253. https://doi.org/10.1016/j.watres.2013.11.047.

15. Boerlin P, Travis R, Gyles CL, Reid-Smith R, Janecko N, Lim H, Nicholson V, McEwen SA, Friendship R, Archambault M. 2005. Antimicrobial resistance and virulence genes of Escherichia coli isolates from swine in Ontario. Appl Environ Microbiol 71:6753–6761. https://doi.org/10.1128/AEM.71.11.6753-6761.2005.

16. Travis RM, Gyles CL, Reid-Smith R, Poppe C, McEwen SA, Friendship R, Janecko N, Boerlin P. 2006. Chloramphenicol and kanamycin resistance among porcine Escherichia coli in Ontario. J Antimicrob Chemother 58:173–177. https://doi.org/10.1093/jac/dkl207.

17. Zhu YG, Johnson TA, Su JQ, Qiao M, Guo GX, Stedtfeld RD, Hashsham SA, Tiedje JM. 2013. Diverse and abundant antibiotic resistance genes in Chinese swine farms. Proc Natl Acad Sci U S A 110:3435–3440. https://doi.org/10.1073/pnas.1222743110.

18. Pal C, Bengtsson-Palme J, Kristiansson E, Larsson DG. 2016. The structure and diversity of human, animal and environmental resistomes. Microbiome 4:54. https://doi.org/10.1186/s40168-016-0199-5.

19. Escudeiro P, Pothier J, DiOniSio F, Nogueira T. 2019. Antibiotic resistance gene diversity and virulence gene diversity are correlated in human gut and environmental microbiomes. mSphere 4:e00135-19. https://doi.org/10.1128/mSphere.00135-19.

20. Madigan T, Johnson JR, Clabots C, Johnston BD, Porter SB, Slater BS, Banerjee R. 2015. Extensive household outbreak of urinary tract infection and intestinal colonization due to extended-spectrum beta-lactamase-producing Escherichia coli sequence type 131. Clin Infect Dis 61:E5–E12. https://doi.org/10.1093/cid/civ273.

21. Jackson BR, Griffin PM, Cole D, Walsh KA, Chai SJ. 2013. Outbreak-associated Salmonella enterica Serotypes and Food Commodities, United States, 1998–2008. Emerg Infect Dis 19:1239–1244. https://doi.org/10.3201/eid1908.121511.

22. Hendry AP, Gotanda KM, Svensson EI. 2017. Human influences on evolution, and the ecological and societal consequences. Philos Trans R Soc B Biol Sci 372:20160028. https://doi.org/10.1098/rstb.2016.0028.

23. Li B, Yang Y, Ma L, Ju F, Guo F, Tiedje JM, Zhang T. 2015. Metagenomic and network analysis reveal wide distribution and co-occurrence of environmental antibiotic resistance genes. ISME J 9:2490–2502. https://doi.org/10.1038/ismej.2015.59.

24. McDonnell L, Armstrong D, Ashworth M, Dregan A, Malik U, White P. 2017. National disparities in the relationship between antimicrobial resistance and antimicrobial consumption in Europe: an observational study in 29 countries. J Antimicrob Chemother 72:3199–3204. https://doi.org/10.1093/jac/dkx248.

25. Gibson MK, Forsberg KJ, Dantas G. 2015. Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. ISME J 9:207–216. https://doi.org/10.1038/ismej.2014.106.

26. Huijbers PM, Blaak H, de Jong MC, Graat EA, Vandenbroucke-Grauls CM, de Roda Husman AM. 2015. Role of the environment in the transmission of antimicrobial resistance to humans: a review. Environ Sci Technol 49:11993–12004. https://doi.org/10.1021/acs.est.5b02566.

27. Beceiro A, Tomas M, Bou G. 2013. Antimicrobial resistance and virulence: a successful or deleterious association in the bacterial world? Clin Microbiol Rev 26:185–230. https://doi.org/10.1128/CMR.00059-12.

28. Woodford N, Turton JF, Livermore DM. 2011. Multiresistant Gram-negative bacteria: the role of high-risk clones in the dissemination of antibiotic resistance. FEMS Microbiol Rev 35:736–755. https://doi.org/10.1111/j.1574-6976.2011.00268.x.

