



Antibiotic Resistance Gene Diversity and Virulence Gene Diversity Are Correlated in Human Gut and Environmental Microbiomes

Pedro Escudeiro,^{a*}  Joël Pothier,^b  Francisco Dionisio,^a  Teresa Nogueira^a

^aCE3c—Centro de Ecologia, Evolução e Alterações Ambientais, Faculdade de Ciências, Universidade de Lisboa, Lisbon, Portugal

^bInstitut de Systématique, Evolution, Biodiversité (ISYEB), Muséum National d'Histoire Naturelle, CNRS, Sorbonne Université, EPHE, Paris, France

ABSTRACT Human beings have used large amounts of antibiotics, not only in medical contexts but also, for example, as growth factors in agriculture and livestock, resulting in the contamination of the environment. Even when pathogenic bacteria are the targets of antibiotics, hundreds of nonpathogenic bacterial species are affected as well. Therefore, both pathogenic and nonpathogenic bacteria have gradually become resistant to antibiotics. We tested whether there is still cooccurrence of resistance and virulence determinants. We performed a comparative study of environmental and human gut metagenomes from different individuals and from distinct human populations across the world. We found a great diversity of antibiotic resistance determinants (AR diversity [ARd]) and virulence factors (VF diversity [VFd]) in metagenomes. Importantly there is a correlation between ARd and VFd, even after correcting for protein family richness. In the human gut, there are less ARd and VFd than in more diversified environments, and yet correlations between the ARd and VFd are stronger. They can vary from very high in Malawi, where antibiotic consumption is unattended, to nonexistent in the uncontacted Amerindian population. We conclude that there is cooccurrence of resistance and virulence determinants in human gut microbiomes, suggesting a possible coselective mechanism.

IMPORTANCE Every year, thousands of tons of antibiotics are used, not only in human and animal health but also as growth promoters in livestock. Consequently, during the last 75 years, antibiotic-resistant bacterial strains have been selected in human and environmental microbial communities. This implies that, even when pathogenic bacteria are the targets of antibiotics, hundreds of nonpathogenic bacterial species are also affected. Here, we performed a comparative study of environmental and human gut microbial communities issuing from different individuals and from distinct human populations across the world. We found that antibiotic resistance and pathogenicity are correlated and speculate that, by selecting for resistant bacteria, we may be selecting for more virulent strains as a side effect of antimicrobial therapy.

KEYWORDS antibiotic resistance, cooccurrence, human gut, metagenomics, microbiome, virulence

Antibiotics are present in microbial communities, not only as a result of the natural lifestyle of microorganisms but also due to the usage of these drugs in agriculture, food industry, livestock, or in human health (1). Therefore, antibiotics can affect bacterial communities as a whole, comprising both pathogenic and nonpathogenic bacteria. Take, for example, the human microbiome, defined as the set of microorganisms that colonize humans. This microbiome is composed of about 3.8×10^{13} bacterial cells (2) spanning thousands of taxa and colonizing our body's surfaces and biofluids,

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

Editor Paul D. Fey, University of Nebraska Medical Center

Copyright © 2019 Escudeiro 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 Teresa Nogueira, teresainogueira@gmail.com.

* Present address: Pedro Escudeiro, Microbiology & Biotechnology Group, Biosystems & Integrative Sciences Institute (BioSI), Faculdade de Ciências da Universidade de Lisboa, Lisbon, Portugal.

P.E. and J.P. contributed equally to this article.

Received 26 February 2019

Accepted 11 April 2019

Published 1 May 2019

including tissues such as skin, mucosa, and, most importantly, the gastrointestinal tract. Thus, even when virulent bacteria are the targets of antibiotics, the administration of these drugs may also affect many nonpathogenic mutualistic or commensal bacterial species present in individuals undergoing treatment (3).

The resistome (the collection of all antibiotic resistance [AR] genes, which exist in both pathogenic and nonpathogenic bacteria [4]) is frequently carried on mobile genetic elements. Similarly, the virulome, the set of genes encoding virulence, can also be carried on the mobilome (5–7). Therefore, many bacterial virulence factors (VFs) are easily spread in bacterial populations by horizontal gene transfer, converting mutualistic or commensal bacteria into potential opportunistic pathogens. Several examples of virulent (and highly virulent) and multiresistant disseminated clones have already been reported throughout the literature (8–10; for a review, see reference 11).

A given bacterial population may constitute a kind of life insurance to other bacterial populations present in the same microbiome through at least two different mechanisms. First, if antibiotic resistance genes are carried by plasmids or other mobile genetic elements, they may transfer to pathogenic cells and save them from the negative effects of antibiotics. Indeed, these genetic elements can spread into the bacterial community by horizontal gene transfer, even crossing species (12–14). Curiously, some bacterial strains from different species have been shown to be extremely good donors of certain plasmids. As such, they are able to amplify the number of plasmids in a bacterial community and spread those plasmids to other bacterial cells (15), a phenomenon probably explained by interactions between different plasmids (16, 17).

Second, even bacterial cells not coding for antibiotic resistance determinants nor receiving mobile elements may be protected by cells coding for certain drug resistance genes. Indeed, gene products that inactivate antibiotics by degrading or modifying antibiotic molecules are also decreasing their concentration in the local environment. (For a review of mechanisms of possible indirect resistance, see reference 18.) This mechanism of indirect resistance has been shown to occur in different systems and to be pervasive (19–23).

In this article, we ask whether the abundance of antibiotic resistance protein families correlates with that of virulence among bacterial populations. For this, we have chosen to study metagenomes. There are three main reasons to use metagenomes to address this question. First, it is known that, for millions of years, bacteria had to cope with the presence of many other species, mostly competing with them (24), but also cooperating, both phenomena relying on the full set of genes of the metagenome (24). Second, horizontal gene transfer promotes genetic relationships between species, thus enforcing cooperation (6), avoiding the emergence of cheaters within the microbiome. Third, the study of metagenomes gives us access to the repertoire of genes involved in adaptation to the environment, given that many of these traits are often encoded in the mobilome and thus can be shared by different, eventually unrelated bacteria. In this context, mining for genes coding for antibiotic resistance and virulence traits in metagenomes is a reliable way to access the selective pressures the population is subject to, as well as the cooccurrence of genetic traits of the whole microbiome.

The present study aims at understanding the relationship between antibiotic resistance genes and those coding for virulence. We show that there is indeed a linkage between the dissemination of virulence factors and genes coding for antibiotic resistance, within natural microbiomes, and that this relationship could be influenced by the behaviors of human populations spanning very different geographical locations across the world.

RESULTS

AR gene families in the metagenomes. We started by evaluating the quantity of different antibiotic resistance (AR) proteins present in the metagenomes under study. We wanted to know how does the diversity of antibiotic resistance gene families' homologues (i.e., AR diversity [ARd], the number of different protein families) vary with

the metagenome protein family richness (i.e., the total number of protein families in a metagenome)? To answer this question, we used a data set composed of natural metagenomes issuing from diverse ecosystems and biomes, such as oceans, coral atolls, deep oceans, Antarctic aquatic environments and snow, soils, hypersaline sediments, sludges, microbial fuel cell biofilms, and animal microbial populations (25), and which from now on will be referred to here as environmental metagenomes, in contrast to those composed solely of human gut metagenomes. In this data set, there is a broad variation in AR gene diversity (ARd) (Fig. 1A). However, there is a correlation between ARd and the protein family richness of the various metagenomes (overall correlation of 0.754). In Fig. 1A, regression lines for human gut microbiome (in blue, slope of 0.0037) and for other biomes (in black, slope of 0.0052) are shown. Nevertheless, the *P* value for the difference of these slopes is 0.014, which is significant but not highly significant, hence not allowing us, in our opinion, to draw solid conclusions from this difference.

