



Sharing of Antimicrobial Resistance Genes between Humans and Food Animals

Huiluo Cao,^a Salim Bougouffa,^b Tae-Jin Park,^c Andes Lau,^a Man-Ki Tong,^a Kin-Hung Chow,^a  Pak-Leung Ho^{a,d}

^aCarol Yu Center for Infection and Department of Microbiology, University of Hong Kong, Hong Kong, People's Republic of China

^bComputational Bioscience Research Center and Bioscience Core Lab, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

^cHME Healthcare Co., Ltd., Suwon-si, Gyeonggi-do, Republic of Korea

^dDepartment of Microbiology, Queen Mary Hospital, Hong Kong, People's Republic of China

ABSTRACT The prevalence and propagation of antimicrobial resistance (AMR) are serious global public health concerns. The large and the ever-increasing use of antibiotics in livestock is also considered a great concern. The extent of the similarity of acquired antibiotic resistance genes (ARGs) between humans and food animals and the driving factors underlying AMR transfer between them are not clear, although a link between ARGs in both hosts was proposed. To address this question, with swine and chicken as examples of food animals, we analyzed over 1,000 gut metagenomes of humans and food animals from over the world. A relatively high abundance and diversity of ARGs were observed in swine compared with those in humans as a whole. Commensal bacteria, particularly species from *Clostridiales*, contribute the most ARGs associated with mobile genetic elements (MGEs) and were found in both humans and food animals. Further studies demonstrate that overrepresented MGEs, namely, Tn4451/Tn4453 and TnAs3, are attributed mainly to the sharing between humans and food animals. A member of large resolvase family site-specific recombinases, TnpX, is found in Tn4451/Tn4453 which facilitates the insertions of the transient circular molecule. Although the variance in the transferability of ARGs in humans is higher than that in swine, a higher average transferability was observed in swine than that in humans. In conclusion, the potential antibiotic resistance hot spots with higher transferability in food animals observed in the present study emphasize the importance of surveillance for emerging resistance threats before they spread.

IMPORTANCE Antimicrobial resistance (AMR) has proven to be a global public health concern. To conquer this increasingly worrying trend, an overarching, One Health approach has been used that brings together different sectors, but the fundamental knowledge of the relationship between humans, food animals, and their environments is not mature yet or is lacking in some aspect. With swine and chicken as examples of food animals, a large global data set of over 1,000 human and food animal gut metagenomes was analyzed with a focus on acquired antibiotic resistance genes (ARGs) associated with mobile genetic elements (MGEs) to answer this question. Outputs from this work open a new avenue to further our understanding of ARG transferability in food animals. It is a necessary milestone to better equip governmental agencies to monitor and pre-empt antibiotic resistance hot spots. This work will assist and give guidance on how to decipher other links within any One Health initiatives with expected positive feedback to human health.

KEYWORDS antibiotic resistance gene, food animals, human gut, metagenome, mobile genetic element

As a serious global public health concern, antimicrobial resistance has been aggravated by the increasing incidence of multidrug resistance in clinical pathogens (1). Genes conferring antimicrobial resistance (AMR) in pathogens are syntenic with mobile

Editor Tricia A. Van Laar, California State University, Stanislaus

Copyright © 2022 Cao et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Pak-Leung Ho, plho@hku.hk.

The authors declare no conflict of interest.

Received 23 August 2022

Accepted 16 September 2022

Published 11 October 2022

genetic elements (MGEs), such as integrons, transposons, plasmids, and prophages, suggesting that they may be disseminated within a mobile multidrug resistance cassette from commensal bacteria to pathogens. Livestock is considered a major source of acquired antibiotic resistance genes (ARGs) (2), especially in low- and middle-income countries (3). Bacteria in animals that are treated with antibiotics can develop antibiotic resistance, and these bacteria, which might carry resistance genes, then can be transmitted from animals to humans and vice versa. This intertransmission of ARGs can occur through food, by direct contact between humans and animals, or through shared environmental resources, such as contaminated water (4). This information emphasizes the importance of the continued prioritization and the surveillance of the antibiotic resistome, particularly in hot spots where there is a high likelihood of resistance gene evolution and transfer between bacterial hosts.

Clinically significant ARGs, such as New Delhi metallo-beta-lactamase (*bla_{NDM}*) and plasmid-borne colistin-resistance genes (*mcr-1*, colistin that is the last resort antibiotic for human health) which are found in common pathogens, have been studied broadly and were observed simultaneously in food animals and humans, indicating potential links between these two types of hosts (5, 6). *mcr-1* was found initially through an active surveillance program in China that monitored ARGs in food animals (7), and a follow-up study on the global distribution posited a possible origin of *mcr-1* in Chinese livestock (8) indicating the potential zoonotic origin of *mcr-1* in humans. These studies suggest that food animals could be an important source of clinically significant ARGs in human populations. However, the extent that ARGs occur in commensal bacteria, which have been proven to transfer to pathogens and are shared by humans and food animals, is undetermined. Monitoring of the resistome in humans and food animals could provide more solid evidence and a comprehensive understanding of the migration of ARGs.

Indeed, the large and expanding use of antibiotics in livestock, particularly in China, is of a great concern in light of the threat of antibiotic resistance (3, 9). In Hong Kong, 80% of the food animals are imported from mainland China (10), ARGs in imported food animals from China are thought to have strong influence on ARGs in Hong Kong populations, and this influence has also been observed in other countries, such as European countries (3). In this work, we exploit swine and chicken as examples to gauge the sharing and the transferability of ARGs with humans on a global scale. The transferable capability of ARGs was evaluated through the corresponding mobile genetic cassette identified in both types of hosts and by the taxonomy of bacteria.

RESULTS

Diversity and relative abundance of ARGs. In present study, a total of 1,487 gut metagenomes of humans and food animals from globally geographical locations, including our recently sequenced 21 human gut microbiota samples, were collated and analyzed (see Fig. S1 in the supplemental material; Table 1; see Table S1 online at <https://doi.org/10.6084/m9.figshare.20514555.v1>). The numbers of observed ARGs in most human samples were lower than those from food animals, including swine and chicken (Fig. 1a), but the distributions of ARGs of sequencing reads in both types of hosts were comparable (see Fig. S2 in the supplemental material). Rarefaction curves indicated the nonsaturation of samplings for samples from both hosts (Fig. 1b; see Fig. S3 in the supplemental material), and therefore, additional sequencing may be needed in the future, especially from those low- and middle-income countries for which data are not available currently. All three beta diversity indices exhibit greater differences within human samples than within food animal samples (Fig. 1c to e).

The relative abundance of ARGs in swine is significantly higher than that in human hosts (Fig. 2a; see Table S2 online at <https://doi.org/10.6084/m9.figshare.20514555.v1>). Considering ARGs per host type and by sampling region, the highest abundance and the highest diversity (both Shannon and Simpson indices) of ARGs were observed in swine from China (pairwise Wilcoxon signed-rank test with Benjamini-Hochberg

TABLE 1 Numbers of metagenomes in each region and host in the present study

Region	No. of metagenomes by host		
	Human	Swine	Chicken
Hong Kong	74	0	0
China	85	97	135
France	12	120	20
Denmark	55	120	20
Netherlands	0	20	20
Germany	62	20	19
Spain	51	20	20
Belgium	0	20	20
Italy	52	20	20
Poland	0	20	20
Bulgaria	0	21	19
Japan	31	0	0
Austria	16	0	0
Sweden	39	0	0
Canada	18	0	0
Peru	15	0	0
Salvador	15	0	0
USA	61	0	0
India	110	0	0
Total	696	478	313

method to correct, $P < 0.05$) (see Fig. S4 in the supplemental material). Within the human cohorts, the Indian population, surprisingly, has the highest relative abundance of ARGs, as supported by the pairwise Wilcoxon signed-rank test (corrected with the Benjamini-Hochberg method, $P < 0.05$) (Fig. S4a). As for humans versus food animals, we detected a higher abundance of ARGs in food animals than that in humans in all countries except France and Italy (Fig. S4b).

We clustered all ARGs by class in both groups of hosts into a heatmap to show those ARG classes that are significantly abundant (Fig. 2b). ARGs conferring resistance to tetracyclines, vancomycin, and macrolides are observed widely in most samples, while those to quinolones, phenicols, and aminoglycosides show large differences in different hosts. For details, refer to Fig. 2c where we highlight ARGs that are statistically different in the relative abundance of ARGs within both types of hosts. All ARGs but *cepA*, which encodes a class A beta-lactamase, are more abundant in food animals than in humans (Fig. 2c). In particular, a few ARGs, for instance, tetracycline resistance genes, exhibit high abundance in both types of hosts. To obtain more details of shared ARGs between different hosts in each region, heatmaps of ARGs for each class of antibiotic were plotted (see Fig. S8 online at <https://doi.org/10.6084/m9.figshare.20522313.v2>). A relatively higher abundance of ARGs against most antibiotic classes were observed in Indian human and Chinese chicken samples, particularly for beta-lactams and ARGs conferring resistance to quinolones. Variations of the relative abundance of ARGs within the same class of antibiotic were observed. Samples from Chinese swine were enriched for ARGs conferring resistance to phenicols, aminoglycosides, and tetracyclines but lack ARGs related to quinolone (Fig. 2b; see Fig. S8 at online <https://doi.org/10.6084/m9.figshare.20522313.v2>). All these results do not show the possible sharing between two types of hosts. Thus, in the following section, the detailed shared ARGs between two types of hosts are described.