29. Petty NK, Ben Zakour NL, Stanton-Cook M, Skippington E, Totsika M, Forde BM, Phan MD, Moriel DG, Peters KM, Davies M, Rogers BA, Dougan G, Rodriguez-Bano J, Pascual A, Pitout JDD, Upton M, Paterson DL, Walsh TR, Schembri MA, Beatson SA. 2014. Global dissemination of a multidrug resistant Escherichia coli clone. Proc Natl Acad Sci U S A 111:5694–5699. https://doi.org/10.1073/pnas.1322678111.

30. Mathers AJ, Peirano G, Pitout J. 2015. The role of epidemic resistance plasmids and international high-risk clones in the spread of multidrug-resistant Enterobacteriaceae. Clin Microbiol Rev 28:565–591. https://doi.org/10.1128/CMR.00116-14.

31. Gomez-Simmonds A, Annavajhala MK, Wang Z, Macesic N, Hu Y, Giddins MJ, O'Malley A, Toussaint NC, Whittier S, Torres VJ, Uhlemann AC. 2018. Genomic and geographic context for the evolution of high-risk carbapenem-resistant Enterobacter cloacae complex clones ST171 and ST78. mBio 9:e00542-18. https://doi.org/10.1128/mBio.00542-18.

32. Haenni M, Bour M, Chatre P, Madec JY, Plesiat P, Jeannot K. 2017. Resistance of animal strains of Pseudomonas aeruginosa to carbapenems. Front Microbiol 8:1847. https://doi.org/10.3389/fmicb.2017.01847.

33. Andersson DI, Hughes D. 2014. Microbiological effects of sublethal levels of antibiotics. Nat Rev Microbiol 12:465–478. https://doi.org/10.1038/nrmicro3270.

34. Barkovskii AL, Babb CM, Hurley D, Shin E. 2015. Origins and environmental mobility of antibiotic resistance genes, virulence factors and bacteria in a tidal creek's watershed. J Appl Microbiol 118:764–776. https://doi.org/10.1111/jam.12735.

35. Dlamini BS, Montso PK, Kumar A, Ateba CN. 2018. Distribution of viru-

lence factors, determinants of antibiotic resistance and molecular fingerprinting of *Salmonella* species isolated from cattle and beef samples: suggestive evidence of animal-to-meat contamination. Environ Sci Pollut Res Int 25:32694–32708. https://doi.org/10.1007/s11356-018-3231-4.

36. Abana D, Gyamfi E, Dogbe M, Opoku G, Opare D, Boateng G, Mosi L. 2019. Investigating the virulence genes and antibiotic susceptibility patterns of *Vibrio cholerae* O1 in environmental and clinical isolates in Accra, Ghana. BMC Infect Dis 19. https://doi.org/10.1186/s12879-019-3714-z.

37. Donia MS, Cimermancic P, Schulze CJ, Wieland Brown LC, Martin J, Mitreva M, Clardy J, Linington RG, Fischbach MA. 2014. A systematic analysis of biosynthetic gene clusters in the human microbiome reveals a common family of antibiotics. Cell 158:1402–1414. https://doi.org/10.1016/j.cell.2014.08.032.

38. Forsberg KJ, Reyes A, Bin W, Selleck EM, Sommer MOA, Dantas G. 2012. The shared antibiotic resistome of soil bacteria and human pathogens. Science 337:1107–1111. https://doi.org/10.1126/science.1220761.

39. Bengtsson-Palme J, Larsson D. 2015. Antibiotic resistance genes in the environment: prioritizing risks. Nat Rev Microbiol 13:396–396. https://doi.org/10.1038/nrmicro3399-c1.

40. Pehrsson EC, Tsukayama P, Patel S, Mejia-Bautista M, Sosa-Soto G, Navarrete KM, Alderon MC, Abrera LC, Hoyos-Arango W, Bertoli MT, Berg DE, Gilman RH, Dantas G. 2016. Interconnected microbiomes and resistomes in low-income human habitats. Nature 533:212–216. https://doi.org/10.1038/nature17672.