Another data set was composed of 110 human gut metagenomes sampled from healthy individuals with ages ranging from 0.05 to 53 years, spanning different regions of the world: the United States (USA), Malawi, and Venezuelan Amerindians from Amazonia (26) (Fig. 2A). Metadata can be retrieved at https://bioinfo.mnhn.fr/abi/public/escudeiro-nogueira/Escudeiro_ARd_VFd_human_gut.xlsx.

VF gene families in the metagenomes. Bacterial virulence factors are often proteins that enable pathogenic bacteria to parasitize a host, including gene products involved in adhesion and invasion, secretion, toxins, and iron acquisition systems (27).

The environmental microbiomes of our data set reveal a great diversity of virulence factor (VFd) protein families' densities (Fig. 1B). Since virulence may also be associated with the colonization of different types of biomes (some virulence traits are involved in adaptation to new environments, by allowing bacteria to adhere to and colonize substrates, to access resources such as iron, among other functions) besides the context of infection, one can expect different types of these genes in environmental microbiomes. Moreover, Fig. 1B and 2B show that in human gut microbiomes, VFd versus protein family richness of the metagenomes achieves an overall correlation coefficient of 0.669. Regression lines for human gut microbiome (in blue, slope of 0.0035) and for other biomes (black, slope of 0.0106) are shown. Here the *P* value for the difference of these slopes is 1.7×10^{-8} , indicating that there could be a difference, but the points are widely spread over all types of biome, and we prefer not to draw conclusions about these data (metadata can be retrieved at https://bioinfo.mnhn.fr/abi/public/escudeiro-nogueira/Escudeiro_ARd_VFd_human_gut.xlsx).

AR/VF correlations. The main purpose of the present work is to evaluate the relationship, if any, between ARd and VFd in microbiomes. Therefore, we excluded from our analysis all the gene products that are homologous both to AR and VF determinants. Thus, we avoid the introduction of a bias due to both correlations of AR and VF diversities (ARd and VFd) with metagenome richness. In the analysis, we systematically analyzed the residuals of the regression of ARd on metagenome richness and the residuals of the regression of VFd on metagenome richness. In Fig. 3 one can see that the residuals of ARd and those of VFd are positively correlated. Although there is a large variation of ARd (Fig. 1A), and VFd (Fig. 1B) in environmental metagenomes, there is a strong correlation between ARd and VFd (Fig. 3): the overall partial correlation between ARd and VFd given metagenome richness is 0.84 ($P = 8 \times 10^{-36}$ [i.e., ~ 0.0]). Indeed, the partial correlation between ARd and VFd given the metagenome richness is the same as the correlation between residuals of ARd and residuals of VFd (the partial correlation coefficient for environmental biomes is 0.86, while the one for the human gut is 0.90). When we draw the regression line for the human gut microbiomes from this data set (human gut), we can see that the ratio of residuals of ARd to residuals of VFd is higher, with a slope of 0.63, than for the case of environmental metagenomes, with a slope of 0.44. The difference in slopes is significant, but not very significant, with a *P* value of 0.03. Despite the fact that the values for ARd and VFd are lower in gut metagenomes (Fig. 1), one can witness a higher ARd/VFd ratio (larger slope) of

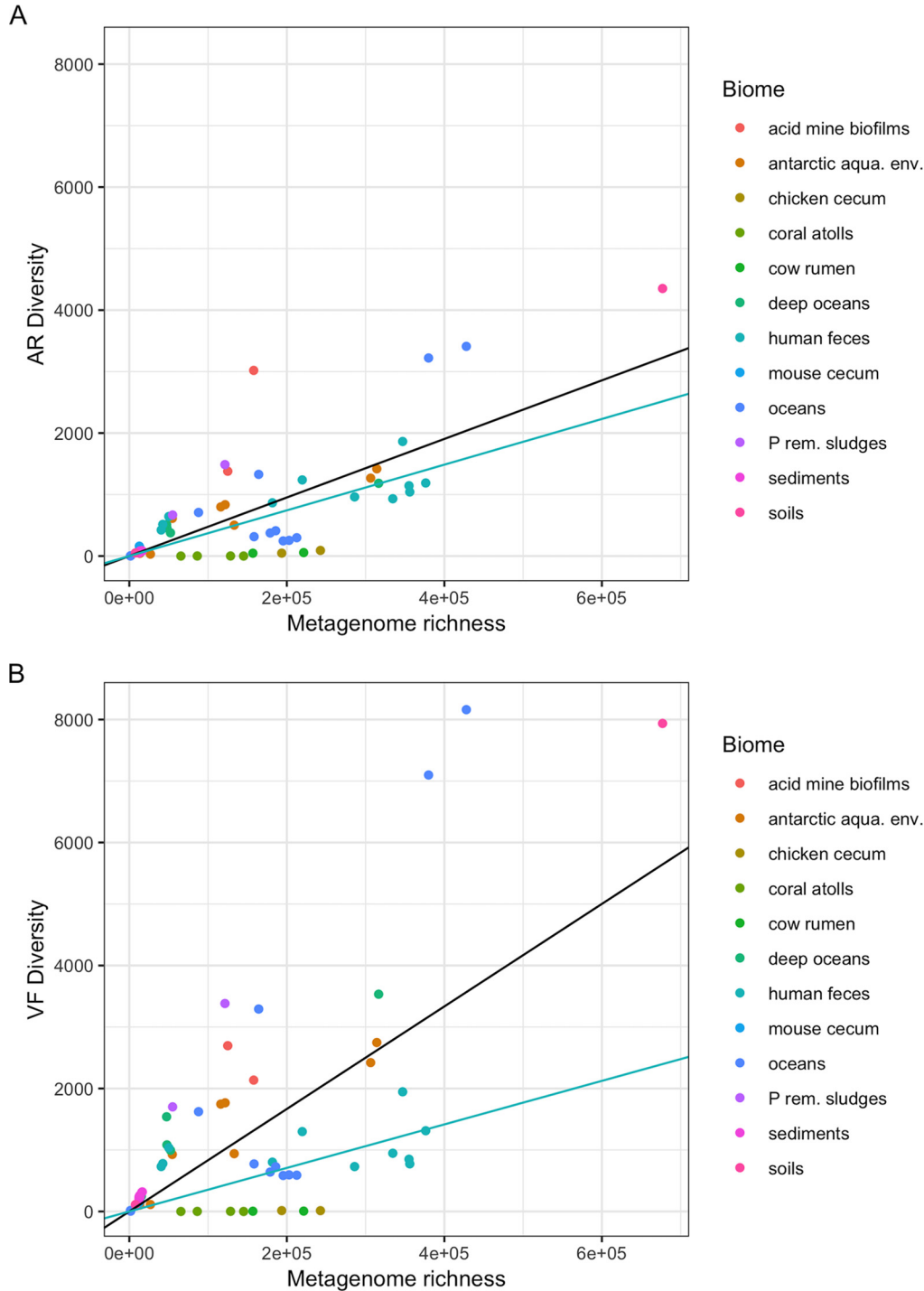


FIG 1 Distribution of the diversity numbers of antibiotic resistance and virulence factor protein families by environmental metagenomes' protein family richness. The vertical axes represent the total diversity count of AR protein families' homologues (ARd) (A) and VF protein families' homologues (VFd) (B) present in metagenomes. The horizontal axes represent the protein family richness of the metagenome: that is, the number of cluster representative sequences (see Materials and Methods). Each dot represents one of the 64 metagenomes. In each panel, the black lines represent the regressions of the ARd (A) or VFd (B) of all the metagenomes on the metagenome richness. The points are scattered, showing that the diversity of AR gene families and even more the diversity of the VF gene families can vary greatly from metagenome to metagenome. The light blue lines represent the linear regressions of the ARd or VFd of the human feces metagenomes subset on metagenome richness. Both panels A and B are drawn on the same scale to allow comparison between the two. (A) Black line (all), slope = 0.004762, $R^2 = 0.7213$, and $P < 2.2e-16$; blue line (human), slope = 0.0037169, $R^2 = 0.9004$, and $P = 2.516e-06$. (B) Black line (all), slope = 0.0083366, $R^2 = 0.6241$, $P = 5.194e-15$; blue line (human), slope = 0.0035407, $R^2 = 0.7791$, and $P = 0.0001432$.

