The Shared Resistome of Human and Pig Microbiota Is Mobilized by Distinct Genetic Elements

Chao Wang, Yuqin Song, Na Tang, Gang Zhang, Sébastien Olivier Leclercq, Jie Feng

State Key Laboratory of Microbial Resources, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China
College of Life Science, University of the Chinese Academy of Sciences, Beijing, China
INRAE, Université de Tours, ISP, Nouzilly, France

ABSTRACT The extensive use of antibiotics in hospitals and in the animal breeding industry has promoted antibiotic resistance in bacteria, which has resulted in the emergence of a large number of antibiotic resistance genes in the intestinal tracts of humans and farmed animals. Genetic exchange of resistance genes between the two ecosystems is now well documented for pathogenic bacteria, but the repertoire of shared resistance genes in the commensal bacterial community and by which genetic modules they are disseminated are still unclear. By analyzing metagenomics data of human and pig intestinal samples collected in Shenzhen, China, a set of 27 highly prevalent antibiotic resistance genes was found to be shared between human and pig intestinal microbiota. The mobile genetic context for 11 of these core antibiotic resistance genes could be identified by mining their carrying scaffolds constructed from the 2 data sets, leading to the detection of 7 integrative and conjugative/mobilizable elements and 2 insertion sequence (IS)-related transposons. The comparison of the relative abundances between these detected mobile genetic elements and their associated antibiotic resistance genes revealed that for many genes, the estimated contribution of the mobile elements to the gene abundance differs strikingly depending on the host. These findings indicate that although some antibiotic resistance genes are ubiquitous across microbiota of human and pig populations, they probably rely on different genetic elements for their dissemination within each population.

IMPORTANCE There is growing concern that antibiotic resistance genes could spread from the husbandry environment to human pathogens through dissemination mediated by mobile genetic elements. In this study, we investigated the contribution of mobile genetic elements to the abundance of highly prevalent antibiotic resistance genes found in commensal bacteria of both human and pig intestinal microbiota originating from the same region. Our results reveal that for most of these antibiotic resistance genes, their abundance is not explained by the same mobile genetic element in each host, suggesting that the human and pig microbial communities promoted a different set of mobile genetic carriers for the same antibiotic resistance genes. These results deepen our understanding of the dissemination of antibiotic resistance genes among and between human and pig gut microbiota.

KEYWORDS antibiotic resistance genes, human and pig microbiota, mobile genetic elements

The problem of bacterial antibiotic resistance is now considered a major health issue worldwide and is directly related to the overuse of antibiotics to treat and prevent disease in humans and livestock animals. Most bacteria can acquire antibiotic resistance genes (ARGs) by horizontal gene transfer. This is mediated by the exchange of mobile genetic elements (MGEs) between bacteria (1). ARGs found on MGEs in human
pathogens are ascribed the highest risk to human health in the risk assessment scheme proposed by Martinez et al. (2). MGEs accelerate the spread of ARGs across different hosts and environments, as shown by the colistin resistance gene mcr-1 and the tigecycline resistance gene tet(X), which both quickly disseminated in farm animals and hospitalized patients because of their MGE localization (3–5). Since the wide dissemination of antibiotic resistance genes between animals and humans is a major potential issue both for the livestock environment and human health, it is a primary necessity to better understand the ARG flow among the different hosts.

As a known reservoir for ARGs (6–8) and a hot spot for horizontal gene transfer (9, 10), the intestines of humans and domestic animals are believed to mediate the spread of ARGs from commensal organisms to pathogens (9, 11). Human and farm animal gut microbiota have also been shown to share a large number of ARGs, suggesting a gene flux between these two ecosystems (8, 12). The pool of MGEs present in the microbiota, which forms the mobilome, is assumed to facilitate the transfer of ARGs among members of the intestine microbial community (13–15). Although several ARG-carrying MGEs have been reported in the intestinal microbiota of both human and animals (9, 13, 14, 16), to which extent they are shared between hosts has never been thoroughly investigated, and the detailed contribution of MGEs to ARGs dissemination between hosts has not yet been studied. This missing information has limited our understanding of the flow of ARGs between humans and domestic animals and prevented us from objectively assessing the role of the livestock practices in the spread of antibiotic resistance.

In this study, we took advantage of the available metagenomic sequences used for the construction of the human and pig gut microbiome reference gene catalogues and for which a general survey of resistance gene content was performed (17, 18). We reevaluated the prevalence of ARGs in these human and pig gut microbiota by focusing on samples coming from the same region and identified 27 ARGs shared by the intestinal microbiota of most humans and pigs. MGEs carrying these ARGs were extracted from the metagenomic assemblies, and their relative abundances were compared to those of their embedded ARGs. Our analyses showed that shared ARGs were mostly carried by integrative and conjugative/mobilizable elements (ICE/IMEs) and that the host has a strong influence on the type of MGE carrying a given ARG. These results indicate that, even for the shared ARGs, the antibiotic resistance MGE’s main carriers are mostly disjointed between pig and human gut microbiota. Our approach provides a unique perspective for reassessing ARG flow between these two ecosystems.

