The ISME Journal (2017) 11, 651–662

www.nature.com/ismej

ORIGINAL ARTICLE

Co-occurrence of antibiotic and metal resistance genes revealed in complete genome collection

Li-Guan Li[1], Yu Xia[1] and Tong Zhang
Environmental Biotechnology Laboratory, Department of Civil Engineering, The University of Hong Kong, Hong Kong

The high frequency of antibiotic resistance is a global public health concern. More seriously, widespread metal pressure in the environment may facilitate the proliferation of antibiotic resistance via coselection of antibiotic resistance genes (ARGs) and metal resistance genes (MRGs). Given the lack of comprehensive understanding of the ARG and MRG coselection, in this study both abundance relationship and genetic linkage between ARGs and MRGs were rigorously investigated by performing a genomic analysis of a large complete genome collection. Many more ARGs were enriched in human-associated bacteria compared with those subjected to less anthropogenic interference. The signatures of ARG and MRG co-occurrence were much more frequent and the distance linkages between ARGs and MRGs were much more intimate in human pathogens than those less human-associated bacteria. Moreover, the co-occurrence structures in the habitat divisions were significantly different, which could be attributed to their distinct gene transfer potentials. More exogenous ARGs and MRGs on the genomes of human pathogens indicated the importance of recent resistance acquisition in resistome development of human commensal flora. Overall, the study emphasizes the potential risk associated with ARG and MRG coselection of both environmental and medical relevance.
The ISME Journal (2017) **11,** 651–662; doi:10.1038/ismej.2016.155; published online 13 December 2016

Introduction

The prevalence and propagation of antibiotic resistance is a serious global public health concern that has been even aggravated by the increasing incidence of multidrug resistance in clinical pathogens (Fischbach and Walsh, 2009; Zowawi *et al.*, 2015; Pamer, 2016). As reported by the WHO (2014), we are now at the brink of a postantibiotic era where antibiotic treatment failures will be common in the coming decades. Much more seriously, frequent emergence of antibiotic resistance has substantially extended beyond medical settings (Forsberg *et al.*, 2012; Wright, 2010). Despite the tremendous efforts directed toward restricting antibiotic consumption worldwide, limited success has been achieved evidenced by the increasing prevalence of antibiotic resistance (Andersson and Hughes, 2010). It has been suggested that the evolution and dissemination of antibiotic resistance is not governed by antibiotics alone, but is instead shaped by a complex array of factors (Andersson and Hughes, 2010). Notably, because common modes of action are shared by metal and antibiotic resistance, there is growing concern that metal stress in the environment may function as a coselection agent in the selection and spread of antibiotic resistance genes (ARGs) (Baker-Austin *et al.*, 2006; Seiler and Berendonk, 2012). Considering the alarming anthropogenic levels of metal pollution in the environment (Charlesworth *et al.*, 2011; Chen *et al.*, 2015), it is of particular concern that metal contamination can exert a long-standing and widespread coselection pressure for antibiotic resistance of both environmental and clinical importance through genetic couplings (Baker-Austin *et al.*, 2006), for example, coresistance (close linkage between two or more different resistance genes), cross-resistance (single genetic element performing both antibiotic and metal resistance) and coregulation (shared regulatory system to antibiotic and metal resistance).

Evidences for coselection of ARGs and MRGs (metal resistance genes) have been reported in a variety of environments over the past several decades. At sites characterized by both high antibiotic and metal burden, such as soil and water impacted by agriculture and aquaculture, coselection has frequently been observed (Ji *et al.*, 2012; Seiler and Berendonk, 2012; Zhu *et al.*, 2013). Moreover, the coselection phenomenon has been exemplified in the bacterial communities inhabiting the human

Correspondence: T Zhang, Environmental Biotechnology Laboratory, Department of Civil Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong 999077, Hong Kong.
E-mail: zhangt@hku.hk
[1]These authors contributed equally to this work.
Received 13 June 2016; revised 30 September 2016; accepted 7 October 2016; published online 13 December 2016

body (Davis *et al.*, 2005; Amsaveni *et al.*, 2015). Another challenge faced by the public for antibiotic and metal resistance coselection is the cotransfer of ARGs and MRGs via MGEs (mobile gene elements), which has been realized in various habitats, such as human gut (Broaders *et al.*, 2013; Ma *et al.*, 2016), animal gut and impacted soil (Johnson *et al.*, 2016; Losada *et al.*, 2016), sediment (Rosewarne *et al.*, 2010) and sludge (Gaze *et al.*, 2011). However, to our knowledge, most studies on the coselection are primarily anecdotal and descriptive. Given the lack of systemic knowledge regarding genetic association between ARGs and MRGs, a comprehensive understanding is urgently needed.

The difficulty in rigorously evaluating the interrelationships between ARGs and MRGs using comprehensive phenotypic and genotypic analysis has hindered our ability to systematically interrogate the coselection. Currently, the increasing number of genomes from a variety of ecosystems have been obtained by the advent in high-throughput sequencing (Loman and Pallen, 2015), providing researchers an opportunity to extensively analyze specific genetic associations in which close linkage on genome can confer great advantage in developing coselection (Summers, 2002). In this study, by coupling a large complete genome collection with the powerful bioinformatics analyses, the genetic relationships between ARGs and MRGs across genome categories (phylogenies/pathogenicities/habitats) were investigated from different perspectives, including the abundance profile of broad-spectrum different ARGs/MRGs types and high-resolution evaluation of their genetic linkages. In this way, we aimed to not only further understand the impact of genetic events in shaping the ARGs and MRGs pools in distinct genome categories but also provide robust evidence of co-occurrence between ARGs and MRGs on specific genomes, particularly those that are substantially relevant to humans. To complement the risk scenario assessment, the underlying co-occurrence structures among ecologies and the cotransfer potential of the two gene types across environments were further analyzed. We believe that the results will shed light on the complex relationship between metal and antibiotic resistance, thereby improving our understanding of the hidden health risk associated with ARG and MRG co-occurrence.

Materials and methods

Complete genome collection

A total of 5436 complete bacterial genomes covering a large diversity of bacteria (including 39 phyla) were downloaded from the NCBI genome database (ftp://ftp.ncbi.nlm.nih.gov/genomes/genbank/bacteria) in August 2016 in GenBank format, and were analyzed following the pipeline well developed in the past 2 years in our group. The taxonomic lineages of the downloaded genomes were retrieved from the NCBI taxonomy database. The full list of genomes used in this study is summarized in Supplementary Table S1. For each genome, the protein coding sequences and genetic location information were extracted from the GenBank files using a self-written Python script (available at https://github.com/LiguanLi/ARG_MRG_Cooccurrence). The coding sequences were used in all downstream searches and analyses of the resistance genes. The habitat information of the genomes was curated using the corresponding metadata available in the IMG database (August 2016) by combing the fields of 'habitat', 'ecology' and 'isolation' (Supplementary Table S2). The potential pathogenicity of each species in the genome collection was obtained from the published database covering all recognized species of human bacterial pathogens (Woolhouse and Gowtage-Sequeria, 2005).

ARGs, MRGs and MGEs retrieval

The Structured ARG (SARG) database, including 23 types and 1277 subtypes in a hierarchical structure (type–subtype–reference sequence), was used to facilitate the detection and quantification of a broad-spectrum of different ARGs (online platform can be accessed through http://smile.hku.hk/SARGs) (Yang *et al.*, 2016). The MRG database ComMet, including 23 metal types, integrated three sources: the highly compact database BacMet (Pal *et al.*, 2014), the copper resistance protein database (Li *et al.*, 2014) and the arsenic metabolism protein database (Cai *et al.*, 2013). ARG and MRG retrieval was achieved by searching against databases. The coding sequences of all genomes were subject to a batch BLASTP search against the SARG database with an e-value cutoff of $1e-5$. Those hits with $\geqslant 90\%$ similarity and $\geqslant 80\%$ alignment length ratio were annotated as ARGs (Yang *et al.*, 2013). A sequence was annotated as an MRG if its BLASTP search against the ComMet met the criteria of e-value $1e-5$, 80% similarity and 90% hit length (Li *et al.*, 2014). To retrieve DNA segments capable of moving inter/intragenomes and keeping heritable stability afterwards (Frost *et al.*, 2005; Stokes and Gillings, 2011), MGEs were identified by string matches to one of the following keywords (Forsberg *et al.*, 2014) in the gene description: transposase, transposon, conjugative, integrase, integron, recombinase, conjugal, mobilization, recombination and plasmid.