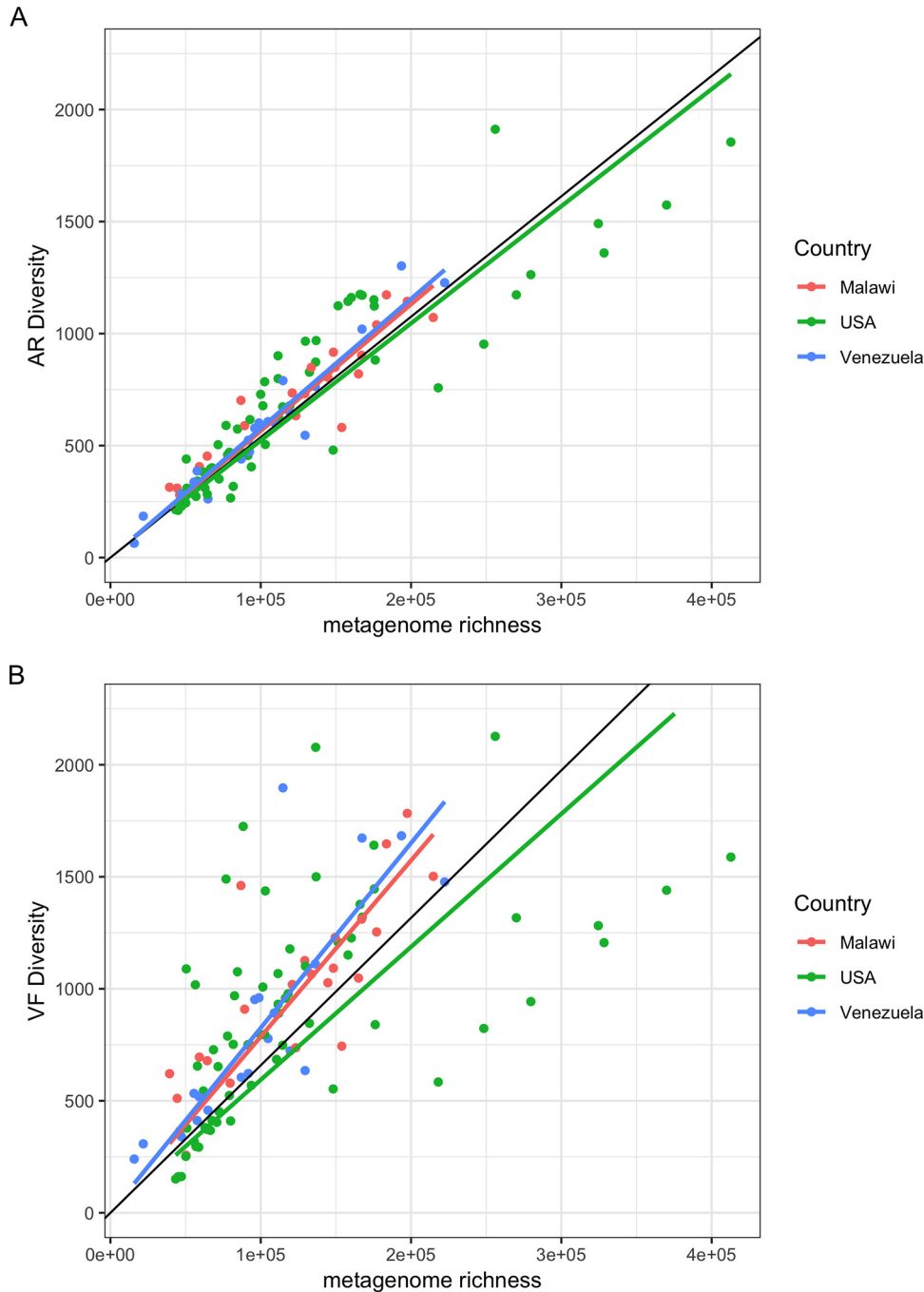


FIG 2 Distribution of the diversity numbers of antibiotic resistance and virulence factor protein families in human gut metagenomes versus metagenome protein family richness. The vertical axes represent the total diversity count of AR protein families' homologues (ARd) (A) and VF protein families' homologues (VFd) (B) present in the 110 human gut metagenomes (see Materials and Methods). As in Fig. 1, the horizontal axes represent the protein family richness of the metagenomes. Each dot represents one of the 110 human gut metagenomes. In each panel, the black lines represent the regressions of the ARd (A) or VFd (B) of all the metagenomes on the metagenome richness. The green lines represent the linear regression of the ARd (A) or VFd (B) of the USA metagenome subset on metagenome richness, the red lines are regression lines for the Malawian metagenomes, and the blue lines are regression lines for the Venezuelan metagenomes. (A) Black line fit (all), slope = 0.00537, $R^2 = 0.9547$, and $P < 2.2e-16$; green line (USA), slope = 0.0052287, $R^2 = 0.9425$, and $P < 2.2e-16$; red line fit (Malawi), slope = 0.005655, $R^2 = 0.9816$, and $P < 2.2e-16$; blue line fit (Venezuela), slope = 0.0057750, $R^2 = 0.9853$, and $P < 2.2e-16$. (B) Black line fit (all), slope = 0.00658, $R^2 = 0.8302$, and $P < 2.2e-1$; green line fit (USA), slope = 0.0059366, $R^2 = 0.7876$, and $P < 2.2e-1$; red line (Malawi), slope = 0.007865, $R^2 = 0.9462$, $P = 1.898e-15$; blue line fit (Venezuela), slope = 0.0082534, $R^2 = 0.9216$, and $P = 1.599e-12$.