RESULTS AND DISCUSSION

Determination of the shared core resistome of human and pig gut microbiota.

To investigate which ARGs are shared between human and pig gut microbiota, we analyzed the relative abundance of ARGs in metagenomics data associated with 145 human and 87 pig intestine samples collected in Shenzhen, Guangdong Province, by mapping the reads on the Comprehensive Antibiotic Resistance Database. Human samples came from an adult cohort composed of 71 patients with type 2 diabetes and 74 control individuals, none of which received antibiotics in the 2 months before sample collection (17). Pig samples came from adult sows belonging to 6 different breeds, and all were reared indoors in the same facility with a diet supplemented with low doses of fluorobenzene and penicillin (18). ARGs detected in these samples encompassed genes encoding three major resistance mechanisms—efflux pumps, antibiotic inactivation, and target protection—and which confer resistance to most major classes of antibiotics. Among the 145 human samples, 360 ARGs were detected in at least one sample (see Table S1 in the supplemental material). Of these samples, 67 ARGs were shared by $\geq 80\%$ of the human samples with a relative abundance (compared with the total 16S rRNA gene abundance) of $\leq 10^{-6}$. This core resistome accounted for 40.85% to 98.40% of the total ARG abundance in human samples (see Fig. S1A and Table S2 in the supplemental material). The total and core ARG contents were highly similar.
between patients with diabetes and control individuals (see Fig. S2 in the supplemental material), indicating that type 2 diabetes has no measurable effect on the human gut resistome. ARGs in the 87 pig samples were more diverse; 572 ARGs were detected in at least one sample (Table S1). The core resistome of the pig samples was composed of 78 ARGs, accounting for 80.56% to 99.79% of the ARG abundance in each pig sample (Fig. S1B; Table S2). The higher diversity of the total resistome in pig samples than that in human samples is probably linked to the intake of antibiotics by the sampled sows, although the lack of details on the administered drugs prevented us from investigating this hypothesis further. On the contrary, the higher richness of the core resistome in pigs than that in humans may be the result of the lower number of samples, as the core gene content is known to decrease with the investigated sample size (19). In addition, sampled pigs all originated from the same farm which also probably influenced the core resistome, due to the known high relatedness in ARG content between individuals from the same herd (20).

Twenty-seven ARGs were present in the core resistome of both human and pig intestine samples (Fig. 1). In this shared core resistome, resistance genes for tetracycline and macrolide-lincosamide-streptogramin B (MLS\textsubscript{B}) were especially abundant, summing up for 59% and 60% of the total ARG abundance in human and pig samples, respectively. Specifically, the 8 core shared tetracycline resistance genes \textit{tet}(Q), \textit{tet}(O), \textit{tet}(W), \textit{tet}(32), \textit{tet}(40), \textit{tet}(M), \textit{tet}(44), \textit{tet}(S), as well as the MLS\textsubscript{B} resistance genes \textit{erm}(F) and \textit{erm}(B), were detected in virtually all human and pig samples (Table S2). Interestingly, two genes, namely, \textit{vanR} and \textit{vanR}G, are designated glycopeptide (vancomycin) resistance genes. However, they encode only the regulatory proteins of the \textit{vanR} and \textit{vanR}G operons (21, 22), and the other genes of these operons are found only sporadically in human or pig gut resistomes (Table S1). Thus, the detection of these two genes is probably the result of nonspecific mapping on other regulatory protein-encoding genes of the microbiome and should not be regarded as true vancomycin resistance genes. None of the other 25 shared core ARGs provided resistance to critically important antibiotics for human health (namely, third- and fourth-generation