Enrichment analysis

After filtering the BLASTP results, both the ARG- and MRG-like sequences were deduplicated and categorized into antibiotic and metal types, respectively. The enrichment of the resistance genes across phylogenies, habitats and pathogenicities was tested by Fisher's exact test, with the *P*-value further adjusted using the Benjamini–Hochberg correction for multiple comparisons. Significant enrichment

was defined as a *P*-value <0.01 and an odds ratio >1.

Genetic co-occurrence analysis

To highly resolve the co-occurrence of ARGs and MRGs on circular genomes, two indexes were introduced, the average minimum distance and incidence of encountering (see Supplementary Methods for a detailed demonstration of the calculation). Briefly, genomes carrying both ARGs and MRGs were used for the co-occurrence analysis. On each genome, the average minimum distance ($\mathrm{Met}A_{\mathrm{min}}$ (bp)) was calculated as the sum of distances of the closest ARG from each MRG divided by the number of MRGs (Equation (1) in Supplementary Methods). To obtain the incidence of encountering, the number of ARGs within the assigned distance (200 bp–100 Kbp, by 200 bp step) from every MRG in each genome was counted, and then the raw counts were averaged within a given genome category using Equation (2) in Supplementary Methods. In particular, to further identify which pathogen species exhibited high ARG and MRG co-occurrence potential, a rigorous filtration of total ARG and MRG numbers $\geqslant 3$ per genome and genome number $\geqslant 3$ per species was performed before the distance analysis. To evaluate the diversity of ARG and MRG co-occurrence, the number of unique ARG subtypes encountered adjacent to MRGs was recorded along the given distance (200 bp–100 Kbp) in each genome, which was used to calculate the Shannon index. Moreover, the nearest ARG types for all MRG types were summarized. In addition, the ARG and MRG cotransfer potential on the genome was analyzed by the average minimum distance from MGEs in pair, where the respective sum of distances of the closest ARG and MRG from each MGE was divided by the number of MGEs. The nearest ARG and MRG from each MGE were identified as a type pair ARG:MRG (e.g., tetracycline:As) in the cotransfer preference analysis.

Cluster analysis of the genomes using the resistance gene profile

To compare the underlying structures of the ARG and MRG co-occurrence profile across the different genome categories, a count matrix of the closest ARG–MRG pairs in all ARG–MRG types (e.g., bacitracin-Cu) was generated for each genome. The count table was then used to generate the Bray–Curtis distance matrix, and a principal coordinate analysis was performed using the R package 'vegan'. Significant clustering was determined using analysis of similarities with a *P*-value <0.001. Random Forest analysis was performed with the R package 'randomForest' to determine the ARG–MRG pairs that were most likely to discriminate genomes between the different categories. Principal coordinate analysis plots were plotted along with these most discriminating ARG–MRG pairs as biplots using the 'ggplot2' package. The biplot position was calculated as the weighted average of the coordinate position of all genomes along the first two principal coordinate analysis axes, where the weight is abundance of the ARG–MRG type in all genomes. The R script for this analysis is available at https://github.com/LiguanLi/ARG_MRG_Cooccurrence.

Tetranucleotide signature analyses

Genome-wide tetranucleotide frequency (TNF) signatures were calculated for each complete genome using self-written function in R. By setting sliding window of 5 Kbp, TNF patterns were calculated across the full length of the genome. Based on the TNF matrix composed by the whole genome and all windows, deviation of TNF in each window from the global TNF pattern was obtained by calculating the Pearson's correlation coefficient (r^2) between TNF counts (in term of log(1+counts)) in each window with that of the whole genome (Daims *et al*., 2015). TNF correlation coefficient of Sliding windows (fragments) containing ARG were summarized based on the type of ARG they carried.

Results and discussion

Broad-spectrum abundance profile

The genes encoding antibiotic and metal resistance were identified in a broad spectrum of phylogenies. Overall, ARGs and MRGs were detected in 48% (2617 genomes in 10 phyla) and 47% (2543 genomes in 15 phyla) of 5436 complete bacterial genomes (Figure 1 and Supplementary Figure S1). Approximately 0.8% of the genomes under analysis had high abundances (>50 hits per genome) of both ARGs and MRGs, for example, genomes in genera *Escherichia*, *Klebsiella*, *Salmonella* and *Shigella*. Much more striking is the apparent propagation of resistance genes in enteric microorganisms, such as *Enterobacteriaceae* accounted for $>90\%$ of the top 100 genomes that carried abundant ARGs or MRGs. Another notable observation is the relatively high correlation ($R^2 \geqslant 0.6$) between the abundances of ARGs and MRGs in *Enterobacteriaceae* (Supplementary Figure S2), which is consistent with the increasingly reported occurrence of typical ARGs and MRGs in intestinal microbial communities that have been subject to antibiotic and metal pressure (Wireman *et al*., 1997; Davis *et al*., 2005; Zhu *et al*., 2013). Albeit the resistance gene profile in the large complete genome collection could provide a comprehensive picture of resistance dissemination at the boundary of our current knowledge, readers should bear in mind the possible biases in the presented results caused by the uncultivable majority of environmental microbes that do not have available genome information and the inevitable uneven distribution across phylogenies/ecologies in current complete genome collection (Supplementary Figure S3).

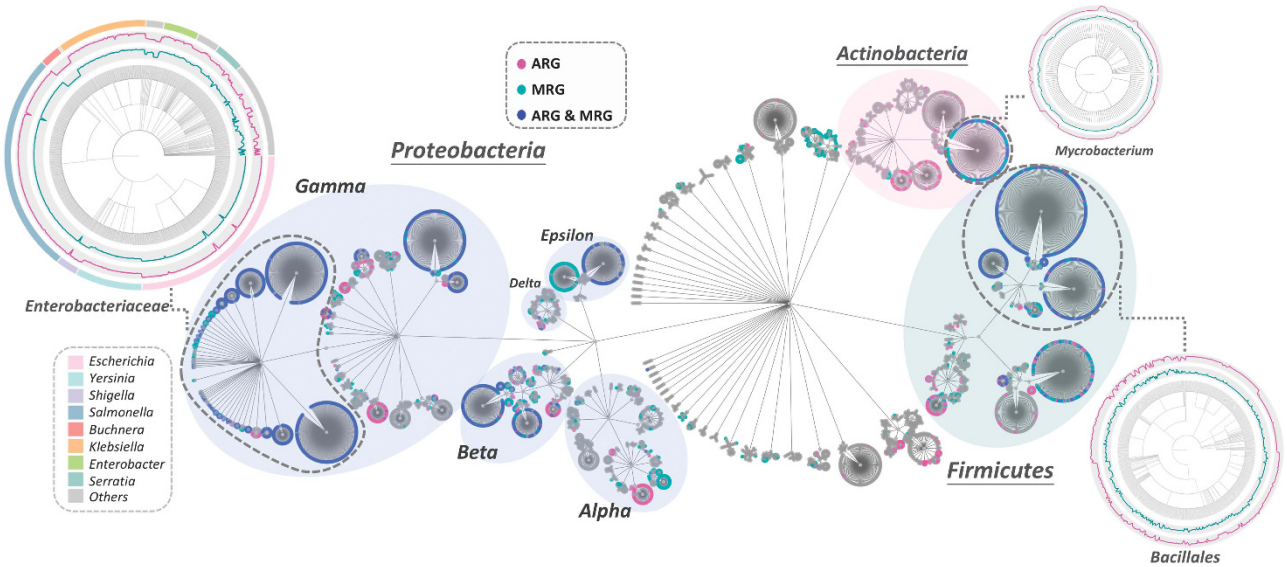

Figure 1 Presence and absence of ARGs and MRGs in the complete genome collection. The phylogenic lineages on circular tree for each strain are as follows: domain, phylum, class, order, family, genus and strain. The strain nodes (gray) are highlighted when they carry resistance genes. Three main phyla (*Proteobacteria*, *Firmicutes* and *Actinobacteira*) are highlighted, of which the main groups carrying resistance genes are presented as circles with the ARG and MRG abundance profile (log 2 transformed) of each genome.