41. Jiang XL, Ellabaan MMH, Charusanti P, Munck C, Blin K, Tong YJ, Weber T, Sommer MOA, Lee SY. 2017. Dissemination of antibiotic resistance genes from antibiotic producers to pathogens. Nat Commun 8:15784. https://doi.org/10.1038/ncomms15784.

42. Berendonk TU, Manaia CM, Merlin C, Fatta-Kassinos D, Cytryn E, Walsh F, Burgmann H, Sorum H, Norstrom M, Pons MN, Kreuzinger N, Huovinen P, Stefani S, Schwartz T, Kisand V, Baquero F, Martinez JL. 2015. Tackling antibiotic resistance: the environmental framework. Nat Rev Microbiol 13:310–317. https://doi.org/10.1038/nrmicro3439.

43. Martinez JL, Baquero F. 2002. Interactions among strategies associated with bacterial infection: pathogenicity, epidemicity, and antibiotic resistance. Clin Microbiol Rev 15:647–679. +. https://doi.org/10.1128/cmr.15.4.647-679.2002.

44. Perry JA, Wright GD. 2013. The antibiotic resistance "mobilome": searching for the link between environment and clinic. Front Microbiol 4:138. https://doi.org/10.3389/fmicb.2013.00138.

45. Newman M. 2006. Modularity and community structure in networks. Proc Natl Acad Sci U S A 103:8577–8582. https://doi.org/10.1073/pnas.0601602103.

46. Uhlemann AC, Otto M, Lowy FD, Deleo FR. 2014. Evolution of community- and healthcare-associated methicillin-resistant *Staphylococcus aureus*. Infect Genet Evol 21:563–574. https://doi.org/10.1016/j.meegid.2013.04.030.

47. Wright MS, Haft DH, Harkins DM, Perez F, Hujer KM, Bajaksouzian S, Benard MF, Jacobs MR, Bonomo RA, Adams MD. 2015. New insights into dissemination and variation of the health care-associated pathogen *Acinetobacter baumannii* from genomic analysis. mBio 6:e00963-13. https://doi.org/10.1128/mBio.00963-13.

48. Cloeckaert A, Zygmunt MS, Doublet B. 2017. Editorial: genetics of acquired antimicrobial resistance in animal and zoonotic pathogens. Front Microbiol 8:2428. https://doi.org/10.3389/fmicb.2017.02428.

49. Forsberg KJ, Patel S, Gibson MK, Lauber CL, Knight R, Fierer N, Dantas G. 2014. Bacterial phylogeny structures soil resistomes across habitats. Nature 509:612–616. https://doi.org/10.1038/nature13377.

50. Prieto AMG, van Schaik W, Rogers MRC, Coque TM, Baquero F, Corander J, Willems R. 2016. Global emergence and dissemination of enterococci as nosocomial pathogens: attack of the clones? Front Microbiol 7:788. https://doi.org/10.3389/fmicb.2016.00788.

51. Kaito C, Saito Y, Ikuo M, Omae Y, Mao H, Nagano G, Fujiyuki T, Numata S, Han X, Obata K, Hasegawa S, Yamaguchi H, Inokuchi K, Ito T, Hiramatsu K, Sekimizu K. 2013. Mobile genetic element SCCmec-encoded *psm-mec* RNA suppresses translation of *agrA* and attenuates MRSA virulence. PLoS Pathog 9:e1003269. https://doi.org/10.1371/journal.ppat.1003269.

52. Penades JR, Chen J, Quiles-Puchalt N, Carpena N, Novick RP. 2015. Bacteriophage-mediated spread of bacterial virulence genes. Curr Opin Microbiol 23:171–178. https://doi.org/10.1016/j.mib.2014.11.019.