The principal coordinate analysis separated ARGs in food animals from those in human hosts for all but a few samples (see Fig. S5 in the supplemental material). The separation of ARGs by host sources is consistent across all regions. But the separation distance between host types varies with Spain and Denmark having the most prominent separation while France had the least prominent separation (Fig. 2d). The human gut

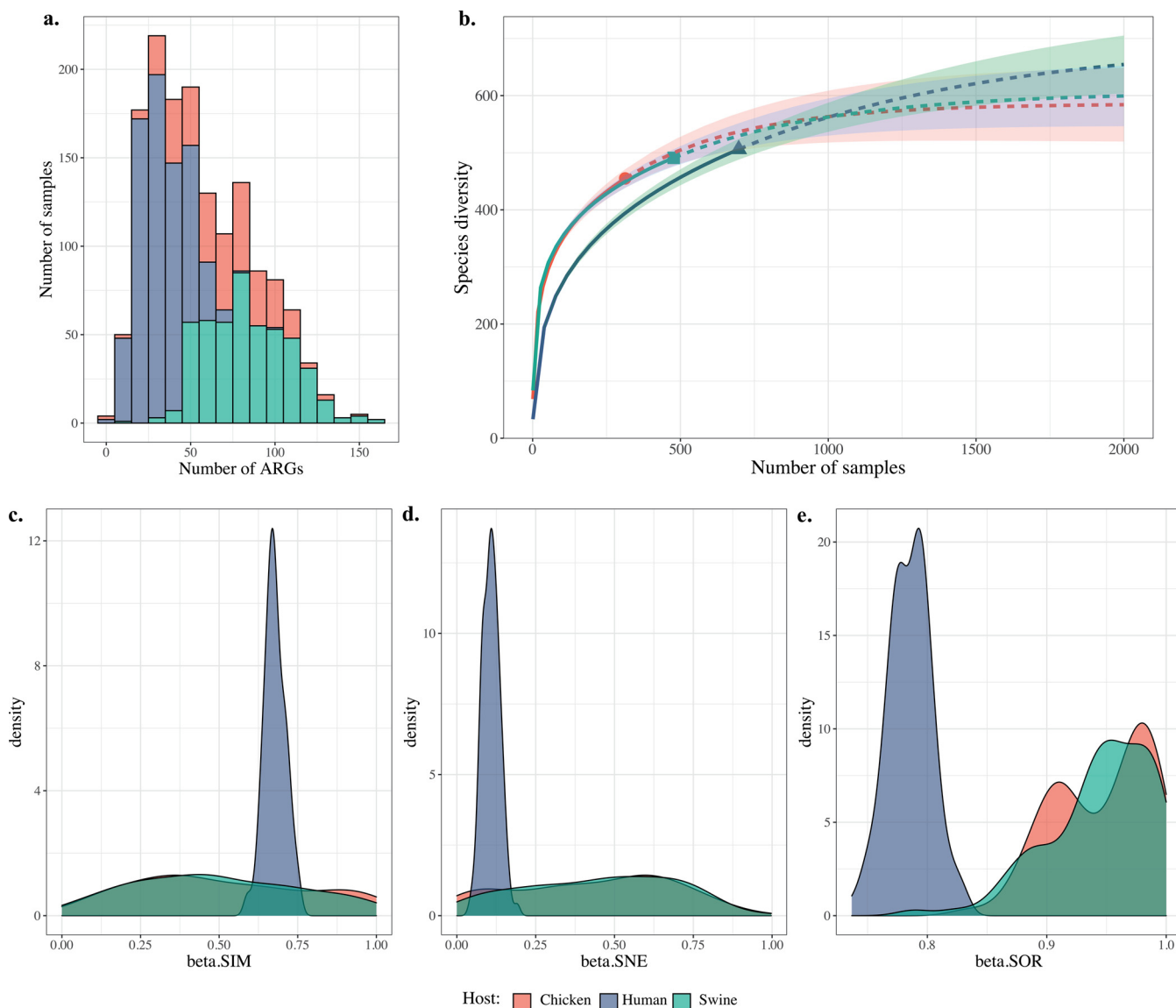


FIG 1 Diversity comparison of acquired ARGs in the hosts human, chicken, and swine. (a) The frequency distribution of ARGs in all samples; x axis shows bins with 10. (b) Rarefaction curve of the numbers of ARGs interpolated and extrapolated for both types of hosts that are calculated by iNEXT and plotted using R. (c to e) Distribution of 3 Sørensen-based multiple-site dissimilarities accounting for the spatial turnover and the nestedness components of beta diversity, and the sum of both values for ARGs in both types of hosts were calculated using R package betapart with 100 resample times for 10 samples. beta.SIM, value of the overall beta diversity, measured as Sorensen dissimilarity; beta.SOR, value of the turnover component, measured as Simpson dissimilarity; beta.SNE, value of the nestedness component, measured as nestedness-resultant fraction of Sorensen dissimilarity. The same color pattern was used for all panels to represent the host type.

resistome is closer to that in swine than that in chicken (Fig. 2d). This finding is consistent with observations in Fig. S5. For these separations between hosts, differences in antibiotic exposure might be an underlying reason. Unexpectedly, there is a closer distance between ARGs from human hosts in Hong Kong to swine hosts in China than to human hosts in China (Fig. 2d). This finding drives us to explore the reasons in the following section since it indicates the potential risk of ARG transfer between two types of hosts.

Sharing of ARGs between humans and food animals. In total, 863 acquired ARGs with different variants from all microbiota were observed in the present study, of which 493 were identified to be shared by all swine gut microbiota, while 457 and 508 ARGs were observed in all chicken and human gut microbiota, respectively (Fig. 3a). Among these 863 ARGs, 345 ARGs were shared by humans and swine and 214 were shared by humans and chicken (Fig. 3a). Except for ARGs conferring resistance to fusidic acid that are absent in human hosts, the frequencies of the occurrence for all

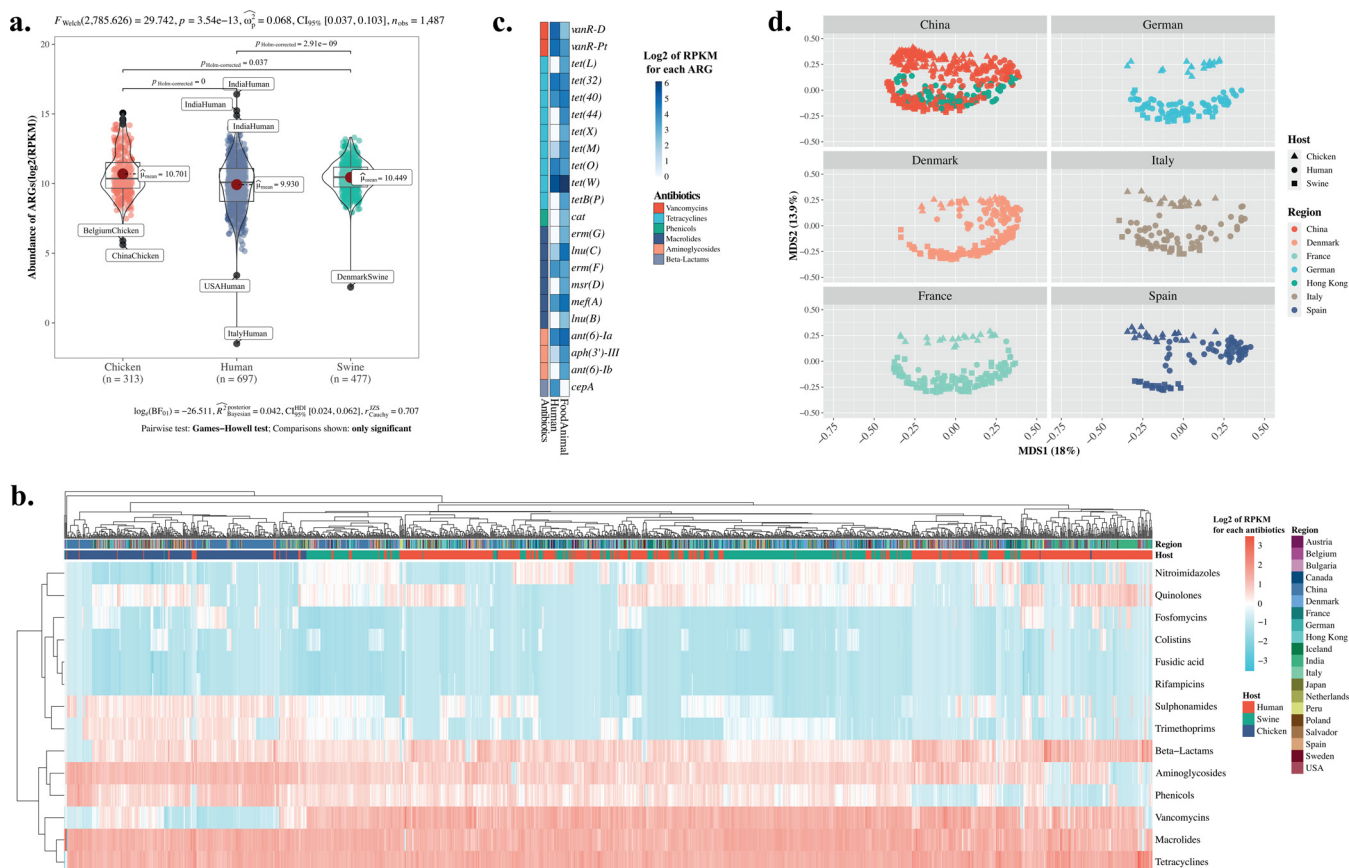


FIG 2 Comparison for the relative abundance of acquired ARGs in human and food animals. (a) Relative abundance of summed ARGs in two types of human and food animal hosts, and each colored point represents one sample. The difference between two types of groups was compared using Wilcoxon rank-sum test with a P value of <0.05 . (b) Heatmaps of the relative abundance of acquired ARGs in human and food animal fecal samples in the present study clustered in each group of hosts in columns and class of ARGs in rows. Relative abundance of each ARG has been standardized with base 2 logarithm and scaled as shown in the scale bar on the top right of the figure. All labels of groups are in Table 1 and Table S1. Assignations of ARGs to antibiotics referred to WHO ATC code J01. (c) Heatmap of significantly different relative abundance of acquired ARGs (represented by \log_2 RPKM in each gene) in two types of hosts, namely, human and food animals, calculated using the Wilcoxon tank-sum test with false-discovery rate (FDR) corrections. Names of ARGs are adapted from the ResFinder database. (d) Principal coordinate analyses of the relative abundance of acquired ARGs for each pair of hosts of human and food animals in the same region. Since most pigs are imported from China in Hong Kong, human in Hong Kong and China and swine from China were plotted in the same faceted panel. Each colored point or shape represents one sample with shapes and colors to differentiate hosts.

classes of ARGs in both types of hosts were comparable (Fig. 3b). When both types of hosts were sampled in the same regions, 303 ARGs were observed in both human and swine hosts (see Tables S3 and S4 online at <https://doi.org/10.6084/m9.figshare.20514555.v1>).