Among the different resistance gene types, an uneven abundance distribution was revealed (Supplementary Figure S4). The five most abundant specific ARG types (excluding multidrug and unclassified ARGs) were beta-lactam (0.7%), macrolide-lincosamide-streptogramin (0.5%), bacitracin (0.5%), aminoglycoside (0.4%) and tetracycline (0.3%). Among the 23 MRG types detected, Zn, Cu, Ni, Mn and Hg resistance genes accounted for more than half (60%) of the total MRGs retrieved. Generally, the resistance genes for commonly consumed antibiotics and metals are more likely to be present in high abundance, which also has been demonstrated in the previous abundance-based surveys (Li *et al*., 2015; Pal *et al*., 2015).

These above lines of evidences indicate that abundance profiles of certain ARG and MRG types might correlate with specific microbial communities, particularly those that are intimately associated with humans. Owing to substantial periods of evolution under extensive anthropogenic activities, 'favorable' gene dissemination in microorganisms may markedly protect them from detrimental environmental conditions (Sandegren and Andersson, 2009; Andersson and Hughes, 2014; Huijbers *et al*., 2015). Therefore, the general resistance gene profile aroused further interest in evaluating the genes that are enriched in distinct genomes across distinct phyla (*Proteobacteria*, *Firmicutes*, *Actinobacteria* and *Bacteroidetes*), pathogenicities (pathogen and non-pathogen) and habitats (human, water and soil) (Supplementary Figure S5). Of the three habitats under investigation, both resistance gene types more tended to be enriched in the human and soil habitats. Across the four major phyla, enrichment was observed for 18 ARG types. Intriguingly, resistance genes for the widely consumed beta-lactam and tetracycline were not enriched in the main phylum *Proteobacteria* in this study, implying that the large-scale and long-term overuse of the antibiotic groups facilitated the development of resistance genes across phylogenetic boundaries. Among MRGs, 19 types were all specifically enriched in different phyla, such as As in *Firmicutes* and Zn in *Proteobacteria*. Regarding the pathogenicity status, many more ARGs were enriched in pathogens compared with non-pathogens, which corresponded with the results of a previous HMM (hidden Markov model)-based ARG analysis (Gibson *et al*., 2015). Taken together, these results suggest a high-risk scenario of widespread antibiotic resistance in human pathogens.

Co-occurrence between ARGs and MRGs

The above observed ARG and MRG abundance profiles within the genome collection, plus previous sporadic studies on coresistance in pathogens (Gillings *et al*., 2015; Wales and Davies, 2015; Di Cesare *et al*., 2016; Johnson *et al*., 2016) and the recently reported abundance correlation in clinically important genera (e.g., *Escherichia*, *Shigella* and *Klebsiella*) (Pal *et al*., 2015) kindled our interest in thoroughly investigating their possible physical genetic co-occurrence, particularly in human-associated bacteria. Therefore, two concepts, the incidence of encountering ARGs along the distance from MRGs and their average minimum distance ($\mathrm{Met}A_{\mathrm{min}}$), were used to further evaluate the possible risk associated with ARG and MRG coselection by examining their likeliness of co-occurrence on

genomes. The ARG and MRG co-occurrence profiles were compared between pathogen and non-pathogen genomes (Figures 2a and b). Generally, the signatures of ARG and MRG co-occurrence were significantly more frequent in pathogen than non-pathogen genomes (within distance range of 100 Kbp) (Figure 2a). As the distance from the MRGs increased, the incidence of ARGs in the pathogen genomes was always greater than in the non-pathogen genomes, indicating that MRGs might have a close relationship with ARG occurrence in pathogens. To further strengthen this argument, a much closer proximity between the ARGs and MRGs in pathogens was observed in the MetA_{min} profile (Figure 2b). To identify which pathogen species contributed to the above distance relationship, a profile of the co-occurrence potential in 16 pathogen species was obtained (Supplementary Figure S6). All the 16 species exhibited a higher incidence of encountering ARGs as the distance from MRGs increased. The trend was particularly evident in *Enterobacter hormaechei*, *Klebsiella oxytoca*, *Klebsiella pneumoniae*, *Enterobacter cloacae* and *Escherichia coli*. Additionally, the MetA_{min} profile further supported substantial potential of MRGs driving ARG occurrence among the 16 pathogen species, as their average MetA_{min} value was 103 Kbp, which was much shorter than those of non-pathogen genomes (380 Kbp). In addition to the pathogenicity, the ARG and MRG co-occurrence profiles of bacteria from different habitats were also examined. In accord with the evidence that humans were regularly exposed to antibiotic and metal-based compounds (Guillemot, 1999; Hambley, 2007; Jernberg *et al.*, 2010), ARG and MRG co-occurrence was remarkable in genomes derived from the human habitat, whereas it was much more moderate in both soil and water habitats (Figures 2c and d).

ARG diversity in close proximity to MRGs

Thus far, diverse ARGs associated with MRGs have been discovered, for example, coupling of resistance genes for methicillin with Zn (Cavaco *et al.*, 2010),