53. Nasser W, Beres SB, Olsen RJ, Dean MA, Rice KA, Long W, Kristinsson KG, Gottfredsson M, Vuopio J, Raisanen K, Caugant DA, Steinbakk M, Low DE, McGeer A, Darenberg J, Henriques-Normark B, Van Beneden CA, Hoff-

mann S, Musser JM. 2014. Evolutionary pathway to increased virulence and epidemic group A *Streptococcus* disease derived from 3,615 genome sequences. Proc Natl Acad Sci U S A 111:E1768–E1776. https://doi.org/10.1073/pnas.1403138111.

54. Salipante SJ, SenGupta DJ, Cummings LA, Land TA, Hoogestraat DR, Cookson BT. 2015. Application of whole-genome sequencing for bacterial strain typing in molecular epidemiology. J Clin Microbiol 53:1072–1079. https://doi.org/10.1128/JCM.03385-14.

55. Liu Y-Y, Wang Y, Walsh TR, Yi L-X, Zhang R, Spencer J, Doi Y, Tian G, Dong B, Huang X, Yu L-F, Gu D, Ren H, Chen X, Lv L, He D, Zhou H, Liang Z, Liu J-H, Shen J. 2016. Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. Lancet Infect Dis 16:161–168. https://doi.org/10.1016/S1473-3099(15)00424-7.

56. Del Franco M, Paone L, Novati R, Giacomazzi CG, Bagattini M, Galotto C, Montanera PG, Triassi M, Zarrilli R. 2015. Molecular epidemiology of carbapenem-resistant *Enterobacteriaceae* in Valle d'Aosta region, Italy, shows the emergence of KPC-2-producing *Klebsiella pneumoniae* clonal complex 101 (ST101 and ST1789). BMC Microbiol 15:260. https://doi.org/10.1186/s12866-015-0597-z.

57. Nogueira T, Rankin DJ, Touchon M, Taddei F, Brown SP, Rocha E. 2009. Horizontal gene transfer of the secretome drives the evolution of bacterial cooperation and virulence. Curr Biol 19:1683–1691. https://doi.org/10.1016/j.cub.2009.08.056.

58. Klumper U, Riber L, Dechesne A, Sannazzarro A, Hansen LH, Sorensen SJ, Smets BF. 2015. Broad host range plasmids can invade an unexpectedly diverse fraction of a soil bacterial community. ISME J 9:934–945. https://doi.org/10.1038/ismej.2014.191.

59. Van Boeckel TP, Gandra S, Ashok A, Caudron Q, Grenfell BT, Levin SA, Laxminarayan R. 2014. Global antibiotic consumption 2000 to 2010: an analysis of national pharmaceutical sales data (vol 14, pg 742, 2014). Lancet Infect Dis 14:742–897. https://doi.org/10.1016/S1473-3099(14)70780-7.

60. ECDC/EFSA/EMA. 2017. ECDC/EFSA/EMA second joint report on the integrated analysis of the consumption of antimicrobial agents and occurrence of antimicrobial resistance in bacteria from humans and food-producing animals. EFSA J 15. http://www.ema.europa.eu/docs/en_GB/document_library/Report/2017/07/WC500232336.pdf.

61. MacDougall C, Polk RE. 2005. Antimicrobial stewardship programs in health care systems. Clin Microbiol Rev 18:638–656. https://doi.org/10.1128/CMR.18.4.638-656.2005.

62. McKenna M. 2013. Antibiotic resistance: the last resort. Nature 499:394–396. https://doi.org/10.1038/499394a.

63. Hu YF, Yang X, Qin JJ, Lu N, Cheng G, Wu N, Pan YL, Li J, Zhu LY, Wang X, Meng ZQ, Zhao FQ, Liu D, Ma JC, Qin N, Xiang CS, Xiao YH, Li LJ, Yang HM, Wang J, Yang RF, Gao GF, Wang J, Zhu B. 2013. Metagenome-wide analysis of antibiotic resistance genes in a large cohort of human gut microbiota. Nat Commun 4:2151. https://doi.org/10.1038/ncomms3151.

64. D'Costa VM, King CE, Kalan L, Morar M, Sung WWL, Schwarz C, Froese D, Zazula G, Calmels F, Debruyne R, Golding GB, Poinar HN, Wright GD. 2011. Antibiotic resistance is ancient. Nature 477:457–461. https://doi.org/10.1038/nature10388.