Since mobile genetic elements are the main contributors for the transferability of ARGs, information about ARGs located on plasmids or chromosomes was determined using contigs they inhabited. Although we have used several methods to classify the contigs into chromosomes, 21.4% of contigs carrying ARGs (21,761 out of 101,945) are unclassified. In total, the highest ratio of plasmids versus chromosomes was observed in chicken, and it was lowest in human (chicken, 2.3 [17,942/7,801]; swine, 1.1 [17,364/15,990]; humans, 0.84 [9,635/11,452]). Within each type of contig, the ratio of contigs carrying ARGs conferring resistance to each class of antibiotics demonstrates that ARGs conferring resistance to aminoglycosides, macrolide, tetracycline, and vancomycin are major classes (Fig. 3c).

To further explore the potential transferability of these ARGs between the two types of hosts, insertion sequence (IS) elements were identified in all contigs carrying ARGs in all samples. Taxonomic assignments of these contigs and whether these contigs are located on a chromosome or on plasmids were determined. The main groups of ARGs associated with ISs in each type of samples were plotted (Fig. 3d). Aminoglycoside- and macrolide-related ARGs are abundant in both types of hosts. ARGs conferring

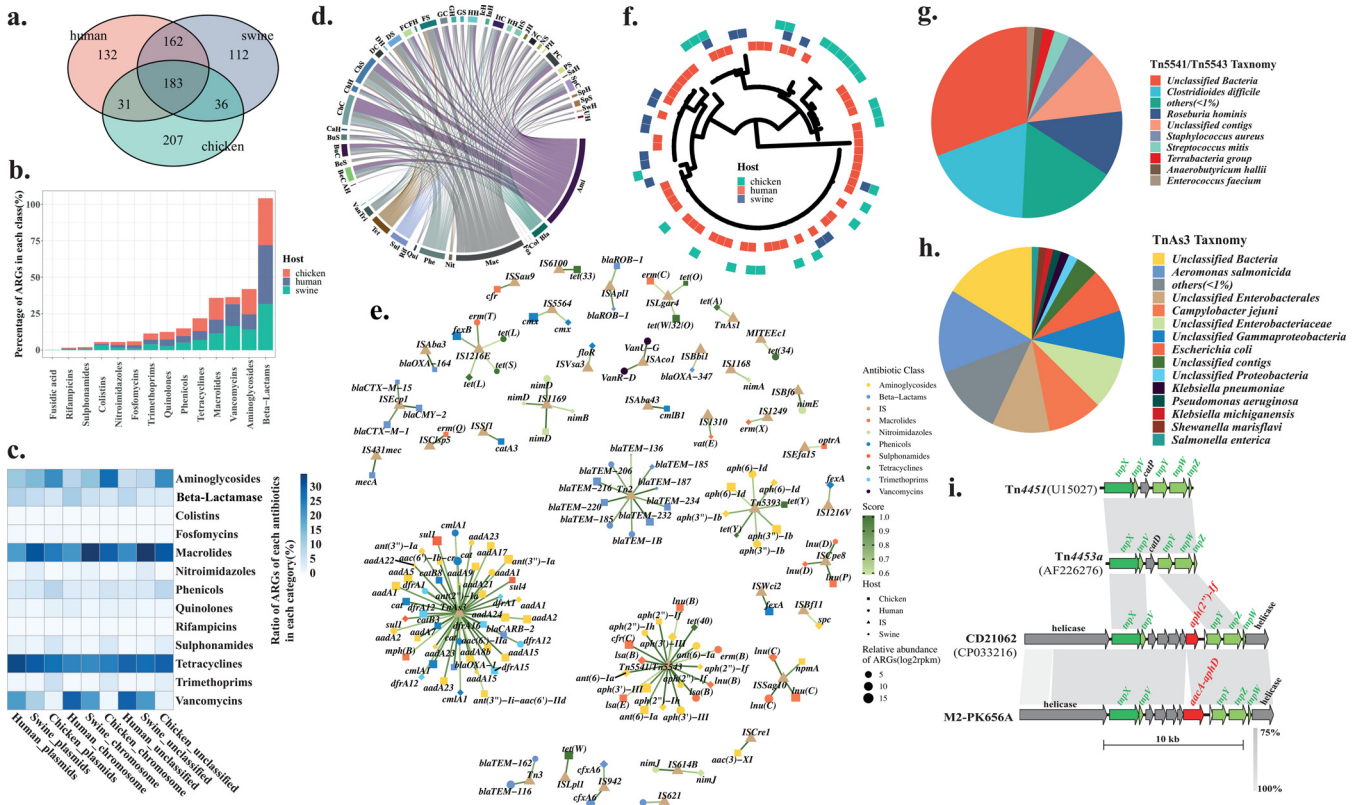


FIG 3 Shared ARGs in humans and food animals. (a) The Venn plot of the numbers of ARGs shared by human and food animals in total. (b) Frequency of ARGs assigned to each antibiotic class in both types of hosts. (c) Heatmap of the frequencies of ARGs assigned to each antibiotic class predicted to occur in plasmids and the chromosome in both types of hosts. (d) Circos diagram shows the presence of the shared antibiotic classes conferred by ARGs associated with mobile genetic elements in both food animals and human hosts (183 ARGs). Links represent the frequencies of ARGs observed in each type of samples. Abbreviations for each type of samples can be found in Table S1; antibiotics classes and their abbreviations are as follows: vancomycin, Van; trimethoprim, Tri; tetracycline, Tet; sulfonamide, Sul; rifampicin, Rif; quinolone, Qui; phenicol, Phe; nitroimidazole, Nit; macrolide, Mac; fosfomycin, Fos; colistin, Col; beta-lactams, Bla; and aminoglycoside, Ami. (e) Network of representative ARGs shared by both types of hosts and their associated insertion sequences. The size of each shape represents the relative abundance of ARGs. The color of each shape indicates the class of antibiotic to which ARGs are resistant. The edges connected by nodes are scores that strand for the ratio of each ARGs within all ARGs associated with the same IS. (f) Phylogenetic trees of representative ARG, *ant(6)-la*, that is associated with the same IS, Tn4451/Tn4453, in both types of hosts, namely, human (red) and food animal (chicken [green] and swine [blue]). Taxonomies of contigs possessing two major groups of ISs, namely, Tn4451/Tn4453 (g) and TnA53 (h) that are associated with ARGs. (i) The alignment of the representative Tn4451/Tn4453-like transposons (CD21062 and one human sample, namely, M2-PK656A from Hong Kong) with typical Tn4451/Tn4453.

resistance to phenicols were observed mainly in swine from China. The main ISs associated with all classes of ARGs were identified and plotted in a network which shows two major clusters (Fig. 3e). Tn4451/Tn4453 is the hub of ARGs conferring resistance to aminoglycosides and macrolides that were observed in both types of hosts (Fig. 3e). Another type of transposon, namely, TnA53, was observed as the hub of aminoglycoside, phenicol, and trimethoprim resistance genes. Both types of hosts are also observed to have these ARGs. To confirm the observation that both types of hosts carry the same ARG associated with the same IS, one phylogenetic tree was constructed using the *ant(6)-la* gene that is connected between two types of hosts through Tn4451/Tn4453 (Fig. 3f). Additionally, 10 ARGs with clinical significance were also clustered and plotted in the network to show the connections of these ARGs in different types of hosts (see Fig. S9 online at <https://doi.org/10.6084/m9.figshare.20522313.v2>), and one representative ARG, namely, *aac(3)*, was chosen to be used in the phylogenetic analysis (see Fig. S6 in the supplemental material). The phylogenetic analysis indicates that several dominant variants contribute to the sharing. The contigs possessing this type of transposon were classified into taxonomic units demonstrating that several dominant groups of commensal bacteria possess these transposons with ARGs (Fig. 3g and h). The gene arrangement in Tn4451/Tn4453 was deciphered, and

one key recombinase encoded by *tnpX* was observed (Fig. 3i). TnpX has been shown to play a major role in the excision of mobile elements (11). All the contigs carrying Tn4451/Tn4453 were observed mainly in human samples from China and Hong Kong, while it also occurred in swine samples from other geographical regions (see Table S2 at <https://doi.org/10.6084/m9.figshare.20514555.v1>). In these Tn4451/Tn4453-carrying contigs, particularly, the recently reported ARG *cfr(C)* encoding resistance to oxazolidinones was observed.

Sharing of ARGs between human and swine in China and Hong Kong. The observation of a close relationship of ARGs between humans and swine gut drove us to further study the ARGs between these two types of hosts by taking those from Hong Kong and China as examples. In total, 5,631 and 3,173 contigs carrying ARGs were identified in human samples from Hong Kong and China, respectively, compared with 13,969 contigs identified in swine of China (more details of these contigs, see Table S5 in the supplemental material and Fig. S8 online at <https://doi.org/10.6084/m9.figshare.20514555.v1>). Among these contigs, IS elements were predicted on 858 (15.2%), 458 (14.4%), and 1,832 (13.1%) in human samples from Hong Kong and China and swine samples in China, respectively.