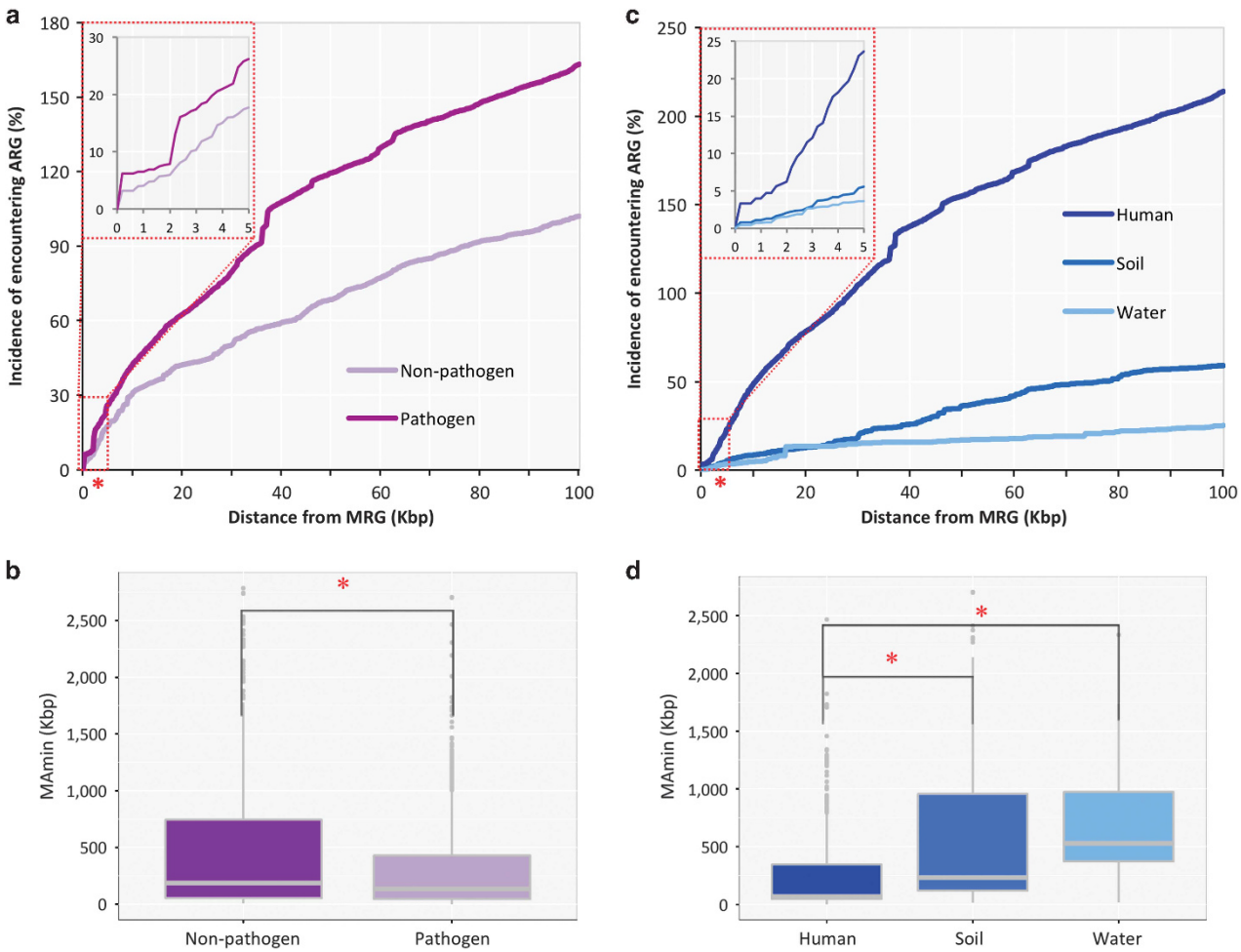

Figure 2 ARG and MRG co-occurrence by pathogenicity (**a** and **b**) and habitat (**c** and **d**). (**a** and **c**) Incidence of encountering ARGs along the distance from MRGs (**P*-value <0.05 determined by Fisher's exact test, otherwise no significance); (**b** and **d**) MetA_{min} index (**P*-value <0.01 determined by Student's *t*-test, otherwise no significance).

tetracycline with Cu (Amachawadia *et al.*, 2013) and multiple antibiotics with Hg (McIntosh *et al.*, 2008; Skurnik *et al.*, 2010; Rodríguez-Blanco *et al.*, 2012). However, there are still gaps in our knowledge of the general diversity pattern on a much broader scale, which was addressed in this study by determining the number of unique ARG subtypes detected along the distance from MRGs in different genome categories (Figure 3). A remarkable higher ARG diversity was revealed in genomes from pathogen species and human habitat, suggesting that MRGs co-occurred with a broader spectrum of ARGs in human-associated pathogens. As genes with closely related functions (e.g., functional interaction to confer resistance to specific antibiotics and metals) tend to be organized into clusters, which are maintained as adjacent genetic elements under strong evolutionary forces (Overbeek *et al.*, 1999; Fang *et al.*, 2008), the nearest ARGs for all MRG were summarized by type (Supplementary Figure S7) to reveal ARG–MRG pairs that tends to be functionally related. The top five ARG types (excluding multidrug and unclassified ARGs) most likely to be related with MRGs are beta-lactam, kasugamycin, bacitracin, aminoglycoside, polymyxin and tetracycline. In particular, the common MRG types had distinct preference for ARG types as their closest genetic neighbor (Figure 4). Specifically, with the exception of multidrug and the unclassified ARGs, Zn was more likely to co-occur with beta-lactam, bacitracin and polymyxin, Cu with beta-lactam, kasugamycin and bacitracin and As with beta-lactam, bacitracin and fosfomycin.

ARG and MRG co-occurrence structures differ based on habitats

Although genomes in different ecologies had distinct ARGs and MRGs co-occurrence profiles, the acquisition and spread of the coupled resistance genes could still cross the ecological boundaries at noticeable rates because of intense gene transfers and anthropogenic disturbances. Additionally, it has long been established that the ARGs harbored in environmental compartments alone could act as a potential reservoir for transfer to human pathogens (Martinez, 2008; Forsberg *et al.*, 2012; Perry and Wright, 2013), which might also be true for the coupled closest ARG–MRG pairs (the nearest ARG for each MRG, hereafter referred to as ARG–MRG) owing to their physical genetic linkages. Therefore, by examining ARG–MRG coupling profile, we were prompted to ascertain whether there is a significant genome grouping pattern by ecologies.

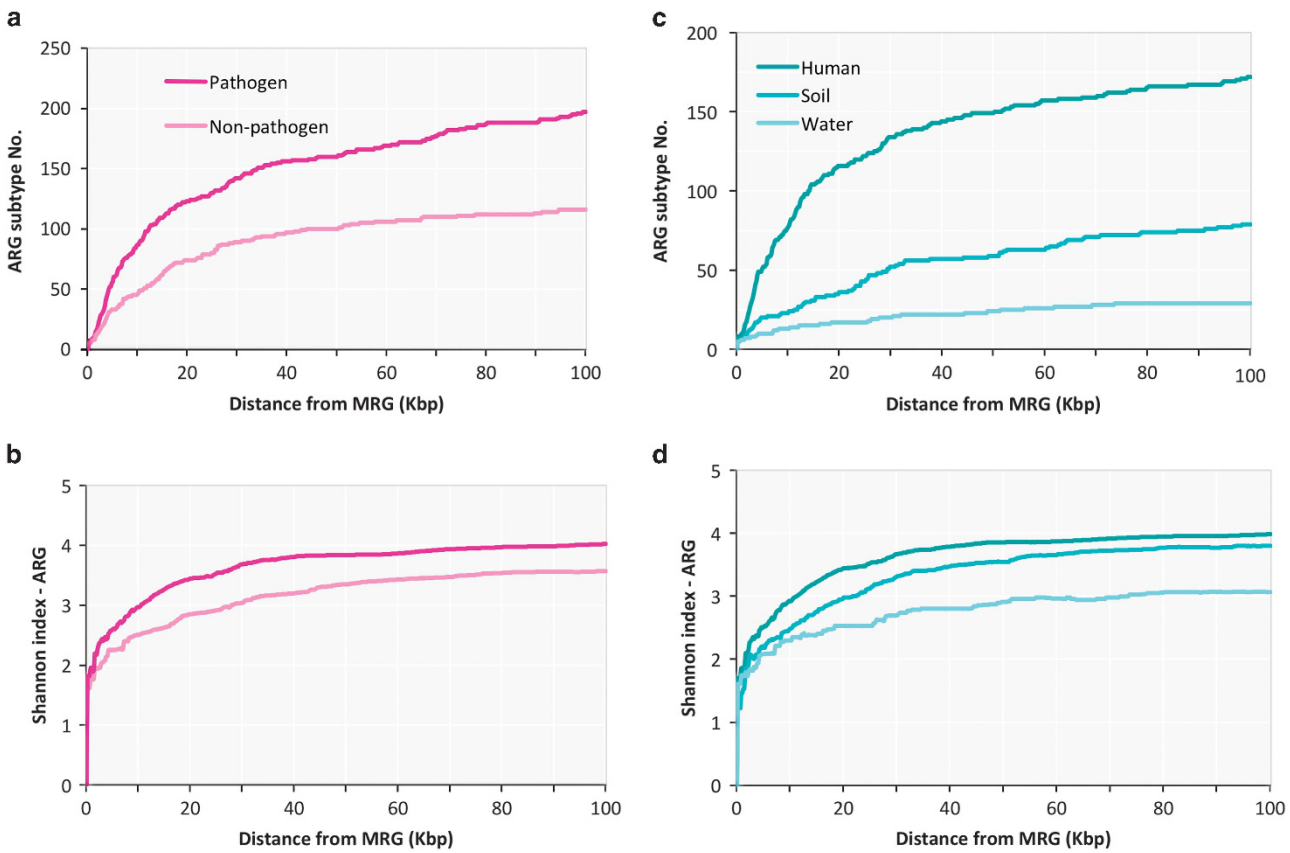

Figure 3 Representations of ARG diversity along the distance from the MRGs by pathogenicity and habitat. (**a** and **c**) The number of unique ARG subtypes along the distance from MRGs. (**b** and **d**) Shannon diversity scores of the ARG subtypes along the distance from the MRGs.