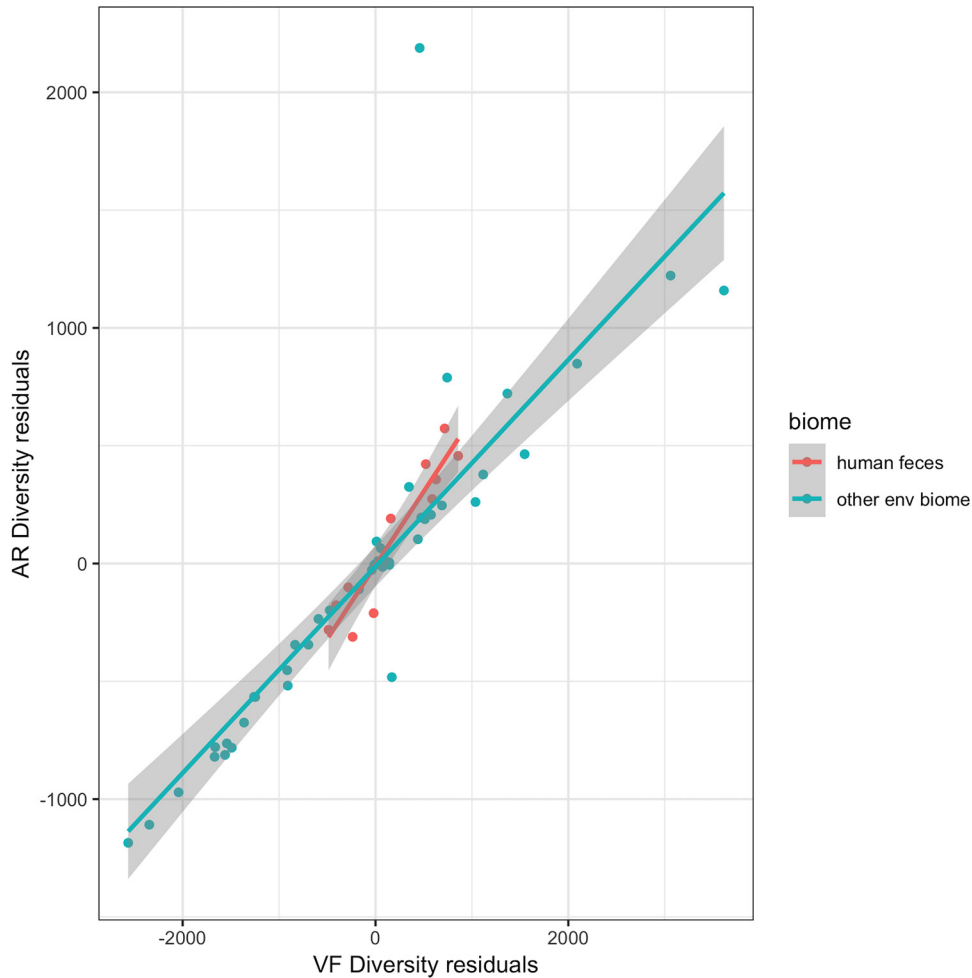


FIG 3 Distribution of AR by VF diversity in environmental metagenomes. Shown are scatter plot ARd residuals versus VFd residuals of each of the environmental and human feces 61 metagenomes (see Materials and Methods). Residuals are taken from the regressions shown in Fig. 1A and B. Each dot represents a metagenome. The blue line represents the regression line of ARd residuals of environmental metagenomes on VFd residuals (with gray shading indicating the 95% confidence interval): slope, 0.438, $R^2 = 0.739$, and $P < 2.2e-16$. The red line represents the regression line of ARd residuals of human feces metagenomes on VFd residuals: slope = 0.625, $R^2 = 0.8829$, and $P = 1.753e-05$. The difference in slopes is significant, but not very significant, with $P = 0.03$.

metagenomes pertaining to the human feces biome relative to the slope calculated for all environmental metagenomes (Fig. 3). This suggests that there could be a greater accumulation of AR over VFd for this particular biome than in environmental ones.

Figure 4 shows residuals of AR diversity versus the residuals of VF diversity in the human gut metagenomes for each population (country) and for the three populations together. The partial Pearson correlation coefficient between ARd and VFd given metagenome richness is lower in the human gut metagenome data set ($r = 0.68$; $P = 3 \times 10^{-22}$ [~ 0.0] [Fig. 4]) than in the environmental samples ($r = 0.86$ as seen above [Fig. 3]).

We can further distinguish different trends upon geographical localization of the human populations under study. The largest contribution to this graph comes from the North American samples, which account for 66/110 (60%) of the individuals (Fig. 4). The Amerindians (Venezuela) account for 21/110 metagenomes (19%), and the Malawians account for 23/110 (21%). These two non-Western human populations have very contrasting lifestyles and exposure to antibiotics: the uncontacted Amerindians harbor a resistome similar to that of pre-antibiotic era, and the people from Malawi have a high prevalence of infectious diseases and widespread use of unprescribed antibiotics (25, 26).

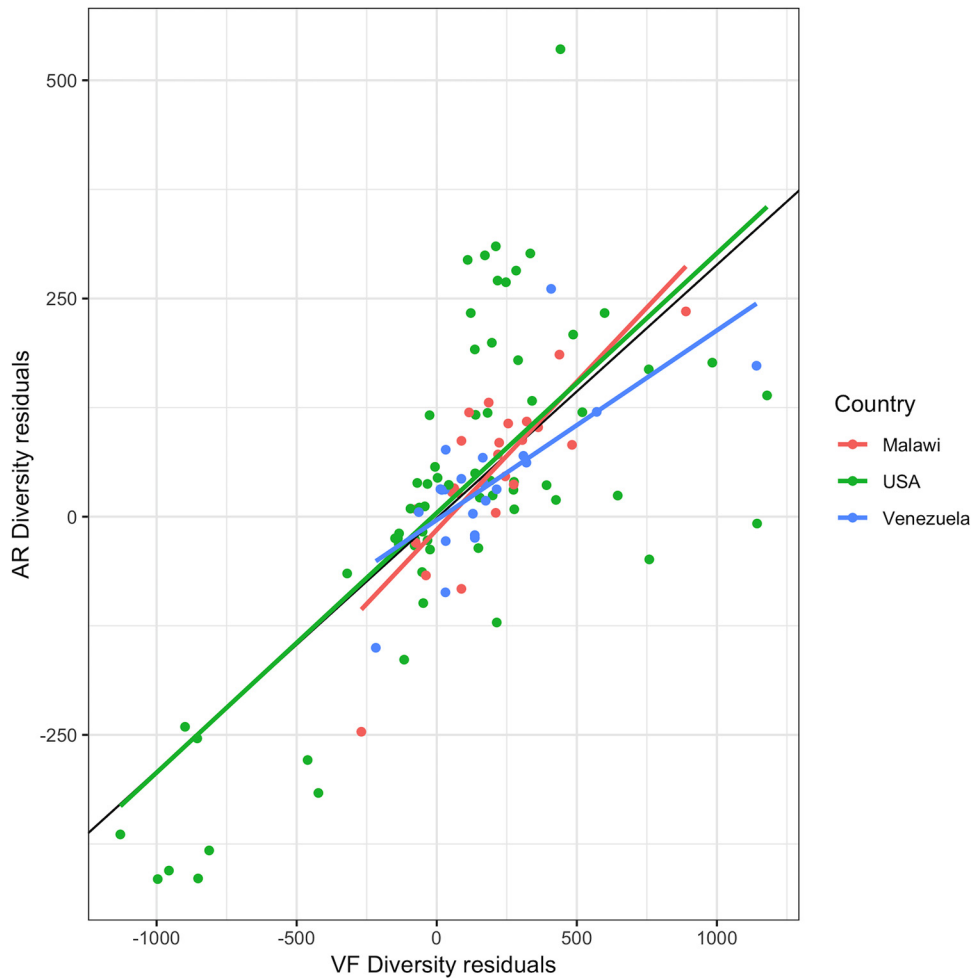


FIG 4 Distribution of AR by VF diversity in human gut metagenomes. Shown are scatter plot ARd residuals versus VFd residuals of each of the 110 human gut metagenomes. Residuals are taken from the regressions shown in Fig. 2A and B. Each dot represents a metagenome. The black line represents the regression of the ARd residuals on VFd residuals for all the metagenomes. The green line represents the linear regression of the ARd residuals on VFd residuals for the USA metagenome subset, the red line is the regression line for the Malawian metagenomes, and the blue line is the regression line for the Venezuelan metagenomes. Black line (all), slope = 0.29015, $R^2 = 0.5226$, and $P < 2.2e-16$; green line (USA), slope = 0.3083, $R^2 = 0.4332$, and $P = 1.904e-09$; red line (Malawi), slope = 0.3324, $R^2 = 0.6627$, and $P = 2.288e-06$; blue line (Venezuela), slope = 0.2052, $R^2 = 0.4882$, and $P = 0.0004258$.

The Pearson partial correlation coefficients for metagenomes in the three populations are positive and in the same range (North Americans, 0.66; Malawians, 0.81; Amerindians, 0.70). Regressions of the residuals of ARd on the residuals of VFd (Fig. 4) show a smaller slope for Amerindians (0.20) than for the United States (0.31) and Malawians (0.33), but the differences are not significant: P values range from 0.08 for the difference in Malawian-Amerindian slopes to 0.72 for the difference in Malawian and North American slopes.