**FIG 1** Heat map of the relative abundance of core ARGs shared by 145 human samples and 87 pig samples. Samples (columns) were clustered by the clustering method ward.D in R. Lines represent ARGs present in at least 80% of individuals in both human and pig data sets. The color range indicates the log\textsubscript{10}-transformed ARG abundance relative to the 16S rRNA gene abundance.
cephalosporins, fluoroquinolones, vancomycin, tigecycline, and colistin). Most of these threatening ARGs were from Gram-negative bacteria, mainly from the Gammaproteobacteria class, and these bacteria are usually present in low abundance in mammalian gut microbiota. The limited resolution of the metagenomic sequencing may thus raise concerns about a possible undersampling of these major carriers. However, other resistance genes commonly found within the Gammaproteobacteria, such as sul1, sul2, or floR, were part of the core resistome from either human or pig samples (Fig. S1), indicating that the low abundance of these bacteria may not be a limiting factor for the detection of widely disseminated ARGs. Accordingly, the core shared resistome reported here is in perfect agreement with that originally observed on these metagenomic data sets, including from nonmedicated pigs (18) or restricted to the nondiabetic human cohort (6). Moreover, 11 of the 27 genes (including almost all the tetracycline and MLSB resistance genes) were previously reported as core and shared between human, cattle, pig, and chicken gut microbiota (12). Fifteen genes were also found to be shared by pig and chicken microbiota in Europe (20), suggesting a worldwide, panspecies dissemination of these core ARGs. Although these 27 ARGs were prevalent in the gut microbiota of humans and pigs, their relative abundances differed significantly between hosts, leading to a perfect separation of pig and human samples according to their ARG profile (Fig. 1). The relative abundance of tet(Q), tet(S), erm(F), erm(B), erm(35), macB, blaoXA-47, aac(6’)-Ie-aph(2’)-Ia, dfrE, dfrF, vanRG, and taeA was significantly higher in human samples than in pig samples (P < 0.05; see Table S3 on the supplemental material). On the contrary, the relative abundance of tet(O), tet(W), tet(32), tet(40), tet(44), lnu(C), mefA, ant(6’)-Ib, vanB, sav1866, and IseF was higher in pig samples than in human samples (P < 0.05; Table S3). The relative abundance of tet(M), cfsA5, mel, and efB8 was similar in both intestine microbiota (P > 0.05; Table S3). The factors responsible for these differences are not known but may be related to species-specific antibiotic consumption (23, 24) or gut microbial composition (25).

**Shared ARGs were found on diverse MGE structures.** The presence of 27 ARGs shared between most human and pig gut microbiota suggests that these genes were efficiently disseminated among and between the 2 ecosystems. To test whether this dissemination was linked to specific MGEs, we looked for MGE genetic hallmarks in the scaffolds that carried the 27 shared core ARGs in the assembled data sets of the 145 human and 87 pig samples (see Materials and Methods). Eleven ARGs were found in genetic contexts related to known MGEs. Nine of them were putatively located in ICE/IMEs, while the other two were connected to insertion sequence (IS)-based transposons (Table 1). ICE/IMEs are known to be important vectors of ARG transfer in Bacteroidetes and Firmicutes (26–28), which collectively represent the vast majority of the metagenomic data sets investigated here (see Fig. S3 in the supplemental material). Such bacterial community distribution is perfectly in line with the established intestinal flora of human and pigs (29, 30); it is therefore reasonable to deduce that ICE/IMEs may play a primary role in the dissemination of ARGs in these ecosystems.

<table>
<thead>
<tr>
<th>MGE type</th>
<th>Inferred MGEs</th>
<th>Family</th>
<th>Carried ARG(s)</th>
<th>No. of nonredundant scaffolds</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICE/IME</td>
<td>Tn916</td>
<td>Tn916</td>
<td>tet(M)</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>CMGE_youdozi10</td>
<td>Unclassified</td>
<td>tet(O), tet(40)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>ICEB_XB1A</td>
<td>CTnDOT/ERL</td>
<td>tet(Q), erm(F)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>CTnBST</td>
<td>Unclassified</td>
<td>blaOXA-347</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>CTnHyb</td>
<td>Unclassified</td>
<td>cfxA5</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Tn4555</td>
<td>Unclassified</td>
<td>cfxA5</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>CTnGERM1</td>
<td>Unclassified</td>
<td>mel, mefA</td>
<td>5</td>
</tr>
<tr>
<td>IS</td>
<td>Tn4001</td>
<td>IS256</td>
<td>aac(6’)-Ie-aph(2’)-Ia</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>MTnSag1</td>
<td>ISSag10</td>
<td>lnu(C)</td>
<td>27</td>
</tr>
</tbody>
</table>

**TABLE 1 Detected ARGs-carrying MGEs**

Wang et al. Applied and Environmental Microbiology March 2021 Volume 87 Issue 5 e01910-20 aem.asm.org
The lack of association with MGEs for the 16 remaining ARGs may result from technical limitations. First, MGEs or even ARGs may not be properly assembled in the metagenomic assemblies, due to the concurrent presence of several variants in the same sample, as previously shown for the tet(O) and tet(W) genes (31). These ARGs may also be part of unknown mobile elements. It has been shown that pig and human gut microbiota harbor a diversity of not-yet-described MGEs carrying known resistance genes (14, 16). In this case, our detection procedure based on high similarity with known MGEs would miss these associations. A last, biological possibility is that some of these ARGs may actually not be located within an MGE but may instead be part of the intrinsic resistome of core members of the gut bacteria. Since all the genes analyzed here were defined as mobile in the literature, such a hypothesis would assume that the aforementioned gut bacteria were the source of the mobilized version of the ARG.