65. Chiu CH, Tang P, Chu CS, Hu SN, Bao QY, Yu J, Chou YY, Wang HS, Lee YS. 2005. The genome sequence of *Salmonella enterica* serovar Choleraesuis, a highly invasive and resistant zoonotic pathogen. Nucleic Acids Res 33:1690–1698. https://doi.org/10.1093/nar/gki297.

66. Cavaco LM, Hasman H, Xia S, Aarestrup FM. 2009. *qnrD*, a novel gene conferring transferable quinolone resistance in *Salmonella enterica* serovar Kentucky and Bovismorbificans strains of human origin. Antimicrob Agents Chemother 53:603–608. https://doi.org/10.1128/AAC.00997-08.

67. Woolhouse MEJ, Gowtage-Sequeria S. 2005. Host range and emerging and reemerging pathogens. Emerg Infect Dis 11:1842–1847. https://doi.org/10.3201/eid1112.050997.

68. Li LG, Xia Y, Zhang T. 2017. Co-occurrence of antibiotic and metal resistance genes revealed in complete genome collection. ISME J 11:651–662. https://doi.org/10.1038/ismej.2016.155.

69. Arango-Argoty G, Garner E, Pruden A, Heath LS, Vikesland P, Zhang L. 2018. DeepARG: a deep learning approach for predicting antibiotic resistance genes from metagenomic data. Microbiome 6:23. https://doi.org/10.1186/s40168-018-0401-z.

70. Yang Y, Jiang XT, Chai BL, Ma LP, Li B, Zhang AN, Cole JR, Tiedje JM, Zhang T. 2016. ARGs-OAP: online analysis pipeline for antibiotic resis-

tance genes detection from metagenomic data using an integrated structured ARG-database. Bioinformatics 32:2346–2351. https://doi.org/10.1093/bioinformatics/btw136.

71. Chen L, Yang J, Yu J, Yao Z, Sun L, Shen Y, Jin Q. 2005. VFDB: a reference database for bacterial virulence factors. Nucleic Acids Res 33:D325–D328. https://doi.org/10.1093/nar/gki008.

72. Yang Y, Li B, Ju F, Zhang T. 2013. Exploring variation of antibiotic resistance genes in activated sludge over a four-year period through a metagenomic approach. Environ Sci Technol 47:10197–10205. https://doi.org/10.1021/es4017365.

73. Bastian M, Heymann S, Jacomy M. 2009. Gephi: an open source software for exploring and manipulating networks. ICWSM 8:361–362.

74. van Passel MWJ, Kuramae EE, Luyf ACM, Bart A, Boekhout T. 2006. The reach of the genome signature in prokaryotes. BMC Evol Biol 6:84. https://doi.org/10.1186/1471-2148-6-84.

75. Pride DT, Meinersmann RJ, Wassenaar TM, Blaser MJ. 2003. Evolutionary implications of microbial genome tetranucleotide frequency biases. Genome Res 13:145–158. https://doi.org/10.1101/gr.335003.

76. Fondi M, Karkman A, Tamminen MV, Bosi E, Virta M, Fani R, Alm E, McInerney JO. 2016. "Every gene is everywhere but the environment selects": global geolocalization of gene sharing in environmental samples through network analysis. Genome Biol Evol 8:1388–1400. https://doi.org/10.1093/gbe/evw077.

77. Nongkhlaw FMW, Joshi SR. 2016. Horizontal gene transfer of the non-ribosomal peptide synthetase gene among endophytic and epiphytic bacteria associated with ethnomedicinal plants. Curr Microbiol 72:1–11. https://doi.org/10.1007/s00284-015-0910-y.

78. Daims H, Lebedeva EV, Pjevac P, Han P, Herbold C, Albertsen M, Jehmlich N, Palatinszky M, Vierheilig J, Bulaev A, Kirkegaard RH, von Bergen M, Rattei T, Bendinger B, Nielsen PH, Wagner M. 2015. Complete nitrification by Nitrospira bacteria. Nature 528:504–509. https://doi.org/10.1038/nature16461.