Cooccurrence of AR and VFs belonging to the cell envelope. Our results suggest that cooccurrence of AR and VFs might be taking place amid bacterial communities. We wondered, however, which were the genetic traits that were more prone to this effect. We then computed the partial correlations, given the metagenome richness, between each of the AR and VF protein families in the metagenomes under study to generate the correlogram presented in Fig. 5. The vast majority of associations fall into the functional category of multidrug efflux pumps (ARs) associated with either secretion systems, as well with iron uptake and adhesion mechanisms (VFs), respectively (Fig. 5). Among the most representative associations between AR and VF traits are those

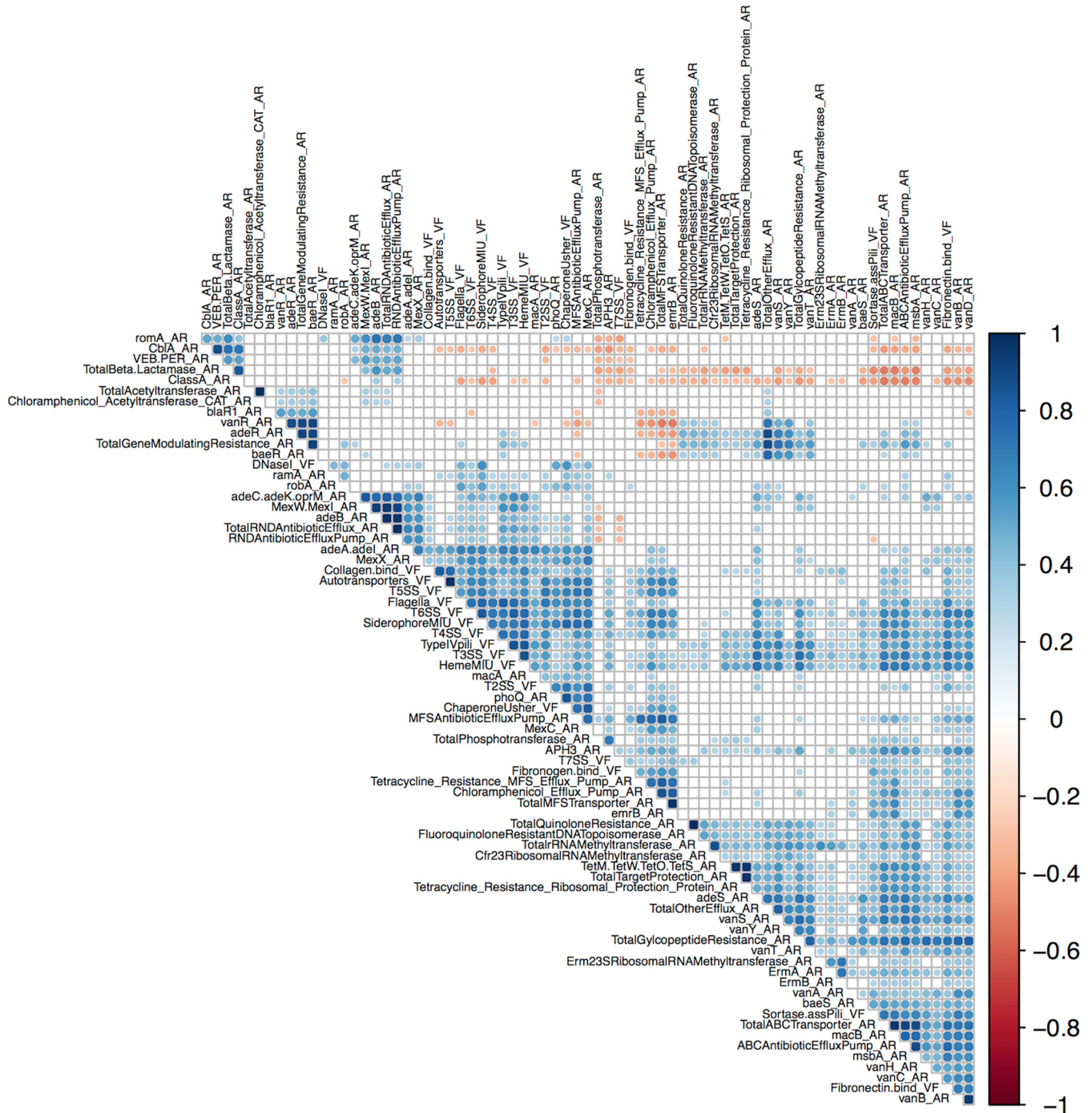


FIG 5 Correlogram of VF and AR protein families. Partial correlations for AR and VF protein families given metagenome richness have been computed. The color indicates negative (blue) or positive (red) partial correlation. The intensity of the color indicates the value. The point sizes indicate the significance of the partial correlation. (White squares indicate no significant correlation.)

belonging to the cell envelope and the general secretion mechanisms. On the other hand, the AR protein families for β -lactamases and general β -lactam resistance mechanisms showed no statistically significant correlations whatsoever.

DISCUSSION

Antibiotic resistance among environmental and human gut metagenomes. As expected, we found different homologues for antibiotic resistance (AR) gene products belonging to different protein families (AR diversity [ARd]) among environmental

metagenomes (Fig. 1A). We have shown that, in the environmental samples, ARd varies a lot from metagenome to metagenome. This may result from the differential microbial community compositions of the metagenomes, whose genetic diversity can be grouped according to the adaptation to the environment in question (25), but also from the fact that the selective pressure for the maintenance of antibiotic resistance genes in the environmental microbiome varies widely from environment to environment (28–30). Antibiotic resistance genes can easily travel and spread in the environment due to wind and runoff waters, but also due to human activities and contact with wild animals, in particular, migratory ones (31). The rise of local temperatures due to climate change also has an effect on increasing antibiotic resistance in human bacterial pathogens (32). Natural selective pressures, together with those resulting from such anthropogenic activities as medicine and agriculture, improve the dissemination of resistance genes throughout different environments.

Human gut metagenomes, on the contrary, have a less diversified repertoire of AR determinants (lower ARd) than that of their environmental counterparts.

We found a very strong correlation between the diversity of AR determinants and the metagenomes' protein family richness independently of the geographic origin of the human populations. These similar densities of ARd can indicate that, in human gut microbiomes, the number of different AR genes is not influenced by the human lifestyle, such as diet, medical care, access to antibiotics, or other cultural habits (33), and/or that the adaptation to the intestinal tract shapes microbial AR diversity as well. This is a surprising result given that antibiotic use can vary from country to country (Center for Disease Dynamics, Economics & Policy [CDDEP]; <https://cddep.org/>), and as a consequence, individuals from different human populations are under different antibiotic exposures. Forslund and coworkers have demonstrated that there are robust differences in the antibiotic resistance arsenals between countries and that these differences follow veterinary and human antibiotic use (34).

Virulence factors among environmental and human gut metagenomes. In what concerns virulence, we can assert that there are a wide diversity and density differences of VFs in environmental metagenomes, which poses as evidence of the plasticity portrayed by environmental bacteria to adapt to different hosts and niches (Fig. 1B). On the other hand, human gut microbiomes harbor a less diverse VF repertoire, especially in the U.S. samples, which seems to indicate an evolution toward adaptation to the human gut or less contact with pathogens, eventually due to vaccinations and sanitation.

Association of AR and VF. In the present study, we intended to evaluate whether the ability that a bacterial community as a whole has to colonize a host (to encode virulent factors) is linked to its survival under different antibiotic pressures (to encode antibiotic resistance). That is, whether the acquisition of new virulence traits is followed by the acquisition of new resistance traits, or vice versa. However, we are unable to determine the route of dissemination of these characteristics into and within the microbiome. It seems conceivable that the AR and VF genes could travel together in bacterial clones arriving the microbiome and that they could eventually spread inside the microbiome through horizontal gene transfer, but in this study, we have no means to say anything about the mechanisms involved in this spreading.

According to Yatsunenko and colleagues, once the human gut metagenome is established at the age of 3 years, it does not diverge much from individual to individual in terms of both phylogenetic diversity and functional richness (26).

The most relevant results presented here are those that show, when corrected for the metagenomes' protein family richness, ARd and VFd are strongly correlated both in environmental samples (Fig. 3) and in human gut samples (Fig. 4), with special emphasis on the human gut.

The North American intestinal samples (Fig. 4) show a wide variety of associations between ARd and VFd, always presenting a statistically significant correlation between these genetic traits. This result, in itself, reinforces our hypothesis that antibiotic

resistance and virulence are in fact coassociated in human gut microbiomes. The U.S. population, like those of other industrialized countries, is culturally exposed to antibiotics from health care facilities such as hospitals, antibiotic therapy, and the use of antibiotics in agriculture and livestock.