Among the 11 ARGs detected on scaffolds harboring ICE/IME hallmarks, 4 were tetracycline resistance genes. tet(Q), the most abundant ARG, was found on four scaffolds showing 91% to 96% nucleotide similarity with a part of ICEBxyXB1A (32), an ICE belonging to the CTnDOT/ERL family (Fig. 2A). These scaffolds also contained the mobA, mobB, mobC, and rteA8 operons as well as rteC, which are all known to be essential for the CTnDOT mobility (33, 34). Interestingly, two of the three scaffolds also carried the shared core MLS\(_{b}\) resistance gene erm(F) directly upstream of tet(Q), a configuration not reported previously (Fig. 2A). CTnDOT/ERL-related ICEs frequently carry the erm(F) gene but are usually not directly adjacent to the tet(Q) gene (35). This arrangement thus represents a candidate new variant of the CTnDOT/ERL family. It enabled this family to acquire the erm(F) without insertion of other MGEs. Two other tetracycline resistance genes, tet(O) and tet(40), were detected in tandem in one scaffold related to the tnpV-sigB element that we reported in a previous study (14). This arrangement is 97% identical at the nucleotide level to a region of CMGE\(_{YOK0816}\) (36), an ICE of unknown family described in Streptococcus suis (Fig. 2B). Another scaffold showed high similarity to CMGE\(_{YOK0816}\) over an additional 4.54 Kbp downstream of the tnpV-sigB element but was lacking the tet(O) gene, raising some doubts about whether tet(O) and tet(40) are always found in tandem on this ICE. The last tetracycline resistance gene, tet(M), was found on nine scaffolds, all of which having high nucleotide sequence similarity to Tn916 elements (three are shown in Fig. 2C), an ICE well known for disseminating the tet(M) gene in diverse commensal and pathogenic Gram-positive bacteria (37). Flanking regions for the Tn916 element in the retrieved scaffolds were diverse, in agreement with the high transfer capacity of this ICE among bacteria.

Among the four remaining ARGs connected to known ICE/IMEs, two were beta-lactam resistance genes. The first one, bla\(_{OXA-347}\), was carried by two nonredundant scaffolds, with both being more than 98% similar on 90% of their length to the ICE CTnBST originally detected in Bacteroides uniformis WH207 (38). Interestingly, bla\(_{OXA-347}\) was not originally reported in CTnBST and was not located in the region matching the ICE reference sequence but instead was in the 2,210-bp downstream region, which also includes an IS\(_{Bbi1}\) transposase and a DNA ligase gene (Fig. 2D). The aforementioned new structure may suggest that the bla\(_{OXA-347}\) gene was integrated into CTnBST through its adjacent IS element, although such an assumption needs to be confirmed by experimental evidence. The other beta-lactamase gene, cfxA5, was located in two unrelated MGEs as follows: Tn4555 (Fig. 2E), a widely distributed IME known to promote cfxA dissemination in Bacteroides (39, 40) and Parabacteroides genera (41); and a region related to CTnHyb, an ICE originally found in a multidrug-resistant Bacteroides fragilis clinical isolate (42) (Fig. 2F). The CTnHyb element was originally reported as a hybrid ICE carrying Gram-positive related genes conferring resistance to tetracycline, kanamycin, metronidazole, and spectinomycin but not carrying cfxA5. The new variant presented here further confirms the highly mosaic nature of this Bacteroidetes MGE.

The last two ARGs connected to ICE/IMEs were the MLS\(_{b}\) resistance genes mel and mefA, which were both present in a structure with 99.79% nucleotide sequence similarity to CTnGERM1 (43) of Bacteroides ovatus (Fig. 2G, three of the five scaffolds are
FIG 2 Genomic organization of ARG-carrying scaffolds showing homology with known ICEs/IMEs, transposons, and IS elements and comparison with related sequences available in GenBank. Coding DNA sequences (CDSs) are colored according to the function of their encoded protein, as follows: purple, antibiotic resistance; blue, integrases or recombinases; brown, transposases; light green, other known functions; and white, hypothetical proteins. Vertical red bars represent mobile element inverted repeats. Gray shading represents similarity at the nucleotide level. Acc., accession number.
shown). Surprisingly, all the scaffolds include \(erm(G)\) in addition to \(mel\) and \(mefA\), a gene also present in the reference sequence of CTnGERM1. The fact that \(erm(G)\) did not come out as a shared core resistance gene in our analysis suggests that the three genes are not always linked. Our results showed that the \(erm(G)\) region of CTnGERM1 seems to be able to transfer as an independent cassette and is not always part of the CTnGERM1 element.