Results reveal that more ARGs (216 ARGs) in human guts from Hong Kong are shared with swine guts from China than those in other geographical regions (Fig. 4a; see Table S5 at <https://doi.org/10.6084/m9.figshare.20514555.v1>). Enrichment of *oqxAB* (encoding efflux pump), *fosA* (fosfomycin resistance gene), and *aph* (encoding aminoglycoside phosphotransferase) variants are ARGs shared among these three host groups. The relative abundance of shared ARGs in each group was shown in Fig. 4b. It is clearly demonstrated that more ARGs conferring resistance to aminoglycosides, macrolides, tetracyclines, and vancomycin are shared by China swine and Hong Kong human samples. When we compared Hong Kong human samples with China swine and human samples, 158 ARGs were shared by these three groups (Fig. 4a). The most abundant ARGs associated with ISs confer resistance to macrolides, followed by those assigned to vancomycin, aminoglycoside, beta-lactam, and tetracycline classes (Fig. 4c).

Mobile genetic elements, including IS, resolvase, phage integrase and recombinase, and conjugation elements were identified on the contigs possessing ARGs. The taxonomy of these contigs carrying ARGs was annotated using Kraken2. Contigs assigned to plasmids were identified as well. In all three groups of microbiomes, Tn4451/Tn4453 was identified as the most common IS, as shown by only considering both types of hosts as a whole (Fig. 4d). Tn4451/Tn4453 was observed to be occupied by different taxonomic groups, and it was correlated with diverse ARGs (Fig. 4e and f). In the Hong Kong human samples, *aph* genes conferring resistance to aminoglycosides were correlated with Tn4451/Tn4453. In this study, Tn4451/Tn4453 was carried mainly by chromosomes from *Clostridioides*, *Roseburia*, and *Streptococcus* species and plasmids in the Hong Kong cohort (Fig. 4e). Interestingly, close strains of *Clostridioides* and *Roseburia* species carrying Tn4451/Tn4453 were also observed in human and swine hosts of China (Fig. 4e). Further study demonstrates that *Clostridioides difficile* is the main host of Tn4451/Tn4453 in three types of hosts in China and Hong Kong. TnAs3 was observed mainly in *Enterobacteriaceae*, including in *Escherichia coli*, *Klebsiella* sp., and *Salmonella* sp., with more diverse ARGs. *Bacteroides* sp. with more variable insertion elements was observed in three groups of hosts.

Since the unique feature of the Hong Kong samples was observed, more details in ARGs were surveyed. In all Hong Kong samples, 310 ARGs were identified, but only one [*erm(B)*] was present in all subjects, compared with 17 to 97 ARGs (median, 46.5) observed in each sample. However, 30 ARGs were observed in half of the subjects from Hong Kong ($n \geq 37$ out of 74), and 112 ARGs were unique in only one subject. Commensal bacteria, including *Clostridiales* (*Ruminococcus* and *Faecalibacterium*) and *Bacteroidales* (*Bacteroides*, *Alistipes*, and *Coproccoccus*) species, were observed as the predominant group carrying IS-associated ARGs, but there was no obvious trend of distributions by populations, age, or sex (see Fig. S7 in the supplemental material). A

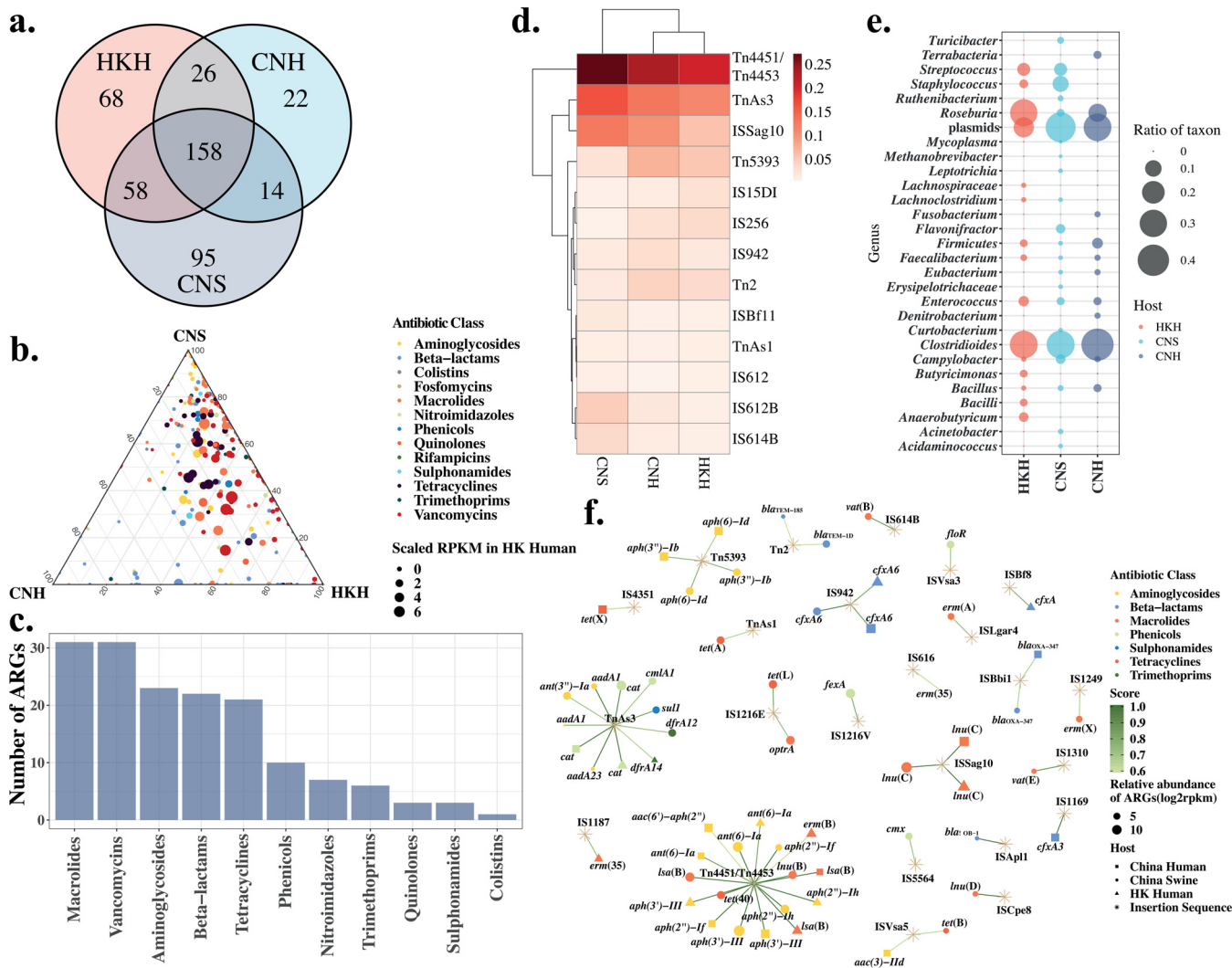


FIG 4 Shared antibiotic resistance genes between CNH, CNS, and HKH groups. (a) The Venn plot of the numbers of ARGs shared by pairs of host groups. (b) Ternary plot of shared ARGs for each pair of hosts. The size of the dot indicates the abundance of ARGs in human of Hong Kong. (c) The number of ARGs in each class shared by three groups of hosts (n = 158). (d) Heatmap of the frequencies of insertion elements associated with ARGs higher than 0.005 in all three groups of hosts. Two-ways clustering in column and row were performed. (e) The frequencies of taxonomic groups on the genera level that carry ARGs and insertion element Tn4451/Tn4453 in all three groups of hosts. (f) Network of ARG-associated insertion sequences in three types of hosts. The size of the shape represents the relative abundance of ARGs. The color stands for the class of antibiotic to which ARGs are resistant. The shape shows the type of hosts or insertion elements. The scores that represent the ratios of ARGs associated with ISs are illustrated as a color gradient of links between nodes. CNS, China swine; HKH, Hong Kong human; CNH, China human.

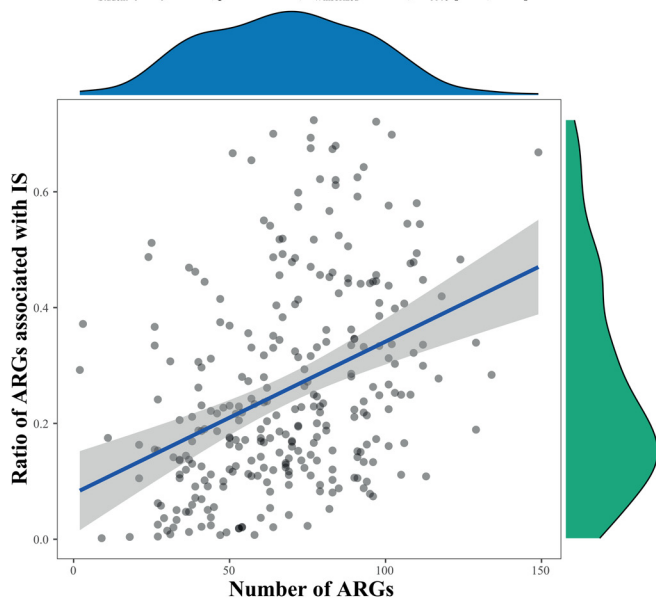
batch of ARGs was observed with a higher abundance than the abundance of contigs that these ARGs were dwelling in, in particular, *tet(34)*, *erm(B)*, *bla_{CTX-M}* and *tet(Q)*. The associations of ARGs with bacterial hosts and IS elements are shown in Fig. S7. Tn4451/Tn4453 was observed with more links with ARGs, particularly quinolone resistance genes in *Bacteroides* (Fig. S7). *Clostridioides* bacteria carry more ARGs assigned to vancomycin, macrolides, and quinolones. ARGs conferring resistance to beta-lactams were overrepresented in *Clostridioides*, *Eubacterium*, and plasmids.

Transferability of ARGs in human and swine hosts and main ARG contributors.