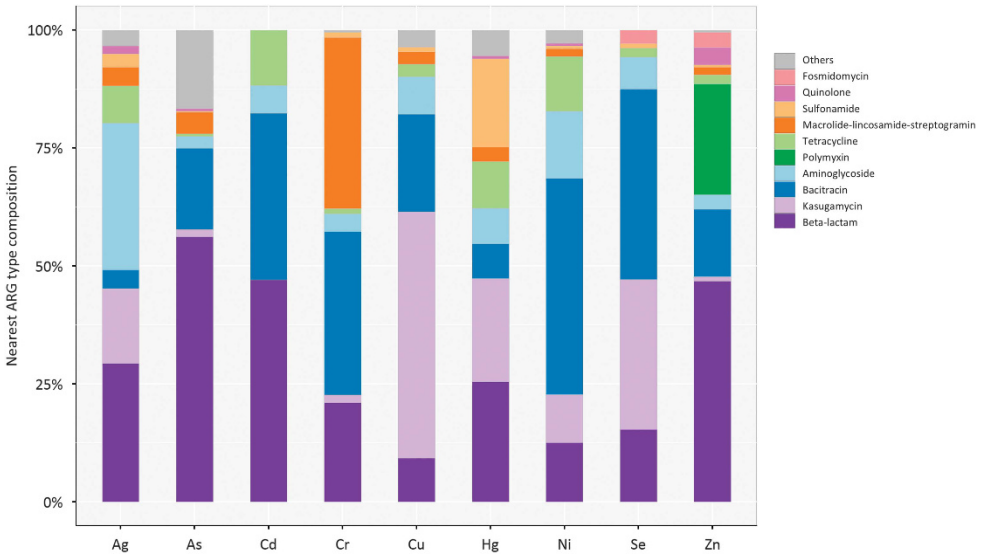

Figure 4 Preference of MRG for the closest ARG across the complete genomes (excluding multidrug and unclassified ARGs): composition of the nearest ARG from nine common MRG types.

A significant genome cluster by habitats was observed (*P*-value <0.001, ANOISM; Supplementary Figure S8), suggesting that although interhabitat resistance gene transfer has frequently been detected (Martinez *et al.*, 2015), these events have not homogenized the ARG–MRG coupling pattern among human, soil and water habitats. Instead, other ecological features among habitats might have a more substantial effect in shaping the resistome. For example, bacterial phylogeny has been evaluated as the primary determinant for the soil resistome (Forsberg *et al.*, 2014). And, much more frequent horizontal gene transfer, even among distant phylogenies, happened within similar environment than between different ecotypes (Beiko *et al.*, 2005; Soucy *et al.*, 2015). In contrast, a non-significant clustering between pathogens and non-pathogens was observed, suggesting that the dissemination of MRG-associated ARG was frequent enough to blur the resistome boundaries between pathogens and non-pathogens. In accordance with previous conclusion, under widespread antibiotic pressure, acquiring favorable genes, such as ARGs, from those non-pathogenic bacteria has markedly accelerated adaptation in pathogens (Perry and Wright, 2013).

Next, we sought to determine whether there were core coupling types across genomes. However, there was no single ARG–MRG pair that was shared among all the genomes, highlighting the extreme diversity of ARG–MRG association. There are 14 coupling types that were shared across $>25\%$ genomes, most of which were couples between multidrug/unclassified ARGs and the MRG types of Cu, Zn, Ni, Cr and Fe. The ARG–MRG pairs that were most likely to discriminate between habitats were determined using the supervised learning technique—Random Forest (Knights *et al.*, 2011). The results reveal that the significant genome cluster in habitats was mainly driven by the coupling occurrence between ARG types of bacitracin/multidrug and MRG types of Cu/Fe/As/Cr/Te/Mn/Sb/Al. Additionally, bacitracin-Cu and bacitracin-Cr were more frequently detected in soil and water rather than in the human-associated habitat.

Transfer potential of ARGs and MRGs

In addition to the co-occurrence of ARGs and MRGs, another concern for the two types of resistance genes is their cotransfer via collecting and recombining multiple extant resistance genes on mobile vehicles, thereby dramatically increasing their dissemination across environmental compartments. In fact, ARGs/MRGs have long been reported to reside on MGEs, including genes with resistance to almost all antibiotic families that are embedded in gene cassettes on integrons (Gaze *et al.*, 2011; Gillings *et al.*, 2015; Wales and Davies, 2015; Di Cesare *et al.*, 2016; Johnson *et al.*, 2016), and multiple resistance genes for commonly consumed antibiotics and metals that are located on plasmids (Carattoli, 2009; Zhang *et al.*, 2011; Popowska and Krawczyk-Balska, 2013; Wales and Davies, 2015; Johnson *et al.*, 2016). However, there is a lack of general information regarding the cotransfer potential of ARGs and MRGs in different populations, particularly those that are seriously impacted by human activities. Therefore, the mobility of both ARGs and MRGs within the genomes of different habitats and pathogenicities was investigated in this study in terms of the nearest distance from MGEs (Figure 5a). The distances of both the nearest ARGs and MRGs from the MGEs were much closer in bacteria from pathogen species and human habitat, reflecting the previous reported substantial capacity to develop resistance genes on MGEs in human pathogens (Perry and Wright, 2013; Gillings *et al.*, 2015; Wales and Davies, 2015). Indeed, equipped on the proper transfer machinery, the