The two ARGs which could be attributed to IS-based transposons were the lincosamide resistance gene \(Inu(C)\) and the aminoglycoside resistance gene \(aac(6')-le-aph(2')-la. \(Inu(C)\) was found in 27 scaffolds always adjacent to \(I\Sigma\)Sag10 (Fig. 2H, three scaffolds are shown), with which it forms the transposon MTnSag1 known to be mobilized by Tn916 transposons and found in numerous human- and gut-associated bacteria (44). The \(aac(6')-le-aph(2')-la\) gene, which encodes a bifunctional enzyme, was found in three scaffolds with 99% to 100% nucleotide similarity to Tn4001, an IS256-based composite transposon (45) (Fig. 2I). On one scaffold, the upstream regions of IS256R included a region sharing 100% nucleotide sequence identity to the plasmid replication protein (\(repE\)) of the Enterococcus faecalis VS83 plasmid pTEF1 (46). However, this homology did not extend in the downstream region of the scaffold nor is it present on the other scaffolds, suggesting that the pTEF1-like plasmid may not be the main carrier of \(aac(6')-le-aph(2')-la\) in the investigated data sets.

**The resistance MGEs contributed differently to their associated ARGs dissemination in human and pig microbiota.** To evaluate the contribution of the above-described MGEs to the prevalence of their associated ARGs in human and pig populations, we investigated the integrity level (completeness) of the ARG-carrying structures in the metagenomic samples by calculating their mapping coverage independently in each sample (Fig. 3A; see Fig. S4 in the supplemental material). Next, we calculated the mapping depth ratio between each ARG and its corresponding MGE-related region to have an estimate of the relative contribution of each MGE on the abundance of ARGs in the two data sets (see equation 2) (Fig. 3B).

When these two statistics were combined, several patterns appeared. A first group of MGE-related regions was characterized by a very high integrity level (>90% coverage) in almost all samples, indicating that these MGEs were part of the core mobilome of pig and human gut microbiota (Fig. 3A). They consisted of MTnSag1 carrying \(Inu(C)\) and of the two identified \(CMGE_{Y060816}\) parts, with one corresponding to the \(tnpV-sigB\) element carrying \(tet(40)\) and \(tet(O)\) and the other corresponding to the scaffold carrying only \(tet(40)\) (scaffold21363_1 in Fig. 2B), here called the \(sigB-parB\) region. Such structures have been reported in microbes from pig intestines (14, 36), but we reported here that they were core elements of both human and pig gut microbiota. Interestingly, while MTnSag1 contributed significantly to the \(Inu(C)\) abundance in both human and pig gut microbiota, the \(CMGE_{Y060816}\) contribution to the \(tet\) gene abundance depended on the investigated region. Both the \(tnpV-sigB\) and the \(sigB-parB\) regions highly contributed to the \(tet(40)\) abundance in human samples, while only the \(tnpV-sigB\) region contributed significantly to the \(tet(40)\) abundance in pig samples (Fig. 3B). On the contrary, the \(tnpV-sigB\) region highly contributed to the \(tet(O)\) abundance in pig but not in human samples. These observations suggested that the \(tnpV-sigB\) element disconnected of \(CMGE_{Y060816}\) was the main contributor for \(tet(40)\) and \(tet(O)\) abundance in pig microbiota, which indicated that this small structure is distributed in a variety of genetic contexts in pig gut microbes and that \(tet(O)\), but not \(tet(40)\), was likely present in genetic contexts other than \(CMGE_{Y060816}\) in human gut microbiota.

A second group of MGE-related regions, including the central regions of ICEBxyXB1A (carrying \(tet(Q)\) and \(erm(F)\)) and of CTnBst (carrying \(bla_{OXA-347}\)), was characterized by a very high integrity level in almost all human samples, while these MGEs were virtually absent from the pig samples (Fig. 3A). These MGEs are primarily found in Bacteroidetes (32, 38); their preferred association with the human samples is therefore consistent with the prevalence of this bacterial lineage in the investigated data set (Fig. 3S). It is worth noting that the relative abundances of \(tet(Q)\) and \(erm(F)\) were very high in pig samples (Fig. 1 and Fig. S4D), suggesting that these ARGs were likely mostly located on other, unrelated genetic
structures in pig gut microbiota. Moreover, even if these regions were almost complete in all human samples, the contribution to their respective ARGs \([\text{erm}(F), \text{tet}(Q), \text{bla}_{OXA-347}]\) was extremely variable among samples but generally lower than 75% (Fig. 3B). Such results indicated that these resistance genes were likely located in other genetic contexts in the human samples as well and that the above MGEs were not the major contributors of the \([\text{erm}(F), \text{tet}(Q), \text{bla}_{OXA-347}]\) spread in the investigated microbiota.