Malawians and Amerindians from Venezuela share, phylogenetically, more similar human gut microbiomes than the North Americans (26), which belong mainly to a different enterotype (26, 35). Yet, in contrast to the Venezuelan individuals, there is a very strong correlation among the Malawian ones. The uncontacted Amerindian microbiome represents a “frozen” relic of a pre-antibiotic era of the human resistome, while the Malawian gut microbiome is much more exposed, both to antibiotics and to colonization by pathogens (34, 36, 37).

Amerindians have no known access to pharmaceutical drugs, as they usually make use of the traditional indigenous medicine. Malawi is one of the poorest countries in Africa, where most people live on less than one dollar a day, many people cope with AIDS and bacterial infections, and many children suffer from severe malnutrition (36). Nutrition has been reported to have a big impact on both the human gut microbiome composition and resistome (33, 38, 39). In Malawi, UNICEF has been implementing a program of ready-to-use therapeutic food (RUTF) to reduce mortality among children. RUTF often contains antibiotics such as co-trimoxazole. It has been questioned, however, whether the success of this therapy is due to renutrition or to the combination with antibiotics (40, 41) that could have an effect on the microbial gut composition. It has been reported that there is widespread resistance to almost all of the antibiotics that are empirically used in Malawi due to the lack of routine microbiological culture and sensitivity testing (41), but also due to self-medication. Human gut samples from the Malawi people of the data set used in the present study are described as having a high overall resistance potential, with an overrepresentation of cephalosporin and tetracycline resistance genes, which may suggest extensive use of old, broad-spectrum antibiotics, a known problem in many developing countries (34).

The corepresentation of AR and VF targeting the cell envelope. Between all the possible statistical AR and VF protein families' correlations, the strongest ones are among proteins belonging to the bacterial cell envelope (Fig. 5). This result is not surprising, as many proteins that belong to the secretory system or to the secretome itself, and thus target the bacterial envelope, are frequently encoded on mobile genetic elements (6). As such, this cooccurrence may also be regarded as a direct consequence of the dynamics of these genetic elements coding for both types of determinants.

In human gut, the strongest correlations involve efflux pumps that can extrude antibiotics nonspecifically and adhesion and iron-scavenging mechanisms of pathogenicity. One possible explanation for the fact that the most frequent AR and VF associations in human gut involve efflux pumps could be that they allow a fast and efficient response to new man-made antibiotic molecules, while in environmental genomes, specific resistance mechanisms targeted to specific antibiotic molecules may have had more time to evolve to specific strategies.

Another possible explanation is that extrusion can also be implicated in pathogenic mechanisms of human-associated pathogens as well as bacterium-host interactions. For example, biofilm formation in a *Staphylococcus aureus* methicillin-resistant (MRSA) strain has been known to be essentially reliant on the activity of fibronectin binding proteins (42), and multidrug efflux pumps have direct implications for the formation and maintenance of such biofilms (43). Furthermore, quinolone-resistant *S. aureus* strains upregulate the production of fibronectin binding proteins when subjected to a sublethal dosage of ciprofloxacin (44). It has also been acknowledged that physiological levels of some cations present within the host promote the upregulation of genes encoding putative efflux transporters, oxidoreductases, and mechanisms of iron uptake in *Acinetobacter baumannii* (45) as in *Burkholderia cenocepacia* (46), which could explain the coassociation of iron acquisition systems with those of multidrug efflux pumps.

We show here that, even after correcting for protein family richness, there is a correlation between the diversity of antibiotic resistance and virulence traits in human gut microbiomes and in environmental microbiomes. Interestingly, this correlation varies from one human population to another. Human antibiotic exposure, due either to therapy or to the environment and food, can have an effect on selecting for potentially pathogenic bacteria in the human gut microbiomes. It can also drive and shape changes in the gene pool of microbiomes. This means that by selecting for resistant bacteria, we might also be selecting for more virulent strains as a side effect of antimicrobial therapy.

MATERIALS AND METHODS

Metagenomic data sets. Our human gut query cohort included 110 previously studied and publicly available metagenomes pertaining to individuals from different regions of Venezuela, Malawi, and the United States, as well as a broad age span (0.05 to 53 years) (see the data set of the project mgp98 “Human gut microbiome viewed across age and geography [WGS]” at the MG-RAST metagenomic analysis server [26]). All of the metagenome files were generated by the same team and project and by using the same bioinformatics pipeline. Our environmental cohort comprised 64 previously selected and publicly available environmental metagenomes belonging to 12 different biomes (see the data set collection of the metagenomes referred to in reference 25). Although Delmont’s team report using a data set comprised of 77 metagenomes, there are only 70 MG-RAST accession numbers present in the article’s appendix, of which only 64 are publicly available. Both project’s metagenomes were downloaded from MG-RAST on 29 January 2016 under the FASTA format, making use of successive calls to its application programming interface (API) (47) using the respective MG-RAST accession numbers available in the aforementioned reference. Each FASTA file comprised clustered protein-coding sequences retrieved from MG-RAST’s file-formatting pipeline (550.cluster.aa90.faa files). The protein sequences enclosed in these files are clustered at 90% identity, containing nonredundant translated sequences. These protein-coding formatted FASTA files contain the translation of one representative from each cluster. Thus, the numbers of protein sequences for a given metagenome used here represent its richness.

BLAST, VFDB, and resfams. For every metagenome present in our query, a blastp (50) search was performed against the VFDB database for bacterial virulence factor families (see the virulence factors of pathogenic bacteria VF.pfasta data set of the Institute of Pathogen Biology, China) (27) and the Resfams AR Proteins database of bacterial antibiotic resistance protein families (the Resfams AR Proteins v1.2 database of the Dantas Laboratory at the Washington University School of Medicine, St. Louis, MO) (48). An alternative approach made use of the Antibiotic Resistance Genes Database (ARDB) (reference antibiotic resistance genes data set resisGenes.pfasta of the Center for Bioinformatics and Computational Biology, University of Maryland) (49), but several hindrances concerning its subclassification by functional antibiotic resistance protein families made us discard the possibility of using such a database to address AR and VF protein family cooccurrence in metagenomes.

The BLAST+ executables package was downloaded on 17 November 2015 (ncbi-blast-2.2.31+ version). The VFDB database was downloaded on 11 November 2013 (31 classified FASTA files of bacterial virulence factor subfamilies), and the Resfams database was downloaded on 29 January 2016 (123 classified FASTA files of bacterial antibiotic resistance protein subfamilies). Every protein included in each of the addressed metagenomes was used as a query in order to search for similarities to either AR or VF protein-coding traits. Hence, we aimed at retrieving the best hit (best-scored alignment) that enabled us to assign an AR or VF function to each of the aforementioned metagenomic proteins. Every blastp search was performed with a very stringent E value cutoff of 10^{-15} (10 orders of magnitude lower than the conservative commonly used E value of 10^{-5}). Next, we filtered the resulting output files so as to retrieve only the alignment hits with >60% coverage of the query size and whose query and subject length ratio are between 0.75 and 1.5. When using a combination of this very stringent E value of 10^{-15} together with coverage and length ratio filters, we expected to only retrieve true homologues, avoiding false positives. Furthermore, from all the generated alignments between our metagenomic query cohort and the preceding databases, we computed the representative counts for the different gene families that coded for either AR or VF traits. Thus, the amounts of different classes (gene families) that are present in a given metagenome represent its diversity in terms of AR or VF traits.

We have also removed hits for proteins that aligned with both antibiotic resistance and virulence factor proteins (25.9% of hits against the VFDB and 29.9% of the hits against the Resfams database). The above-mentioned filters and algorithms were implemented by making use of Unix scripting languages (GNU Awk version 4.0.1 and Z-Shell version 5.0.2) in a Linux environment.