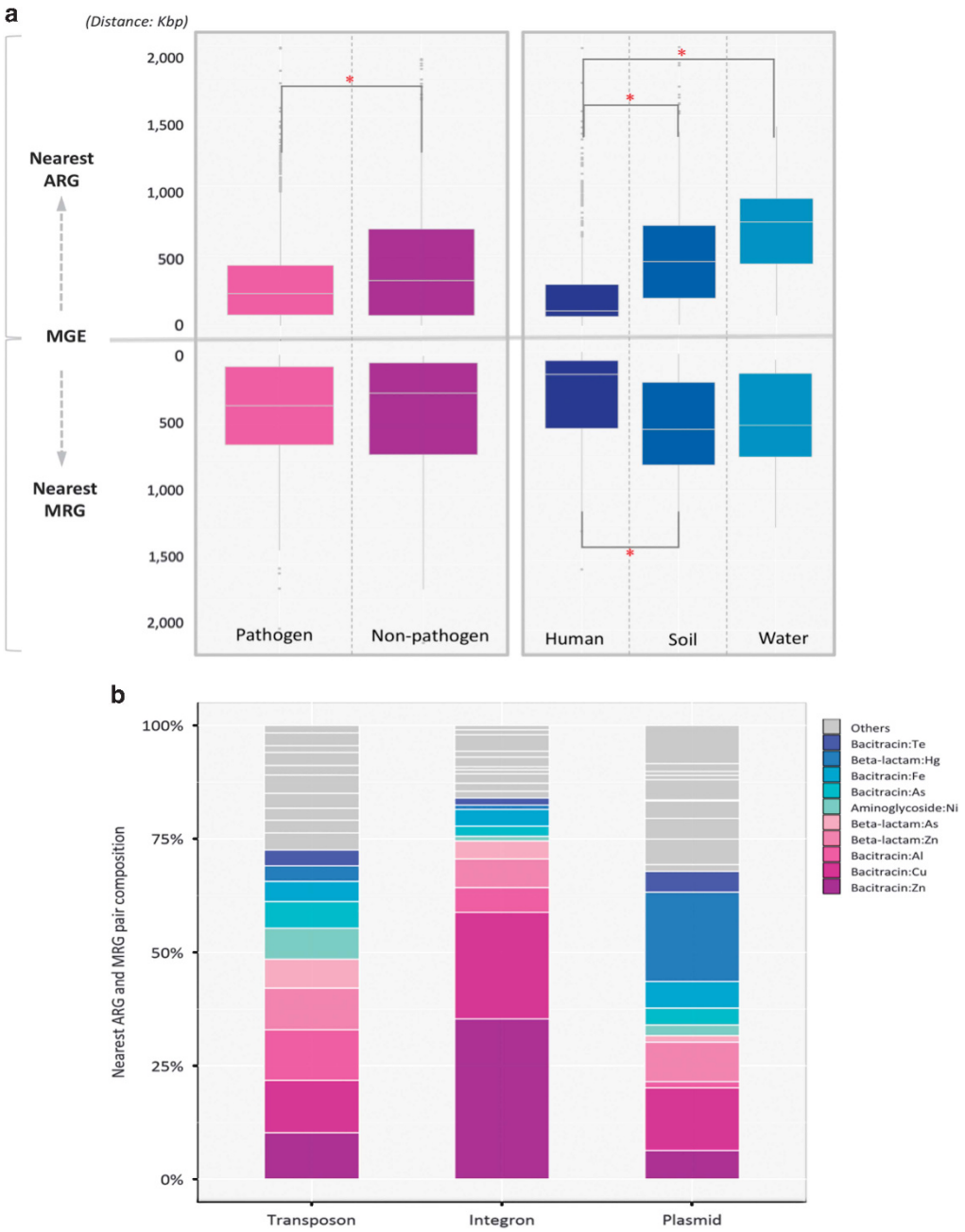

Figure 5 ARG and MRG cotransfer potential. (**a**) Boxplot of the average distance of both the nearest ARG and MRG from each MGE in each genome across the different categories. (*P-value<0.01 determined by Student's t-test, otherwise no significance). (**b**) Cotransfer preference of the three MGE types (transposon, integron and plasmid) for the nearest ARG and MRG pair across all complete genomes (top 10 in color).

efficient vehicles in gene shuffling, resistance gene dissemination can markedly be enhanced, benefiting the bacterial communities under environmental stress (Martinez *et al*., 2015; Soucy *et al*., 2015). Characterized by long exposure to selective pressure from antimicrobial and non-antimicrobial agents in human habitat, massive genetic exchanges have conferred efficient defense system to specific bacterial lineages (Jernberg *et al*., 2010; Juhas, 2015). In contrast to single fixed occurrence, the high cotransfer potential of both ARGs and MRGs could theoretically provide a greater ecological fitness, thus allowing the bacteria to adapt to environmental stress (Summers, 2002). Recognition of genetically linked transmissible resistance in human pathogens highlights the challenges of controlling resistome expansion in sites with human health importance, as antibiotic resistance can also be enriched through cotransfer with resistance genes for ubiquitous metal pressure.

In further analyses, the cotransfer machineries of both resistance gene types were investigated based on their transfer preference by transposons,

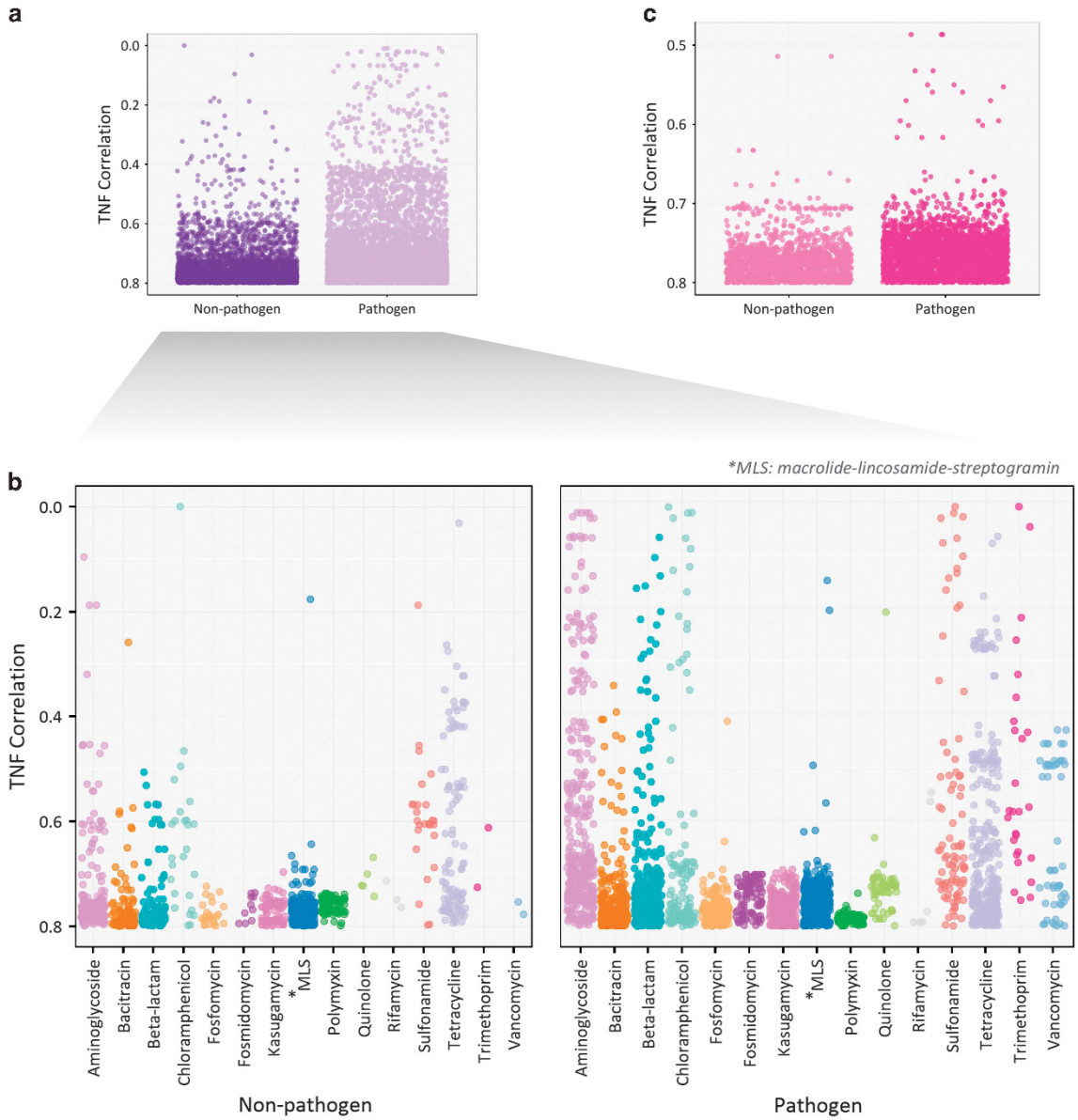

Figure 6 TNF correlation between a resistance gene carrying fragment and its global genome sequence in pathogen and non-pathogen groups. (**a** and **b**) TNF correlation profile of fragments with ARGs and (**c**) TNF correlation profile of fragments with both ARGs and MRGs.

integrons and plasmids (Figure 5b). Generally, a wide spectrum of ARG and MRG pairs might be cotransferred by any of the mobile machineries investigated in this study. These pairs include bacitracin and Cu/Zn/Al/Fe/Cr/As/Hg/Te, aminoglycoside and Ni/Fe/Zn, tetracycline and Zn/Fe, and beta-lactam and As/Zn/Hg. Although no obvious preference for transposons, the normal constituent of most bacterial genomes (Kleckner, 1981), was observed, integrons were much more likely to transfer pairs such as bacitracin and Cu/Zn, whereas plasmids were preferable to carry bacitracin and Cu/Zn and beta-lactam and Hg. It is of particular concern that the ARG and MRG pairs that were once carried by MGE types with high mobility potential, such as class 1 integrons or broad-host-range plasmids, can readily move across species or even phylum boundaries (Klumper *et al.*, 2015), thereby increasing the probability of spreading the resistance gene pairs to very diverse bacteria.