The third group of MGE-related regions showed the opposite trend, with a high or perfect integrity in all pig samples but only in very few human samples. One of these scaffolds corresponded to the CTnGERM1 region carrying \(\text{mefA}\) and \(\text{mel}\) which highly contributed to the \(\text{mel}\) abundance but not to the \(\text{mefA}\) abundance in pig samples (Fig. 3B and S4F). Interestingly, in the few human samples where the CTnGERM1 region was completely detected, its abundance almost totally explained both \(\text{mel}\) and \(\text{mefA}\) abundances. In streptococci, the \(\text{mefA}\) and \(\text{mel}\) genes were part of a two-component efflux pump and were always found together (47). It was thus intriguing that CTnGERM1 explains only the \(\text{mel}\) abundance in pig samples, suggesting that \(\text{mefA}\) may be present without \(\text{mel}\) in another genomic context(s) in pig gut microbiota. In addition to the CTnGERM1 region, the two scaffolds corresponding to the \(\text{cfxA5}\)-carrying regions of \(\text{Tn4555}\) and of CTnHyb showed a complete integrity in all pig samples but only in half and in 1 of the 145 human samples. The case of CTnHyb was interesting because the mapping revealed the existence of two main variants (Fig. S4H). The first variant was

---

**FIG 3** The integrity level of detected MGEs and their contribution to their associated ARG abundance. (A) Distribution of the integrity degree of each MGE among samples, for human (red) and pig (blue) data sets. Integrity is displayed in degree of completeness (0% to 24%, 25% to 49%, 50% to 69%, 70% to 89%, and 90% to 100%), with shading corresponding to the proportion of samples within each range. (B) Contribution of each MGE to the abundance of the ARGs they carry (see Materials and Methods). Colors as described for panel A. The size of the bubble is proportional to the number of samples in which a given MGE has the same contribution level. For each MGE, only samples showing an integrity level \(\geq 90\)% were considered.
100% identical to the mapped scaffold and was found in all pig samples but only once in human samples (Fig. 3A). The second variant lacked some regions, including the region where cfxA5 was located, and was found in both human and pig samples. Moreover, the lacking fragments in this second variant matched exactly the regions from the scaffold not found on the CtnHyb reference sequence (Fig. S4H). This result suggested that the published cfxA5-free CtnHyb was the original element found in both human and pig microbiota and that the cfxA5-carrying CtnHyb successfully spread in the pig gut microbiota after its cfxA5 acquisition. Accordingly, CtnHyb was the major contributor of the cfxA5 abundance in pig samples (Fig. 3B). On the other hand, the abundance of cfxA5 in human microbiota was mostly explained by the presence of Tn4555 (when present) rather than CtnHyb (Fig. 3B), consistent with the presence of only the cfxA5-free variants of CtnHyb in human samples.

Finally, a last group of MGE-related regions included those corresponding to Tn916 [carrying tet(M)] and Tn4001 [carrying aac(6’)-le-aph(2’)-la], which showed a variable integrity level among pig and human samples, although they tended to be more complete in human than in pig samples (>90% coverage in >50% of the human samples but only in <50% of the pig samples). When present, Tn916 highly contributed to the tet(M) abundance in human samples (61/75 of the samples show a contribution higher than 80%). On the contrary, for the 8 pig microbiota with a totally detected Tn916, only 1 provides a major contribution to the tet(M) abundance. Tn916 was thus an important contributor of tet(M) dissemination in human but not in pig gut microbiota. Finally, the contribution of Tn4001 to the aac(6’)-le-aph(2’)-la abundance did not show any clear pattern of association among samples, suggesting that this gene was present in various genetic contexts in human and pig gut bacterial communities. Several studies indicated that some variants of Tn4001 were commonly found in Gram-positive bacteria, notably in diverse gut-associated enterococci (48–50). These variants either lacked the 5’, 3’, or both IS256 elements, or might still harbor partial IS256 regions downstream of the aac(6’)-le-aph(2’)-la gene (48). The possible presence of these various structures in our samples might have blurred the mapping estimates on the IS256, which could have in turn impacted our calculation of Tn4001 contributions on the aac (6’)-le-aph(2’)-la abundances.