In the present study, the transferability of ARGs in both types of hosts was evaluated using an index, namely, Antibiotic Resistance Gene Transferability Index (ARGTI) that was calculated from the ratio of ARGs associated with mobile genetic elements compared with total ARGs in individual samples (Fig. 5). Generally, the ARGTI in a human host (median, 0.0831) is lower than that in swine (median, 0.1051) and chicken (median, 0.2180). A higher variance of ARGTI in human is observed than that in swine, consistent with the higher beta-diversity of ARGs in human hosts. Along with the increase

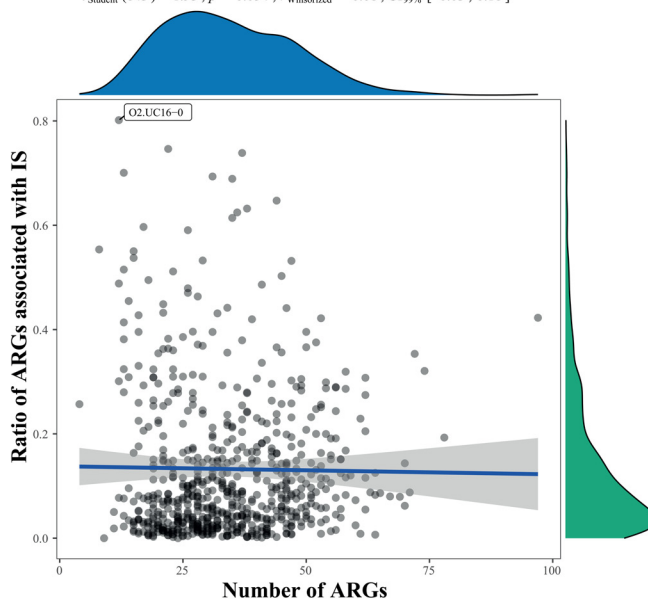
a. Host: Chicken

$t_{\text{Student}}(288) = 7.04, p = 1.4e-11, \hat{r}_{\text{Winsorized}} = 0.38, \text{CI}_{99\%} [0.25, 0.50]$



b. Host: Human

$t_{\text{Student}}(643) = 1.93, p = 0.054, \hat{r}_{\text{Winsorized}} = 0.08, \text{CI}_{99\%} [-0.03, 0.18]$



c. Host: Swine

$t_{\text{Student}}(473) = -4.52, p = 7.81e-06, \hat{r}_{\text{Winsorized}} = -0.20, \text{CI}_{99\%} [-0.31, -0.09]$

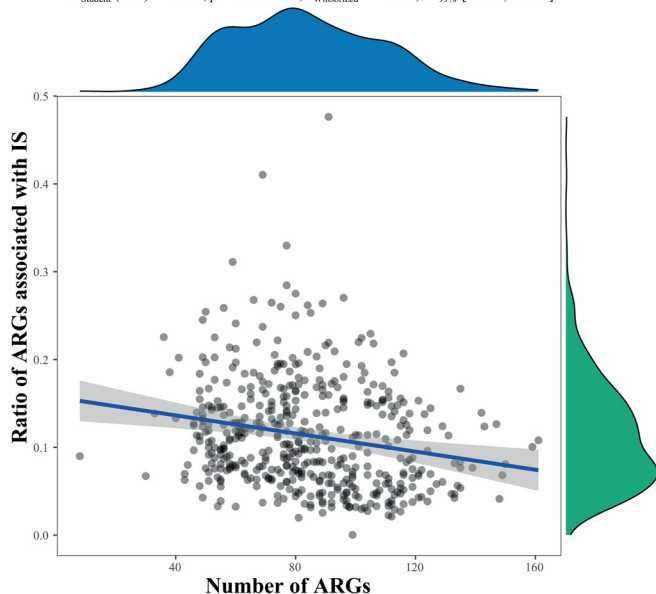


FIG 5 Model of the ARG transferability in different hosts. The transferability (y axis) was plotted with the number of ARGs in each sample (x axis). The transferability of ARGs in each sample is represented by the transmission index that was calculated from the ARGs associated with insertion elements compared with the total ARGs in each sample. The percentage bend correlation coefficient robust test was done for each type of hosts with confidence intervals at a 0.99 level.

of observed ARGs in each sample, ARGTI is negatively correlated with the number of ARGs in humans and swine, compared with a positive correlation in chicken (Fig. 5).

Linear discriminant analysis effect size (LEfSe) was used to identify discriminative ARG types between human and swine hosts (Fig. 6) demonstrating that ARGs conferring resistance to macrolides and beta-lactams are the main contributors.

DISCUSSION

One Health studies have been performed to survey the relatedness of bacterial isolates carrying ARGs in human, animal, and environmental commensals, e.g., for

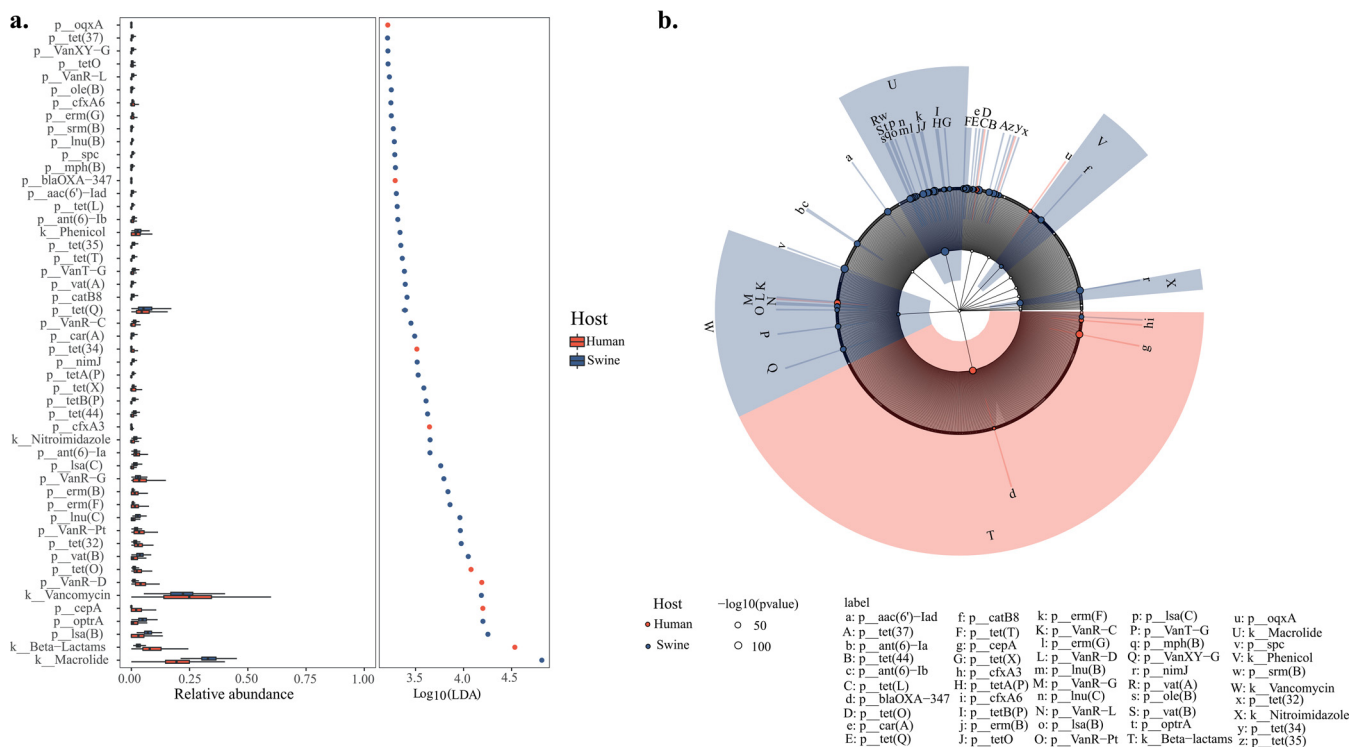


FIG 6 Linear discriminant analysis effect size (LEfSe) was used to identify discriminative ARGs between human and food animals. (a) Histograms of linear discriminant analysis (LDA) scores of the relative abundance of ARGs inhabiting two types of hosts. Key types of differently abundant ARGs were identified using the LEfSe algorithm in different hosts, and their relative abundance was presented accordingly. (b) Cladograms derived from LEfSe analysis of differential ARGs in both types of hosts. The central point denotes the root of the tree and expanded to each ring representing the next lower taxonomic level from antibiotic classes to ARGs. Each circle's diameter represents the relative abundance of the ARGs.

Klebsiella pneumoniae (12) and *E. coli* (13). The case studied extensively is about methicillin-resistant *Staphylococcus aureus* (MRSA) CC398 in humans and pigs (14, 15). High pig density leading to a significant increase of MRSA CC398 in hospitals in Spain was proposed recently (16). The abundance of genes conferring resistance to six classes of antibiotics together with class 1 integrase and the abundance of IS6100-type transposons were observed in one study on swine farms from China (17). However, there is not much convincing evidence to show the relationships between pathogens of *E. coli* from food animal to bloodstream infection in human based on whole-genome studies (13), particularly commensal bacteria. Indeed, it is a challenge to identify sources of AMR, and horizontal gene transfer appears to play a more significant role than clonal expansion in the spread of AMR. In the present study, ARGs associated with mobile genetic elements and their bacterial hosts were investigated comprehensively in human and food animal fecal metagenomes. Through comparisons between food animals, with swine and chicken as examples, and human hosts, particularly swine samples from China and human samples from China and Hong Kong, the predominantly shared ARGs were identified. Their associated mobile genetic elements and bacterial hosts were identified. These shared ARGs are important to human health and deserve to be studied further in depth with more human microbiota data. Therefore, we hypothesize that the strong effects of ARGs in swine on the human population are greater than we have expected thus far. However, a caveat of current metagenome data set, particularly regarding ARGs from the Hong Kong population that were constrained to one study on colorectal cancer, is that researchers should be cautious before drawing any solid conclusions (18). On another hand, the transferable capability of ARGs should be validated with more information on MGEs that are integrated with ARGs, although one index has been developed in the present study.