The above remarkable resistance mobility aroused our interest in further identifying the exogenous-originated resistance genes across the genome collection, especially those from human pathogens. As naturally occurring antibiotic resistance in bacteria have been evolved and selected under various offensive environmental conditions before the era of modern antibiotics, ARGs have been part of well-inherited genomic information (Allen *et al.*, 2010). By comparing TNF signature to genome-wide TNF pattern, it is not surprising to find that >60% of ARG/MRG-carrying fragments (single ARG or both

ARG and MRG could be allocated on the fragment) were endemic components within the genome (correlation coefficient $\geqslant 0.5$). More interestingly, compared with non-pathogen genomes, human pathogens contained more ARG/MRG-carrying fragments with divergent TNF signature from genome-wide TNF pattern (correlation coefficient <0.5) (Figure 6), suggesting the frequent incidence of exogenous acquisition of ARG/MRGs in human pathogens. Further categorization of those ARG-carrying exogenous fragments in pathogen genomes showed that they were mainly (65%) composed by the genes conferring resistance against aminoglycoside, beta-lactam, tetracycline and vancomycin. Unsurprisingly, long history exposure to the antibiotics of aminoglycoside, beta-lactam and tetracycline has shaped flexible gene transfer system in pathogens to tackle with selective pressure imposed by these common antibiotics (Davies and Davies, 2010). The phenomenon with particular concern is that the resistance genes for vancomycin, the last resort medicine to defend against infections caused by Gram-positive bacteria, was also found to be exogenous in human pathogen, suggesting a recent acquisition of these genes likely from environmental sources. This frequent horizontal gene transfer incidence among community, together with previously observed high abundance of ARGs against vancomycin in human feces samples (Li *et al.*, 2015), implied rapid adaptation towards newly introduced antibiotics of human pathogens under current infectious treatment. In addition, majority of these exogenous ARG-carrying fragments (68%) from clinically important pathogen species of *Escherichia coli*, *Shigella flexneri*, *Acinetobacter baumannii* and *Enterococcus faecium*, suggesting gene transfer, has had crucial role in adaptation to antibiotic stress in these pathogen species (Dijkshoorn *et al.*, 2007; Stokes and Gillings, 2011; Didelot *et al.*, 2016). Overall, both genetic linkage and composition-based analysis revealed surprisingly high extent and relevance of horizontal gene transfer events to resistome development in human-associated bacteria, especially human pathogens.

In future studies, the resistance phenotypes of bacteria should be extensively investigated to provide an actual reference for the coselection potential in genotype. Moreover, the affiliated genes within the neighborhood of the co-occurring ARGs and MRGs should also be taken into account to understand other genetic factors that are responsible for antibiotic and metal resistance coselection. In addition to metals, there are increasing evidences of the ability of other non-antibiotic agents (e.g., detergents and disinfectants) to select ARGs (Gaze *et al.*, 2011; Wellington *et al.*, 2013; Forbes *et al.*, 2016; Hartmann *et al.*, 2016). One example of particular concern is the promoter role of quaternary ammonium compound (widely used detergent and disinfectant) in antibiotic resistance coselection through either gene cluster on integrons (Gaze *et al.*, 2011) or the overexpression of common efflux pumps (Tezel and Pavlostathis, 2015). Therefore, identifying all potential coselection agents and their roles in antibiotic resistance dissemination in human-associated environment is necessary, which will contribute to risk assessment of antibiotic resistance under current clinical/environmental management.

Conclusions

Overall, by integrating powerful genomic analyses with a large complete genome collection, the ARG and MRG co-occurrence was investigated from multiple aspects. The results provide insight into the ARG and MRG co-occurrence profile in distinct genome categories, where the intimate genetic linkage of the two resistance types and their high cotransfer potential were observed in genomes from pathogen species and human habitat. Our results also revealed the importance of horizontal gene transfer in shaping resistome of human commensal flora, especially clinically important pathogens, and highlighted its profound health impact on modern antibiotic usage.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

We thank the Hong Kong General Research Fund (172099-/14E) for providing the financial support for this study. Dr Li-Guan Li and Dr Yu Xia thank The University of Hong Kong for providing the postdoctoral fellowships.

References

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.

Amachawadia RG, Scotta HM, Alvaradob CA, Maininia TR, Vinascoa J, Drouillardb JS *et al.* (2013). Occurrence of the transferable copper resistance gene tcrB among fecal enterococci of U.S. feedlot cattle fed copper-supplemented diets. *Appl Environ Microbiol* **79**: 4369–4375.

Amsaveni R, Sureshkumar M, Reshma J, Mary UI, Vivekanandhan G. (2015). Screening of multi-drug and metal resistant *Aeromonas* species from diverse sources. *Am J Infect Dis* **11**: 41–47.

Andersson DI, Hughes D. (2010). Antibiotic resistance and its cost: is it possible to reverse resistance? *Nat Rev Microbiol* **8**: 260–271.

Andersson DI, Hughes D. (2014). Microbiological effects of sublethal levels of antibiotics. *Nat Rev Microbiol* **12**: 465–478.

Baker-Austin C, Wright MS, Stepanauskas R, McArthur JV. (2006). Co-selection of antibiotic and metal resistance. *Trends Microbiol* **14**: 176–182.

Beiko RG, Harlow TJ, Ragan MA. (2005). Highways of gene sharing in prokaryotes. *Proc Natl Acad Sci USA* **102**: 14332–14337.

Broaders E, Gahan CG, Marchesi JR. (2013). Mobile genetic elements of the human gastrointestinal tract: potential for spread of antibiotic resistance genes. *Gut Microbes* **4**: 271–280.

Cai L, Yu K, Yang Y, Chen BW, Li XD, Zhang T. (2013). Metagenomic exploration reveals high levels of microbial arsenic metabolism genes in activated sludge and coastal sediments. *Appl Microbiol Biotechnol* **97**: 9579–9588.

Carattoli A. (2009). Resistance plasmid families in *Enterobacteriaceae*. *Antimicrob Agents Chemother* **53**: 2227–2238.

Cavaco LM, Hasman H, Stegger M, Andersen PS, Skov R, Fluit AC *et al.* (2010). Cloning and occurrence of czrC, a gene conferring cadmium and zinc resistance in methicillin-resistant *Staphylococcus aureus* CC398 isolates. *Antimicrob Agents Chemother* **54**: 3605–3608.

Charlesworth S, De Miguel E, Ordonez A. (2011). A review of the distribution of particulate trace elements in urban terrestrial environments and its application to considerations of risk. *Environ Geochem Health* **33**: 103–123.

Chen H, Teng Y, Lu S, Wang Y, Wang J. (2015). Contamination features and health risk of soil heavy metals in China. *Sci Total Environ* **512-513**: 143–153.

Daims H, Lebedeva EV, Pjevac P, Han P, Herbold C, Albertsen M *et al.* (2015). Complete nitrification by *Nitrospira* bacteria. *Nature* **528**: 504–509.

Davies J, Davies D. (2010). Origins and evolution of antibiotic resistance. *Microbiol Mol Biol Rev* **74**: 417–433.

Davis IJ, Roberts AP, Ready D, Richards H, Wilson M, Mullany P. (2005). Linkage of a novel mercury resistance operon with streptomycin resistance on a conjugative plasmid in *Enterococcus faecium*. *Plasmid* **54**: 26–38.

Di Cesare A, Eckert EM, D'Urso S, Bertoni R, Gillan DC, Wattiez R *et al.* (2016). Co-occurrence of integrase 1, antibiotic and heavy metal resistance genes in municipal wastewater treatment plants. *Water Res* **94**: 208–214.