Statistical analysis. To test for relationships in metagenomes between the presence of both antibiotic resistance and virulence factor traits, we proceeded as follows. Our expectation is that the diversity number for homologues of antibiotic resistance protein families (henceforth denoted as ARd) and the diversity number for homologues of virulence factor protein families (denoted as VFd) in each metagenome (clustered at 90% identity) increases with the protein family richness of the latter (in that the protein family richness here means the total number of cluster representative proteins). Given the potential diversity of these protein families, one does not expect ARd and VFd to level off with a metagenome’s protein family richness. Therefore, we assume a linear relationship between ARd and metagenome’s protein family richness, with a fixed 0 intercept. The same is assumed for VFd. Our results will show that these assumptions are reasonable. Thus, to avoid spurious correlation, we corrected the

diversity of ARd and VFd in a metagenome by taking into account its protein family richness: we have computed the partial correlations between ARd and VFd given the metagenome richness, and equivalently, when plotting ARd versus VFd, we have instead taken residuals of regression of ARd on metagenome richness and residuals of regression of VFd on metagenome richness. Recall that the partial correlation measures the degree of association between two random variables, with the effect of a set of controlling random variables removed. For example, the partial correlation between A and B given a third variable, C , correlated with A and B is the correlation between the residuals e_a and e_b , resulting from the linear regression of A with C and of B with C . If r_{AB} , r_{AC} and r_{BC} are the correlation coefficients between A and B , A and C , and B and C , the partial correlation between A and B given C , $r_{AB.C}$ can be computed as

$$r_{AB.C} = \frac{r_{AB} - r_{AC}r_{BC}}{\sqrt{1 - r_{AC}^2}\sqrt{1 - r_{BC}^2}}$$

ACKNOWLEDGMENTS

We thank Eduardo P. C. Rocha for critical reading of the manuscript and constant interest and suggestions since the beginning of this work, Guillaume Sapriel for suggestions to improve the methods, Pedro David, Eric Ducaud, Guillaume Achaz, and all the members of our teams for fruitful discussions, and Octávio Paulo, Francisco Pina-Martins, and Sophie Brouillet for technical help and granting access to the servers. We also thank the anonymous referees who helped to improve this manuscript.

P.E., J.P., and T.N. conceived and designed the experiments. P.E. and J.P. performed the experiments. P.E., J.P., F.D., and T.N. analyzed the data. P.E., J.P., F.D., and T.N. wrote the paper.

REFERENCES

- Castanon J. 2007. History of the use of antibiotic as growth promoters in European poultry feeds. *Poult Sci* 86:2466–2471. <https://doi.org/10.3382/ps.2007-00249>.
- Sender R, Fuchs S, Milo R. 2016. Revised estimates for the number of human and bacteria cells in the body. *PLoS Biol* 14:e1002533. <https://doi.org/10.1371/journal.pbio.1002533>.
- Francino MP. 2015. Antibiotics and the human gut microbiome: dysbioses and accumulation of resistances. *Front Microbiol* 6:1543. <https://doi.org/10.3389/fmicb.2015.01543>.
- Wright GD. 2007. The antibiotic resistome: the nexus of chemical and genetic diversity. *Nat Rev Microbiol* 5:175–186. <https://doi.org/10.1038/nrmicro1614>.
- Svara F, Rankin DJ. 2011. The evolution of plasmid-carried antibiotic resistance. *BMC Evol Biol* 11:130. <https://doi.org/10.1186/1471-2148-11-130>.
- Nogueira T, Rankin DJ, Touchon M, Taddei F, Brown SP, Rocha E. 2009. Horizontal gene transfer of the secretome drives the evolution of bacterial cooperation and virulence. *Curr Biol* 19:1683–1691. <https://doi.org/10.1016/j.cub.2009.08.056>.
- Nogueira T, Touchon M, Rocha E. 2012. Rapid evolution of the sequences and gene repertoires of secreted proteins in bacteria. *PLoS One* 7:e49403. <https://doi.org/10.1371/journal.pone.0049403>.
- Dlamini BS, Montso PK, Kumar A, Ateba CN. 2018. Distribution of virulence factors, determinants of antibiotic resistance and molecular fingerprinting of *Salmonella* species isolated from cattle and beef samples: suggestive evidence of animal-to-meat contamination. *Environ Sci Pollut Res Int* 25:32694. <https://doi.org/10.1007/s11356-018-3231-4>.
- Barkovskii A, Babb C, Hurlley D, Shin E. 2015. Origins and environmental mobility of antibiotic resistance genes, virulence factors and bacteria in a tidal creek's watershed. *J Appl Microbiol* 118:764–776. <https://doi.org/10.1111/jam.12735>.
- Abana D, Gyamfi E, Dogbe M, Opoku G, Opare D, Boateng G, Mosi L. 2019. Investigating the virulence genes and antibiotic susceptibility patterns of *Vibrio cholerae* O1 in environmental and clinical isolates in Accra, Ghana. *BMC Infect Dis* 19:76. <https://doi.org/10.1186/s12879-019-3714-z>.
- Beceiro A, Tomas M, Bou G. 2013. Antimicrobial resistance and virulence: a successful or deleterious association in the bacterial world? *Clin Microbiol Rev* 26:185–230. <https://doi.org/10.1128/CMR.00059-12>.
- Sommer MOA, Church GM, Dantas G. 2010. The human microbiome harbors a diverse reservoir of antibiotic resistance genes. *Virulence* 1:299–303. <https://doi.org/10.4161/viru.1.4.12010>.
- Andersson DI, Hughes D. 2014. Microbiological effects of sublethal levels of antibiotics. *Nat Rev Microbiol* 12:465–478. <https://doi.org/10.1038/nrmicro3270>.
- Vega NM, Gore J. 2014. Collective antibiotic resistance: mechanisms and implications. *Curr Opin Microbiol* 21:28–34. <https://doi.org/10.1016/j.mib.2014.09.003>.
- Dionisio F, Matic I, Radman M, Rodrigues OR, Taddei F. 2002. Plasmids spread very fast in heterogeneous bacterial communities. *Genetics* 162:1525–1532.
- Gama JA, Zilhão R, Dionisio F. 2017. Multiple plasmid interference—pledging allegiance to my enemy's enemy. *Plasmid* 93:17–23. <https://doi.org/10.1016/j.plasmid.2017.08.002>.
- Gama JA, Zilhão R, Dionisio F. 2017. Conjugation efficiency depends on intra and intercellular interactions between distinct plasmids: plasmids promote the immigration of other plasmids but repress co-colonizing plasmids. *Plasmid* 93:6–16. <https://doi.org/10.1016/j.plasmid.2017.08.003>.
- Wright G. 2005. Bacterial resistance to antibiotics: enzymatic degradation and modification. *Adv Drug Deliv Rev* 57:1451–1470. <https://doi.org/10.1016/j.addr.2005.04.002>.
- Nicoloff H, Andersson DI. 2016. Indirect resistance to several classes of antibiotics in cocultures with resistant bacteria expressing antibiotic-modifying or -degrading enzymes. *J Antimicrob Chemother* 71:100–110. <https://doi.org/10.1093/jac/dkv312>.
- Sorg RA, Lin L, van Doorn GS, Sorg M, Olson J, Nizet V, Veening J-W. 2016. Collective resistance in microbial communities by intracellular antibiotic deactivation. *PLoS Biol* 14:e2000631. <https://doi.org/10.1371/journal.pbio.2000631>.
- Domingues IL, Gama JA, Carvalho LM, Dionisio F. 2017. Social behaviour involving drug resistance: the role of initial density, initial frequency and population structure in shaping the effect of antibiotic resistance as a public good. *Heredity* 119:295–301. <https://doi.org/10.1038/hdy.2017.33>.
- Ponomarova O, Patil KR. 2015. Metabolic interactions in microbial communities: untangling the Gordian knot. *Curr Opin Microbiol* 27:37–44. <https://doi.org/10.1016/j.mib.2015.06.014>.
- Rosenthal AZ, Matson EG, Eldar A, Leadbetter JR. 2011. RNA-seq reveals cooperative metabolic interactions between two termite-gut spirochete species in co-culture. *ISME J* 5:1133–1142. <https://doi.org/10.1038/ismej.2011.3>.
- Foster KR, Bell T. 2012. Competition, not cooperation, dominates inter-