Conclusions. Metagenomic sequencing data have been extensively used to get insights into ARGs and MGE prevalence and abundance in various environments, but very few attempts have been made to link the putatively mobile ARGs to their respective MGEs. Here, we analyzed ARGs and MGEs by deep mining the metagenomes of human and pig fecal samples collected in Shenzhen, Guangdong Province, China. Unlike in previous studies, here we investigated a widely distributed subset of all detected ARGs, likely harbored by the commensal flora, and focused on the MGEs carrying them. Our analysis revealed that most of the investigated ARG-carrying MGEs made different contributions to their associated ARG abundance in human and pig samples, with no systematic links with the prevalence of the MGE in these ecosystems. At a more general point of view, such findings suggest that ARGs shared by human and pig microbiota may have been mobilized independently or that ecological specificity favored the dissemination of different genetic vehicles within each ecosystem. Our results also support the recent view that the gut microbiota may not be an ARG reservoir for bacterial pathogens as important as previously assumed. Tn916 and the CtnDOT-like ICEBxyXB1A, which are well-known MGEs widely distributed in pathogens, were found only occasionally within microbiota (especially in pigs) or, when detected, contributed moderately to their associated ARG abundance. On the other hand, CMGE_YY00815r, the tnpV-sigB element and the MTnsag1 transposon, which collectively carry tet(40), tet(O), and lnu(C), were found in high abundance in almost 100% of the samples of both microbiota while being only recently characterized and rarely reported in pathogens. In summary, the findings presented here indicate that although numerous antibiotic resistance genes are ubiquitous in human and animal gut microbiota, their widespread distribution seems to not be connected to specific MGEs or to

Distinct MGE Mediated the Transfer of Shared Resistome Applied and Environmental Microbiology

March 2021 Volume 87 Issue 5 e01910-20 aem.asm.org 9
MGEs usually found in bacterial pathogens. Whether these conclusions hold true for resistance genes that are less widely distributed but more critical, such as the extended-spectrum and metallo-\(\beta\)-lactamases, deserves to be investigated to fully understand the exact role of intensive antibiotic use in husbandry in the development of the human gut resistome.

MATERIALS AND METHODS

Data collection and bacterial diversity analysis. We selected and downloaded the raw sequence data of 145 human samples from the NCBI Sequence Read Archive database (accession no. SRA045646) and 87 pig fecal samples from the European Nucleotide Archive (PRJEB11755) to perform an in-depth analysis of their ARG repertoire and associated MGEs. These two sample data sets were both isolated in Shen Zhen, China, and sequenced using either an Illumina GAIIx or a HiSeq 2000 platform, producing paired-end reads of 75 to 90 bp and 100 bp for an average amount of raw reads per sample of 2.61 Gbp and 5.76 Gbp for human and pig data sets, respectively (17, 18).

Bacterial taxonomic information for the pig samples was downloaded from the GigaDB database at the class and phylum level. Since such information was not available for the human data set, raw paired-end reads of each human sample were analyzed with the Kraken2 software (51) using the parameters --report and --use-names. The resulting taxonomic diversity reports were combined and summarized with an in-house perl script, using the taxonomic classification given by Kraken2.

Calculation of ARG and identification of the resistome. We assessed the relative abundance of ARGs according to the method described by Li et al. (52), which considers the sequence lengths of the reference ARGs and normalizes them by the 16S rRNA gene sequence length and abundance. All clean reads from the 145 human samples and 87 pig samples were aligned against the sequences of all 4,080 antibiotic resistance ontologies downloaded from the Comprehensive Antibiotic Resistance Database (CARD v3.0.1) (53), using BLASTX with the parameters -E value 1e-5 and -perc_identity 90. Alignments shorter than 20 amino acids were discarded. Then, the mapping results were normalized by the length and read numbers of 16S RNA genes according to the method described by Li et al. (52), using the following equation:

\[
\text{Relative Abundance}_{\text{ARG}} = \sum_{i=1}^{n} \frac{N_{\text{ARG-like sequence}} \times L_{\text{reads}}}{N_{16S \text{ sequence}} \times L_{16S \text{ sequence}}}
\]

where \(N_{\text{ARG-like sequence}}\) represents the number of reads with high similarity to the ARG reference sequences, \(L_{\text{ARG reference sequence}}\) represents the length of the corresponding ARG, \(N_{16S \text{ sequence}}\) is the number of 16S rRNA sequence reads identified from the metagenomic data (52), and \(L_{16S \text{ sequence}}\) is the length of the reads and is equal to 101 bp in this study; \(L_{16S \text{ sequence}}\) is the average length of the 16S rRNA sequence and was set to 1,432 bp.