Commensal bacteria affiliated with two major phyla, namely, *Firmicutes* and *Bacteroides*, have been observed as inhabiting the most abundant microbial habitats in human and food

animals, including swine (19, 20). In the present study, *Clostridiales*, including *Clostridioides* and *Roseburia* species carrying Tn4451/Tn4453 with diverse ARGs, were observed to be shared by human and swine hosts in China. The mobilizable transposons Tn4451/Tn4453 could be excised, which was mediated by the product of the *tnpX* gene which encodes a member of the large resolvase (11). The transposase TnpX was identified widely in human and swine gut microbiota. This report is the first one about contributions of Tn4451/Tn4453 to antibiotic resistance in the wide type of human and food animal hosts. Noticeably, this type of transferable cassettes carrying ARGs was observed in several commensal opportunistic bacterial hosts, including *Clostridioides*, *Campylobacter*, and *Bacteroides*, indicating their potential transferability. Overall, from what we know of modern food production, it is clear that antibiotic use does contribute to the resistance of our commensal and pathogenic microflora, but the scale of this contribution has not yet been quantified. We find that ARGs are more likely to be spread by conjugation via integrative conjugative elements or integrative mobilizable elements than transduction via prophages. Particularly, a new multidrug resistance gene, *cfr(C)*, identified in recent studies was observed within this type of common transposon, namely, Tn4451/Tn4453, in commensal bacteria. The variable commensal bacterial hosts of this type of transposon suggest the potential transmission of this mechanism via the foodborne route, warranting enhanced efforts to monitor its spread in foodborne pathogens.

Within the human host, ARGs are diverse in populations from different regions with India having the most prominently high observations followed by the human cohort in Hong Kong and China. China and India have been reported as the hot pots of AMR in food animals (3). The link between swine and host in China and Hong Kong proposed in the study possibly might explain the high abundance of ARGs in India. Unfortunately, the resistome data in food animals in India is lacking currently. All these data suggest the importance of One Health to surveil AMR in humans and food animals together, although the similarity degree between swine and human hosts in different regions is different. On another hand, even in the same region, the heterogeneous compositions of ARGs were observed. However, there are no correlations between age and ARGs in the cohort from Hong Kong, and different age groups also possibly carry different types of relative abundances of ARGs, e.g., in the Hong Kong cohort. This information necessitates the development of personal and specific therapies to treat infections in people. Considering the diversity of ARGs, there is no association between ARGs and age, as even the abundance of ARGs in old people is higher than that in younger people. However, the dominant group of bacteria carrying ARGs associated with mobile genetic elements is similar within the Hong Kong cohort. Additionally, several types of ARGs were observed with an extraordinarily high abundance compared with contigs they were from in the Hong Kong cohort. All these data possibly show insights into the choice of antibiotics for therapy.

The main strengths of this study are that ARGs from diverse groups of the Hong Kong cohort were surveyed and mobile genetic elements correlated with the diversity of ARGs. Shared relationships between swine and humans in China were explored. Nonetheless, the study has several limitations. First, because of the weakness of the sequencing methods, namely, the short-read sequencing Illumina platform, we suppose that a high ratio of mobile genetic elements has been disrupted from associated ARGs, as shown in the whole-genome sequencing on culture pathogens, leading us to evaluate the transferable capability of ARGs. However, the low ratio of contigs possessing ISs was observed, and it also excludes the intrinsic existence of fewer ARGs located adjacent to mobile genetic elements, leading to less of an opportunity for them to transfer within human gut microbiota. Second, minor groups of bacteria have still not been studied completely. Third, sampling size was insufficient to make solid conclusions, as was shown in the rarefaction curves of the diversity of ARGs in both hosts. All these limitations might be resolved in the future with more efforts on samplings and sequencing depths with different sequencing platforms, including the Nanopore sequencer. Finally, the intestinal source locations (e.g., ileum, cecum, and colon) for the majority of the gut metagenomes were not available. Therefore, the ARGs could not be further analyzed by intestinal source locations, although it is recognized that the

microbiome compositions of gut locations are different in terms of microbial abundance and diversity (21).

In conclusion, potential AMR hot spots were observed in the present study with a comprehensive surveillance of the resistome, especially those ARGs that are the most likely candidates for horizontal transfer to pathogens. This study emphasizes on the importance of surveillance for detecting emerging resistance threats before they spread, so as to mitigate threats. The knowledge from this study will benefit pathogen control and prevention efforts to protect human health.

MATERIALS AND METHODS

Sampling and sequencing. In the present study, 21 subjects (all volunteers) who met the following inclusion criteria were recruited: (i) not currently suffering from any acute illness, including fever, acute gastroenteritis, or acute respiratory tract infection; (ii) not currently taking any antibiotics in the past 3 months; (iii) no intake of probiotic supplements; and (iv) no history of hospitalization in the past 3 months. In total, 11 male and 10 female subjects were recruited.

In addition, metagenomes ($n = 1,465$) of slaughter pigs from European countries ($n = 181$) (22) and China ($n = 295$) (20) and poultries from China ($n = 135$) and European countries ($n = 178$), as well of humans from several regions (23–28), including one study on a Hong Kong population (18), were integrated with 21 newly sequenced metagenomes of a Hong Kong cohort and one swine from Hong Kong for further analyses of the resistome (Table 1 and Fig. 1).

A standardized fecal sampling protocol to process stool samples was followed in the present study (29). Sequencing was conducted on a NextSeq platform (Illumina, San Diego, CA) at Novogene Co., Ltd. (Beijing, China).

The study met the standards for the ethical treatment of participants and was approved by the Institutional Review Board of the University of Hong Kong/Hospital Authority Hong Kong West Cluster (reference number UW 18–278). Study participants gave informed consent.

Metagenome assembling. To be consistent, all samples used in the present study were processed using the same pipeline from quality control, gene assembly and gene annotations, to taxonomic study.

An in-house pipeline was developed to perform calling of ARG contigs assembled from raw reads in the present study. Briefly, quality control was performed using Trimmomatic (30). Host (human/swine) DNA removal was done by alignment against host genomes using BWA-MEM (31) and SAMtools (32). Reference genomes of the hosts are WTSI_X_Y_pig V2 ([GCF_000003025.6](https://www.ncbi.nlm.nih.gov/genome/10867)) for swine and Hg19 ([GCA_000001405.26](https://www.ncbi.nlm.nih.gov/genome/10867)) for human. The metagenome module of SPAdes was used to do assemble (33). Qualified reads were submitted to the Metagenomic Phylogenetic Analysis pipeline (MetaPhlan2) for taxonomic profiling (34). Taxonomic identification on strain level was performed using StrainPhlan (35). Kraken2 was also used to validate taxonomic analyses using contigs (36). MetaBAT2 (37) was used to bin genomes from contigs, and CheckM was performed to assess the quality of bins (38).

ARG calling. Genes was called using MetaGeneMark from contigs longer than 500 bp (39) and searched against the most recent version of the ARG databases of ResFinder (updated in September 2020) (40). A BLAST E value of $1e-10$, a 70% query coverage, and an 80% similarity as suggested previously (23) were adopted to filter searching results using in-house python scripts.

For ARGs, the classification of antibiotics was inferred based on the WHO ATC code J01 (https://www.whooc.no/atc_ddd_index/). For each class of antibiotics, the sum of the relative abundance of each antibiotic resistance gene mapped to the same antibiotic was categorized for this type of antibiotics (23). All matrix files were generated using custom python and R scripts.

Rarefaction curves for ARGs in both hosts were extrapolated using iNEXT (41). Sørensen-based multi-site dissimilarities in both types of hosts were calculated using R package betapart with 100 resample times for 10 samples (42). Shannon and Simpson diversities were calculated using R package vegan (<https://cran.r-project.org/web/packages/vegan/>). The number of present and shared ARGs was calculated for the compared sample types, and Venn diagrams were drawn in R using the draw.pairwise.venn command of the VennDiagram package (<https://CRAN.R-project.org/package=VennDiagram>). Negative binomial general linearized models (GLMs) were used to predict the means and standard errors (SEs) of the relative sum abundances in different sample types with Turkey's *post hoc* test to correct since the distributions of ARGs do not fit a normal or Poisson distribution due to overdispersion. For principal coordinate analysis (PCoA), raw matrix files were used to generate the Bray-Curtis distance matrix that was submitted to the R package vegan. A permutational multivariate analysis of variance (PERMANOVA) between different groups was done with *adonis* in vegan with the similarity index using 9,999 permutations, and the resulting *P* values were corrected with the Benjamini and Hochberg procedure for multiple testing using the *p.adjust* command in R. All plots, bar charts, PCoAs, and heatmaps were generated using the ggplot2 package (<https://ggplot2.tidyverse.org/>).

Mobile genetic elements. Integrons were identified using IntegronFinder v.1.4 with the `-local_max` option (43). Integrative conjugative elements (ICEs) and integrative mobilizable elements (IMEs) were detected using MacSyFinder v.1.0.2 with CONJScan profiles (44). Integrases were called using the PFAM profile PF00589 for tyrosine recombinases and profiles PF00239 and PF07508 for serine recombinases (45). Insertion sequences (ISs) were detected using ISfinder and ISEScan on contigs (46). Prophages and viruses were identified using an in-house script with *vpn* to run PHASTER (47) and Virfinder (48), respectively. Plasmids were predicted using PlasFlow (49) and PLSDB (50). Reads were aligned using BWA-MEM

to all genes predicted using MetaGeneMark. pileup.sh from the bbmap suite v37.25 was used to calculate reads per kilobase per million (RPKM) and the number of reads mapped to each gene (BBMap short read aligner, <http://sourceforge.net/projects/bbmap>).

ARGs coordinated with MGEs and taxonomy. The co-occurrence of ARGs and MGEs was surveyed through identifying coordinates (<10 kb) between ARGs and MGEs on the same contig using in-house python scripts. On another hand, the Bray-Curtis distance matrix of species, ARGs, and MGEs was calculated with the *vegdist* command in the *vegan* package and compared to observe possible correlations between taxonomy and the resistome and mobilome that occurred in at least half of all samples. Comparisons were done using Mantel's test from *vegan*. Network analyses were conducted in R using the *vegan*, *igraph* (<https://igraph.org/r/>), and *Hmisc* packages. Associations of ARGs between bacterial hosts and mobile genetic elements were also visualized with the *HiveR* R package (<https://CRAN.R-project.org/package=HiveR>).