- actions among culturable microbial species. *Curr Biol* 22:1845–1850. <https://doi.org/10.1016/j.cub.2012.08.005>.
25. Delmont TO, Malandain C, Prestat E, Larose C, Monier J-M, Simonet P, Vogel TM. 2011. Metagenomic mining for microbiologists. *ISME J* 5:1837–1843. <https://doi.org/10.1038/ismej.2011.61>.
 26. Yatsunenkov T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, Magris M, Hidalgo G, Baldassano RN, Anokhin AP, Heath AC, Warner B, Reeder J, Kuczynski J, Caporaso JG, Lozupone CA, Lauber C, Clemente JC, Knights D, Knight R, Gordon JI. 2012. Human gut microbiome viewed across age and geography. *Nature* 486:222–227. <https://doi.org/10.1038/nature11053>.
 27. Chen L, Xiong Z, Sun L, Yang J, Jin Q. 2012. VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors. *Nucleic Acids Res* 40:D641–D645. <https://doi.org/10.1093/nar/gkr989>.
 28. Martinez JL. 2008. Antibiotics and antibiotic resistance genes in natural environments. *Science* 321:365–367. <https://doi.org/10.1126/science.1159483>.
 29. Martínez JL, Baquero F. 2002. Interactions among strategies associated with bacterial infection: pathogenicity, epidemicity, and antibiotic resistance. *Clin Microbiol Rev* 15:647–679. <https://doi.org/10.1128/CMR.15.4.647-679.2002>.
 30. Fitzpatrick D, Walsh F. 2016. Antibiotic resistance genes across a wide variety of metagenomes. *FEMS Microbiol Ecol* 92:fiv168. <https://doi.org/10.1093/femsec/fiv168>.
 31. Allen HK, Donato J, Wang HH, Cloud-Hansen KA, Davies J, Handelsman J. 2010. Call of the wild: antibiotic resistance genes in natural environments. *Nat Rev Microbiol* 8:251–259. <https://doi.org/10.1038/nrmicro2312>.
 32. MacFadden D, McGough S, Fisman D, Santillana M, Brownstein J. 2017. Antibiotic resistance increases with local temperature. *Open Forum Infect Dis* 4(Suppl 1):S179. <https://doi.org/10.1093/ofid/ofx163.327>.
 33. Ghosh TS, Gupta SS, Nair GB, Mande SS. 2013. In silico analysis of antibiotic resistance genes in the gut microflora of individuals from diverse geographies and age-groups. *PLoS One* 8:e83823. <https://doi.org/10.1371/journal.pone.0083823>.
 34. Forslund K, Sunagawa S, Kultima JR, Mende DR, Arumugam M, Typas A, Bork P. 2013. Country-specific antibiotic use practices impact the human gut resistome. *Genome Res* 23:1163–1169.
 35. Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, Fernandes GR, Tap J, Bruls T, Batto J-M, Bertalan M, Borruel N, Casellas F, Fernandez L, Gautier L, Hansen T, Hattori M, Hayashi T, Kleerebezem M, Kurokawa K, Leclerc M, Levenez F, Manichanh C, Bjørn Nielsen H, Nielsen T, Pons N, Poulain J, Qin J, Sicheritz-Ponten T, Tims S, Torrents D, Ugarte E, Zoetendal EG, Wang J, Guarner F, Pedersen O, de Vos WM, Brunak S, Doré J, MetaHIT Consortium, Weissenbach J, Dusko Ehrlich S, Bork P. 2011. Enterotypes of the human gut microbiome. *Nature* 473:174–180. <https://doi.org/10.1038/nature09944>.
 36. WHO. 2016. Global Health Observatory (GHO) data. WHO, Geneva, Switzerland. <https://www.who.int/gho/en/>.
 37. UNICEF. 2015. Committing to child survival: a promise renewed—progress report 2015. UNICEF, New York, NY. https://www.unicef.org/publications/index_83078.html.
 38. Ghosh TS, Gupta SS, Bhattacharya T, Yadav D, Barik A, Chowdhury A, Das B, Mande SS, Nair GB. 2014. Gut microbiomes of Indian children of varying nutritional status. *PLoS One* 9:e95547. <https://doi.org/10.1371/journal.pone.0095547>.
 39. Pehrsson EC, Tsukayama P, Patel S, Mejía-Bautista M, Sosa-Soto G, Navarrete KM, Calderon M, Cabrera L, Hoyos-Arango W, Bertoli MT, Berg DE, Gilman RH, Dantas G. 2016. Interconnected microbiomes and resistomes in low-income human habitats. *Nature* 533:212–216. <https://doi.org/10.1038/nature17672>.
 40. Alcoba G, Kerac M, Breyse S, Salpeteur C, Galetto-Lacour A, Briand A, Gervais A. 2013. Do children with uncomplicated severe acute malnutrition need antibiotics? A systematic review and meta-analysis. *PLoS One* 8:e53184. <https://doi.org/10.1371/journal.pone.0053184>.
 41. Makoka MH, Miller WC, Hoffman IF, Cholera R, Gilligan PH, Kamwendo D, Malunga G, Joaki G, Martinson F, Hosseinipour MC. 2012. Bacterial infections in Lilongwe, Malawi: aetiology and antibiotic resistance. *BMC Infect Dis* 12:67. <https://doi.org/10.1186/1471-2334-12-67>.
 42. McCourt J, O'Halloran DP, McCarthy H, O'Gara JP, Geoghegan JA. 2014. Fibronectin-binding proteins are required for biofilm formation by community-associated methicillin-resistant *Staphylococcus aureus* strain LAC. *FEMS Microbiol Lett* 353:157–164. <https://doi.org/10.1111/1574-6968.12424>.
 43. Soto SM. 2013. Role of efflux pumps in the antibiotic resistance of bacteria embedded in a biofilm. *Virulence* 4:223–229. <https://doi.org/10.4161/viru.23724>.
 44. Bisognano C, Kelley WL, Estoppey T, Francois P, Schrenzel J, Li D, Lew DP, Hooper DC, Cheung AL, Vaudaux P. 2004. A recA-LexA-dependent pathway mediates ciprofloxacin-induced fibronectin binding in *Staphylococcus aureus*. *J Biol Chem* 279:9064–9071. <https://doi.org/10.1074/jbc.M309836200>.
 45. Hood MI, Jacobs AC, Sayood K, Dunman PM, Skaar EP. 2010. *Acinetobacter baumannii* increases tolerance to antibiotics in response to monovalent cations. *Antimicrob Agents Chemother* 54:1029–1041. <https://doi.org/10.1128/AAC.00963-09>.
 46. Drevinek P, Holden MTG, Ge Z, Jones AM, Ketchell I, Gill RT, Mahenthalingam E. 2008. Gene expression changes linked to antimicrobial resistance, oxidative stress, iron depletion and retained motility are observed when *Burkholderia cenocepacia* grows in cystic fibrosis sputum. *BMC Infect Dis* 8:121. <https://doi.org/10.1186/1471-2334-8-121>.
 47. Gibson MK, Forsberg KJ, Dantas G. 2015. Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. *ISME J* 9:207–216. <https://doi.org/10.1038/ismej.2014.106>.
 48. Wilke A, Bischof J, Harrison T, Brettin T, D'Souza M, Gerlach W, Matthews H, Paczian T, Wilkening J, Glass EM, Desai N, Meyer F. 2015. A RESTful API for accessing microbial community data for MG-RAST. *PLoS Comput Biol* 11:e1004008. <https://doi.org/10.1371/journal.pcbi.1004008>.
 49. Liu B, Pop M. 2009. ARDB—Antibiotic Resistance Genes Database. *Nucleic Acids Res* 37:D443–D447. <https://doi.org/10.1093/nar/gkn656>.
 50. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402. <https://doi.org/10.1093/nar/25.17.3389>.