We identified the core resistome of each human and pig data sets as genes with a relative abundance of \(\geq 10^{-4}\) in at least 80% of human and pig samples, respectively. There were 27 ARGs shared between the human and pig core resistomes.

Identification of ARG-carrying structures. First, the human metagenomic assembly for the corresponding samples were downloaded from the GigaDB database (http://gigadb.org/dataset/view/id/100036). Since no assembly was publicly available for the pig data set, a metagenomic assembly was produced locally from the pooled pig raw reads using the SOAPdenovo assembler v2.04 (54) with the kmer size set to 31. The human and pig scaffolds were then aligned to the sequence of the core ARGs. A scaffold was considered an ARG-carrying scaffold when it shared >80% identity with an ARG gene over >70% of the ARG length. In addition, ARG-carrying scaffolds of <500 bp were discarded. To reduce redundancy, scaffolds that totally overlapped longer scaffolds were also discarded. A total of 288 non-redundant scaffolds carrying 18 of the 27 shared core ARGs were identified. No flanking sequence could be found for the remaining nine ARGs. Of the 288 scaffolds, 62 had more than 90% similarities at the nucleic acid sequence with entries from ISFinder (55), ICEberg (56), and GenBank plasmid database, and together carried 11 of the 27 core ARGs. The other 226 scaffolds did not show high similarity to any known mobile element and were thus not considered for further analyses. The ARG-carrying scaffolds were annotated using Prokka v1.13.2 (57), and each encoded protein sequence was compared to the ARG reference sequence downloaded from the Comprehensive Antibiotic Resistance Database (CARD v3.0.1) (53), using BLASTX with the parameters -E value 1e-5 and -perc_identity 90. Alignments shorter than 20 amino acids were discarded. Then, the mapping results were normalized by the length and read numbers of 16S RNA genes according to the method described by Li et al. (52), using the following equation:

\[
\text{Relative Abundance}_{\text{ARG}} = \sum_{i=1}^{n} \frac{N_{\text{ARG-like sequence}} \times L_{\text{reads}}}{N_{16S \text{ sequence}} \times L_{16S \text{ sequence}}}
\]

where \(N_{\text{ARG-like sequence}}\) represents the number of reads with high similarity to the ARG reference sequences, \(L_{\text{ARG reference sequence}}\) represents the length of the corresponding ARG, \(N_{16S \text{ sequence}}\) is the number of 16S rRNA sequence reads identified from the metagenomic data (52), and \(L_{16S \text{ sequence}}\) is the length of the reads and is equal to 101 bp in this study; \(L_{16S \text{ sequence}}\) is the average length of the 16S rRNA sequence and was set to 1,432 bp.

To assess the importance of MGEs detected in the scaffolds to the dissemination of ARGs in pig and human microbiota, we used the ratio in mapping depth between these MGEs and their associated ARGs as a proxy for the contribution of the MGE to the ARG abundance in each sample. The sequencing depth of different parts of a continuous genetic region is theoretically consistent, and an MGE totally contributing to a given ARG abundance should have a mapping depth at...
Distinct MGE Mediated the Transfer of Shared Resisitome


de least equal to that of the ARG. We calculated the mapping depth ratio (MDR) between each MGE and its corresponding ARG(s), for each sample as follows:

\[ MDR = \frac{MMD_{MGE}}{MMD_{ARG}} \times 100 \]  

(2)

where MMD_{MGE} and MMD_{ARG} represent the median mapping depth of each ARG and of its flanking MGE region (excluding the ARG of interest), respectively.

**Data availability.** Raw read sequencing data for 145 human samples and pig fecal samples are available under the NCBI Sequence Read Archive database (accession no. SRA045646) and European Nucleotide Archive (PRJEB11755), respectively (17, 18). The detailed information has been released at GigADB online at http://gigadb.org/dataset/view/id/100036 and http://gigadb.org/dataset/view/id/100187 (17, 18). All data sets used in this study are listed in Table S1.

**SUPPLEMENTAL MATERIAL**

Supplemental material is available online only.

- **SUPPLEMENTAL FILE 1**, XLSX file, 0.9 MB.
- **SUPPLEMENTAL FILE 2**, XLSX file, 0.3 MB.
- **SUPPLEMENTAL FILE 3**, XLSX file, 0.01 MB.
- **SUPPLEMENTAL FILE 4**, PDF file, 1.5 MB.

**ACKNOWLEDGMENTS**

This work, including the efforts of C.W., was funded by National Natural Science Foundation of China (NSFC) (31872632). This work, including the efforts of J.F., was funded by National Natural Science Foundation of China (NSFC) (31861133001 and 31870134).

We thank the anonymous reviewers which helped us to greatly improve the quality of the manuscript.

**REFERENCES**