Variations of ARGs. In the present study, a list of clinically significant ARGs was selected for further analyses on variants. For each ARG, genes (length, >70% coverage) present in pairs of hosts in the same regions were extracted and aligned using MAFFT with the L-INS-i option (51). The phylogenetic tree for each alignment was built using IQ-TREE (52), and the associated features including the relative abundance of each gene, flanking IS element, whether or not on plasmid, host source and geographic region were aligned with the tree using Ggtree (53). All these genes were also clustered using PopPUNK (54) and visualized with Cytoscape (<https://cytoscape.org>).

Modeling transferability of ARGs. In the present study, the antibiotic resistance gene transferability index (ARGTI) was developed to evaluate the risk of spread of ARGs within and between hosts.

The ARGTI is

$$\text{ARGTI} = \sum_{i=1}^n \frac{S_{a-mi}}{S_{ai}} \quad (1)$$

in which, S_{a-mi} stands for the sum of the abundance of ARGs that are associated with mobile genetic elements, while S_{ai} represents the sum of the abundance of ARGs in individual microbiota.

Effectors on ARGs. Linear discriminant analysis effect size (LEfSe) was used to identify discriminative ARG types between human and swine (55). The alpha value for the factorial Kruskal-Wallis test and the pairwise Wilcoxon rank-sum test is 0.01. Multiclass analysis was performed using the all-against-all comparison with 3.0 as the threshold on the linear discriminant analysis (LDA) score. The false-discovery rate (FDR) correction for *P* values of the discriminative ARGs was performed using the Benjamini-Hochberg method. Significance was determined by using the *q* value of 0.01.

Data availability. The DNA sequences from 22 fecal samples are deposited in the European Nucleotide Archive under the project accession number [PRJNA701076](https://www.ebi.ac.uk/ena/record/PRJNA701076). All data, summarized figures, tables, and codes used in the study have been deposited online at https://github.com/huiluocao/gut_resistome.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, TIF file, 2.6 MB.

FIG S2, TIF file, 0.9 MB.

FIG S3, TIF file, 2 MB.

FIG S4, TIF file, 2.8 MB.

FIG S5, TIF file, 2.4 MB.

FIG S6, TIF file, 1.7 MB.

FIG S7, TIF file, 2.6 MB.

ACKNOWLEDGMENTS

We thank the Computational Bioscience Research Centre at The King Abdullah University of Science and Technology for support on computational resources.

This work is supported by grants from (i) the Consultancy Service for Enhancing Laboratory Surveillance of Emerging Infectious Disease for the HKSAR Department of Health, (ii) the Health and Medical Research Fund, Health Bureau, HKSAR (reference: CID-HKU-1-13) and (iii) the Seed Fund for Basic Research, University of Hong Kong (reference: 201711159159).

H.C. and P.-L.H. designed the study. M.-K.T., A.L., and K.-H.C. collected fecal samples and did genomic DNA extractions for sequencing. H.C., T.-J.P., and S.B. performed bioinformatics analysis, and H.C. plotted all figures. H.C. and P.-L.H. wrote the manuscript. All authors reviewed the manuscript.

We declare no competing interests.

REFERENCES

- Crofts TS, Gasparri AJ, Dantas G. 2017. Next-generation approaches to understand and combat the antibiotic resistome. *Nat Rev Microbiol* 15: 422–434. <https://doi.org/10.1038/nrmicro.2017.28>.
- Aarestrup FM. 2015. The livestock reservoir for antimicrobial resistance: a personal view on changing patterns of risks, effects of interventions and the way forward. *Philos Trans R Soc Lond B Biol Sci* 370:20140085. <https://doi.org/10.1098/rstb.2014.0085>.
- Van Boeckel TP, Pires J, Silvester R, Zhao C, Song J, Criscuolo NG, Gilbert M, Bonhoeffer S, Laxminarayan R. 2019. Global trends in antimicrobial resistance in animals in low- and middle-income countries. *Science* 365: eaaw1944. <https://doi.org/10.1126/science.aaw1944>.
- 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>.
- Wong SCY, Tse H, Chen JHK, Cheng VCC, Ho P-L, Yuen K-Y. 2016. Colistin-resistant Enterobacteriaceae carrying the mcr-1 gene among patients in Hong Kong. *Emerg Infect Dis* 22:1667–1669. <https://doi.org/10.3201/eid2209.160091>.
- Ho P-L, Wang Y, Liu MC-J, Lai EL-Y, Law PY-T, Cao H, Chow K-H. 2018. IncX3 epidemic plasmid carrying bla_{NDM-5} in *Escherichia coli* from swine in multiple geographic areas in China. *Antimicrob Agents Chemother* 62: e02295-17. <https://doi.org/10.1128/AAC.02295-17>.
- 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](https://doi.org/10.1016/S1473-3099(15)00424-7).
- Wang R, van Dorp L, Shaw LP, Bradley P, Wang Q, Wang X, Jin L, Zhang Q, Liu Y, Rieux A, Dorai-Schneiders T, Weinert LA, 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. <https://doi.org/10.1038/s41467-018-03205-z>.
- Van Boeckel TP, Glennon EE, Chen D, Gilbert M, Robinson TP, Grenfell BT, Levin SA, Bonhoeffer S, Laxminarayan R. 2017. Reducing antimicrobial use in food animals. *Science* 357:1350–1352. <https://doi.org/10.1126/science.aao1495>.
- Ho PL, Chow KH, Lai EL, Lo WU, Yeung MK, Chan J, Chan PY, Yuen KY. 2011. Extensive dissemination of CTX-M-producing *Escherichia coli* with multidrug resistance to 'critically important' antibiotics among food animals in Hong Kong, 2008–10. *J Antimicrob Chemother* 66:765–768. <https://doi.org/10.1093/jac/dkq539>.
- Lyras D, Rood JI. 2000. Transposition of Tn4451 and Tn4453 involves a circular intermediate that forms a promoter for the large resolvase, TnpX. *Mol Microbiol* 38:588–601. <https://doi.org/10.1046/j.1365-2958.2000.02154.x>.
- Ludden C, Moradigaravand D, Jamroz D, Gouliouris T, Blane B, Naydenova P, Hernandez-Garcia J, Wood P, Hadjirin N, Radakovic M, Crawley C, Brown NM, Holmes M, Parkhill J, Peacock SJ. 2019. A One Health study of the genetic relatedness of *Klebsiella pneumoniae* and their mobile elements in the East of England. *Clin Infect Dis* 70:219–226. <https://doi.org/10.1093/cid/ciz174>.
- Ludden C, Raven KE, Jamroz D, Gouliouris T, Blane B, Coll F, de Goffau M, Naydenova P, Horner C, Hernandez-Garcia J, Wood P, Hadjirin N, Radakovic M, Brown NM, Holmes M, Parkhill J, Peacock SJ. 2019. One Health genomic surveillance of *Escherichia coli* demonstrates distinct lineages and mobile genetic elements in isolates from humans versus livestock. *mBio* 10:e02693-18. <https://doi.org/10.1128/mBio.02693-18>.
- Sieber RN, Skov RL, Nielsen J, Schulz J, Price LB, Aarestrup FM, Larsen AR, Stegger M, Larsen J. 2018. Drivers and dynamics of methicillin-resistant livestock-associated *Staphylococcus aureus* CC398 in pigs and humans in Denmark. *mBio* 9:e02142-18. <https://doi.org/10.1128/mBio.02142-18>.
- van Duijkeren E, Schwarz C, Bouchard D, Cattri B, Pomba C, Baptiste KE, Moreno MA, Rantala M, Ružauskas M, Sanders P, Teale C, Wester AL, Ignate K, Kunsagi Z, Jukes H. 2019. The use of aminoglycosides in animals within the EU: development of resistance in animals and possible impact on human and animal health: a review. *J Antimicrob Chemother* 74: 2480–2496. <https://doi.org/10.1093/jac/dkz161>.
- Ceballos S, Aspiroz C, Ruiz-Ripa L, Reynaga E, Azcona-Gutiérrez JM, Rezusta A, Seral C, Antoñanzas F, Torres L, López C, López-Cerero L, Cercenado E, Zarazaga M, Torres C; Study Group of clinical LA-MRSA. 2019. Epidemiology of MRSA CC398 in hospitals located in Spanish regions with different pig-farming practices: a multicentre study. *J Antimicrob Chemother* 74:2157–2161. <https://doi.org/10.1093/jac/dkz180>.
- Johnson TA, Stedtfeld RD, Wang Q, Cole JR, Hashsham SA, Looft T, Zhu Y-G, Tiedje JM. 2016. Clusters of antibiotic resistance genes enriched together stay together in swine agriculture. *mBio* 7:e02214-15. <https://doi.org/10.1128/mBio.02214-15>.
- Yu J, Feng Q, Wong SH, Zhang D, Liang QY, Qin Y, Tang L, Zhao H, Stenvang J, Li Y, Wang X, Xu X, Chen N, Wu WKK, Al-Aama J, Nielsen HJ, Kiillerich P, Jensen BAH, Yau TO, Lan Z, Jia H, Li J, Xiao L, Lam TYT, Ng SC, Cheng AS-L, Wong VW-S, Chan FKL, Xu X, Yang H, Madsen L, Datz C, Tilg H, Wang J, Brünner N, Kristiansen K, Arumugam M, Sung JJ-Y, Wang J. 2017. Metagenomic analysis of faecal microbiome as a tool towards targeted non-invasive biomarkers for colorectal cancer. *Gut* 66:70–78. <https://doi.org/10.1136/gutjnl-2015-309800>.
- Martinez-Guryn K, Leone V, Chang EB. 2019. Regional diversity of the gastrointestinal microbiome. *Cell Host Microbe* 26:314–324. <https://doi.org/10.1016/j.chom.2019.08.011>.
- Xiao L, Estellé J, Kiillerich P, Ramayo-Caldas Y, Xia Z, Feng Q, Liang S, Pedersen AØ, Kjeldsen NJ, Liu C, Maguin E, Doré J, Pons N, Le Chatelier E, Prifti E, Li J, Jia H, Liu X, Xu X, Ehrlich SD, Madsen L, Kristiansen K, Rogel-Gaillard C, Wang J. 2016. A reference gene catalogue of the pig gut microbiome. *Nat Microbiol* 1:16161. <https://doi.org/10.1038/nmicrobiol.2016.161>.
- Quan J, Cai G, Ye J, Yang M, Ding R, Wang X, Zheng E, Fu D, Li S, Zhou S, Liu D, Yang J, Wu Z. 2018. A global comparison of the microbiome compositions of three gut locations in commercial pigs with extreme feed conversion ratios. *Sci Rep* 8:4536. <https://doi.org/10.1038/s41598-018-22692-0>.
- Munk P, Knudsen BE, Lukjancenko O, Duarte ASR, Van Gompel L, Luiken REC, Smit LAM, Schmitt H, Garcia AD, Hansen RB, Petersen TN, Bossers A, Ruppé E, Lund O, Hald T, Pamp SJ, Vigre H, Heederik D, Wagenaar JA, Mevius D, Aarestrup FM, EFFORT Group. 2018. Abundance and diversity of the faecal resistome in slaughter pigs and broilers in nine European countries. *Nat Microbiol* 3:898–908. <https://doi.org/10.1038/s41564-018-0192-9>.
- Hu Y, Yang X, Qin J, Lu N, Cheng G, Wu N, Pan Y, Li J, Zhu L, Wang X, Meng Z, Zhao F, Liu D, Ma J, Qin N, Xiang C, Xiao Y, Li L, Yang H, Wang J, Yang R, 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>.
- Lloyd-Price J, Mahurkar A, Rahnavard G, Crabtree J, Orvis J, Hall AB, Brady A, Creasy HH, McCracken C, Giglio MG, McDonald D, Franzosa EA, Knight R, White O, Huttenhower C. 2017. Strains, functions and dynamics in the expanded Human Microbiome Project. *Nature* 551:256. <https://doi.org/10.1038/nature24485>.
- Qin J, Li Y, Cai Z, Li S, Zhu J, Zhang F, Liang S, Zhang W, Guan Y, Shen D, Peng Y, Zhang D, Jie Z, Wu W, Qin Y, Xue W, Li J, Han L, Lu D, Wu P, Dai Y, Sun X, Li Z, Tang A, Zhong S, Li X, Chen W, Xu R, Wang M, Feng Q, Gong M, Yu J, Zhang Y, Zhang M, Hansen T, Sanchez G, Raes J, Falony G, Okuda S, Almeida M, LeChatelier E, Renault P, Pons N, Batto J-M, Zhang Z, Chen H, Yang R, Zheng W, Li S, Yang H, et al. 2012. A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* 490:55–60. <https://doi.org/10.1038/nature11450>.
- Dhakan DB, Maji A, Sharma AK, Saxena R, Pulikkan J, Grace T, Gomez A, Scaria J, Amato KR, Sharma VK. 2019. The unique composition of Indian gut microbiome, gene catalogue, and associated fecal metabolome deciphered using multi-omics approaches. *Gigascience* 8:giz004. <https://doi.org/10.1093/gigascience/giz004>.
- Wirbel J, Pyl PT, Kartal E, Zych K, Kashani A, Milanese A, Fleck JS, Voigt AY, Palleja A, Ponnudurai R, Sunagawa S, Coelho LP, Schrotz-King P, Vogtmann E, Habermann N, Niméus E, Thomas AM, Manghi P, Gandini S, Serrano D, Mizutani S, Shiroma H, Shiba S, Shibata T, Yachida S, Yamada T, Waldron L, Naccarati A, Segata N, Sinha R, Ulrich CM, Brenner H, Arumugam M, Bork P, Zeller G. 2019. Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer. *Nat Med* 25:679–689. <https://doi.org/10.1038/s41591-019-0406-6>.
- Li J, Jia H, Cai X, Zhong H, Feng Q, Sunagawa S, Arumugam M, Kultima JR, Prifti E, Nielsen T, Juncker AS, Manichanh C, Chen B, Zhang W, Levenez F, Wang J, Xu X, Xiao L, Liang S, Zhang D, Zhang Z, Chen W, Zhao H, Al-Aama JY, Edris S, Yang H, Wang J, Hansen T, Nielsen HB, Brunak S, Kristiansen K, Guarnier F, Pedersen O, Doré J, Ehrlich SD, Bork P, Wang J, MetaHIT Consortium. 2014. An integrated catalog of reference genes in the human gut microbiome. *Nat Biotechnol* 32:834–841. <https://doi.org/10.1038/nbt.2942>.
- Costea PI, Zeller G, Sunagawa S, Pelletier E, Alberti A, Levenez F, Tramontano M, Driessen M, Hercog R, Jung F-E, Kultima JR, Hayward MR, Coelho LP, Allen-Vercoe E, Bertrand L, Blaut M, Brown JRM, Carton T, Cools-Portier S, Daigneault M, Derrien M, Druésne A, de Vos WM, Finlay BB, Flint HJ, Guarner F, Hattori M, Heilig H, Luna RA, van Hylckama Vlieg J,

- Junick J, Klymiuk I, Langella P, Le Chatelier E, Mai V, Manichanh C, Martin JC, Mery C, Morita H, O'Toole PW, Orvain C, Patil KR, Penders J, Persson S, Pons N, Popova M, Salonen A, Saulnier D, Scott KP, Singh B, et al. 2017. Towards standards for human fecal sample processing in metagenomic studies. *Nat Biotechnol* 35:1069–1076. <https://doi.org/10.1038/nbt.3960>.
30. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
31. Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv*. <https://doi.org/10.48550/arXiv.1303.3997>.
32. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
33. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
34. Truong DT, Franzosa EA, Tickle TL, Scholz M, Weingart G, Pasolli E, Tett A, Huttenhower C, Segata N. 2016. MetaPhlan2 for enhanced metagenomic taxonomic profiling. *Nat Methods* 13:101. <https://doi.org/10.1038/nmeth0116-101b>.
35. Truong DT, Tett A, Pasolli E, Huttenhower C, Segata N. 2017. Microbial strain-level population structure and genetic diversity from metagenomes. *Genome Res* 27:626–638. <https://doi.org/10.1101/gr.216242.116>.
36. Wood DE, Salzberg SL. 2014. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* 15:R46. <https://doi.org/10.1186/gb-2014-15-3-r46>.
37. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, Wang Z. 2019. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ Preprints* 7:e27522v1. <https://doi.org/10.7717/peerj.7359>.
38. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25:1043–1055. <https://doi.org/10.1101/gr.186072.114>.
39. Zhu WH, Lomsadze A, Borodovsky M. 2010. Ab initio gene identification in metagenomic sequences. *Nucleic Acids Res* 38:e132. <https://doi.org/10.1093/nar/gkq275>.
40. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV. 2012. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother* 67:2640–2644. <https://doi.org/10.1093/jac/dks261>.
41. Hsieh TC, Ma KH, Chao A. 2016. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods Ecol Evol* 7: 1451–1456. <https://doi.org/10.1111/2041-210X.12613>.
42. Baselga A. 2010. Partitioning the turnover and nestedness components of beta diversity. *Global Ecol Biogeogr* 19:134–143. <https://doi.org/10.1111/j.1466-8238.2009.00490.x>.
43. Cury J, Jove T, Touchon M, Neron B, Rocha EPC. 2016. Identification and analysis of integrons and cassette arrays in bacterial genomes. *Nucleic Acids Res* 44:4539–4550. <https://doi.org/10.1093/nar/gkw319>.
44. Cury J, Touchon M, Rocha EPC. 2017. Integrative and conjugative elements and their hosts: composition, distribution and organization. *Nucleic Acids Res* 45:8943–8956. <https://doi.org/10.1093/nar/gkx607>.
45. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A. 2016. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res* 44:D279–D285. <https://doi.org/10.1093/nar/gkv1344>.
46. Xie ZQ, Tang HX. 2017. ISEScan: automated identification of insertion sequence elements in prokaryotic genomes. *Bioinformatics* 33:3340–3347. <https://doi.org/10.1093/bioinformatics/btx433>.
47. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* 44:W16–W21. <https://doi.org/10.1093/nar/gkw387>.
48. Ren J, Ahlgren NA, Lu YY, Fuhrman JA, Sun F. 2017. VirFinder: a novel k-mer based tool for identifying viral sequences from assembled metagenomic data. *Microbiome* 5:69. <https://doi.org/10.1186/s40168-017-0283-5>.
49. Krawczyk PS, Lipinski L, Dziembowski A. 2018. PlasFlow: predicting plasmid sequences in metagenomic data using genome signatures. *Nucleic Acids Res* 46:e35. <https://doi.org/10.1093/nar/gkx1321>.
50. Galata V, Fehlmann T, Backes C, Keller A. 2019. PLSDB: a resource of complete bacterial plasmids. *Nucleic Acids Res* 47:D195–D202. <https://doi.org/10.1093/nar/gky1050>.
51. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30:772–780. <https://doi.org/10.1093/molbev/mst010>.
52. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 32:268–274. <https://doi.org/10.1093/molbev/msu300>.
53. Yu G, Lam TT, Zhu H, Guan Y. 2018. Two methods for mapping and visualizing associated data on phylogeny using Ggtree. *Mol Biol Evol* 35: 3041–3043. <https://doi.org/10.1093/molbev/msy194>.
54. Lees JA, Harris SR, Tonkin-Hill G, Gladstone RA, Lo SW, Weiser JN, Corander J, Bentley SD, Croucher NJ. 2019. Fast and flexible bacterial genomic epidemiology with PopPUNK. *Genome Res* 29:304–316. <https://doi.org/10.1101/gr.241455.118>.
55. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C. 2011. Metagenomic biomarker discovery and explanation. *Genome Biol* 12:R60. <https://doi.org/10.1186/gb-2011-12-6-r60>.