Didelot X, Walker AS, Peto TE, Crook DW, Wilson DJ. (2016). Within-host evolution of bacterial pathogens. *Nat Rev Microbiol* **14**: 150–162.

Dijkshoorn L, Nemec A, Seifert H. (2007). An increasing threat in hospitals: multidrug-resistant *Acinetobacter baumannii*. *Nat Rev Microbiol* **5**: 939–951.

Fang G, Rocha EPC, Danchin A. (2008). Persistence drives gene clustering in bacterial genomes. *BMC Genomics* **9**: 4.

Fischbach MA, Walsh CT. (2009). Antibiotics for emerging pathogens. *Science* **325**: 1089–1093.

Forbes S, Knight CG, Cowley NL, Amezquita A, McClure P, Humphreys G *et al.* (2016). Variable effects of exposure to formulated microbicides on antibiotic susceptibility in *Firmicutes* and *Proteobacteria*. *Appl Environ Microbiol* **82**: 3591–3598.

Forsberg KJ, Patel S, Gibson MK, Lauber CL, Knight R, Fierer N *et al.* (2014). Bacterial phylogeny structures soil resistomes across habitats. *Nature* **509**: 612–616.

Forsberg KJ, Reyes A, Wang B, Selleck EM, Sommer MO, Dantas G. (2012). The shared antibiotic resistome of soil bacteria and human pathogens. *Science* **337**: 1107–1111.

Frost LS, Leplae R, Summers AO, Toussaint A. (2005). Mobile genetic elements: the agents of open source evolution. *Nat Rev Microbiol* **3**: 722–732.

Gaze WH, Zhang L, Abdouslam NA, Hawkey PM, Calvo-Bado L, Royle J *et al.* (2011). Impacts of anthropogenic activity on the ecology of class 1 integrons and integron-associated genes in the environment. *ISME J* **5**: 1253–1261.

Gibson MK, Forsberg KJ, Dantas G. (2015). Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. *ISME J* **9**: 207–216.

Gillings MR, Gaze WH, Smalla K, Tiedje JM, Zhu Y-G. (2015). Using the class 1 integron-integrase gene as a proxy for anthropogenic pollution. *ISME J* **9**: 1269–1279.

Guillemot D. (1999). Antibiotic use in humans and bacterial resistance. *Curr Opin Microbiol* **2**: 494–498.

Hambley TW. (2007). Developing new metal-based therapeutics: challenges and opportunities. *Dalton Trans* **43**: 4929–4937.

Hartmann EM, Hickey R, Hsu T, Betancourt Roman CM, Chen J, Schwager R *et al.* (2016). Antimicrobial chemicals are associated with elevated antibiotic resistance genes in the indoor dust microbiome. *Environ Sci Technol* **50**: 9807–9815.

Huijbers PM, Blaak H, de Jong MC, Graat EA, Vandenbroucke-Grauls CM, de Roda Husman AM. (2015). Role of the environment in the transmission of antimicrobial resistance to humans: a review. *Environ Sci Technol* **49**: 11993–12004.

Jernberg C, Lofmark S, Edlund C, Jansson JK. (2010). Long-term impacts of antibiotic exposure on the human intestinal microbiota. *Microbiology* **156**: 3216–3223.

Ji X, Shen Q, Liu F, Ma J, Xu G, Wang Y *et al.* (2012). Antibiotic resistance gene abundances associated with antibiotics and heavy metals in animal manures and agricultural soils adjacent to feedlots in Shanghai; China. *J Hazard Mater* **235-236**: 178–185.

Johnson TA, Stedtfeld RD, Wang Q, Cole JR, Hashsham SA, Looft T *et al.* (2016). Clusters of antibiotic resistance genes enriched together stay together in swine agriculture. *MBio* **7**: e02214–e02215.

Juhas M. (2015). Horizontal gene transfer in human pathogens. *Crit Rev Microbiol* **41**: 101–108.

Kleckner N. (1981). Transposable elements in prokaryotes. *Annu Rev Genet* **15**: 341–404.

Klumper U, Riber L, Dechesne A, Sannazzarro A, Hansen LH, Sorensen SJ *et al.* (2015). Broad host range plasmids can invade an unexpectedly diverse fraction of a soil bacterial community. *ISME J* **9**: 934–945.

Knights D, Costello EK, Knight R. (2011). Supervised classification of human microbiota. *FEMS Microbiol Rev* **35**: 343–359.

Li B, Yang Y, Ma LP, Ju F, Guo F, Tiedje JM *et al.* (2015). Metagenomic and network analysis reveal wide distribution and co-occurrence of environmental antibiotic resistance genes. *ISME J* **9**: 2490–2502.

Li LG, Cai L, Zhang XX, Zhang T. (2014). Potentially novel copper resistance genes in copper-enriched activated sludge revealed by metagenomic analysis. *Appl Microbiol Biotechnol* **98**: 10255–10266.

Loman NJ, Pallen MJ. (2015). Twenty years of bacterial genome sequencing. *Nat Rev Microbiol* **13**: 787–794.

Losada L, DebRoy C, Radune D, Kim M, Sanka R, Brinkac L *et al.* (2016). Whole genome sequencing of diverse Shiga toxin-producing and non-producing *Escherichia coli* strains reveals a variety of virulence and novel antibiotic resistance plasmids. *Plasmid* **83**: 8–11.

Ma L, Xia Y, Li B, Yang Y, Li LG, Tiedje JM *et al.* (2016). Metagenomic assembly reveals hosts of antibiotic resistance genes and the shared resistome in pig, chicken, and human feces. *Environ Sci Technol* **50**: 420–427.

Martinez JL. (2008). Antibiotics and antibiotic resistance genes in natural environments. *Science* **321**: 365–367.

Martinez JL, Coque TM, Baquero F. (2015). What is a resistance gene? Ranking risk in resistomes. *Nat Rev Microbiol* **13**: 116–123.

McIntosh D, Cunningham M, Ji B, Fekete FA, Parry EM, Clark SE *et al.* (2008). Transferable, multiple antibiotic and mercury resistance in Atlantic Canadian isolates of *Aeromonas salmonicida* subsp. *salmonicida* is associated with carriage of an IncA/C plasmid similar to the *Salmonella enterica* plasmid pSN254. *J Antimicrob Chemother* **61**: 1221–1228.

Overbeek R, Fonstein M, D'Souza M, Pusch GD, Maltsev N. (1999). The use of gene clusters to infer functional coupling. *Proc Natl Acad Sci USA* **96**: 2896–2901.

Pal C, Bengtsson-Palme J, Kristiansson E, Larsson DGJ. (2015). Co-occurrence of resistance genes to antibiotics, biocides and metals reveals novel insights into their co-selection potential. *BMC Genomics* **16**: 964.

Pal C, Bengtsson-Palme J, Rensing C, Kristiansson E, Larsson DGJ. (2014). BacMet: antibacterial biocide and metal resistance genes database. *Nucleic Acids Res* **42**: D737–D743.

Pamer EG. (2016). Resurrecting the intestinal microbiota to combat antibiotic-resistant pathogens. *Science* **352**: 535–538.

Perry JA, Wright GD. (2013). The antibiotic resistance 'mobilome': searching for the link between environment and clinic. *Front Microbiol* **4**: 138.

Popowska M, Krawczyk-Balska A. (2013). Broad-host-range IncP-1 plasmids and their resistance potential. *Front Microbiol* **4**: 44.

Rodríguez-Blanco A, Lemos ML, Osorio CR. (2012). Integrating conjugative elements as vectors of antibiotic, mercury, and quaternary ammonium compound resistance in marine aquaculture environments. *Antimicrob Agents Chemother* **56**: 2619–2626.

Rosewarne CP, Pettigrove V, Stokes HW, Parsons YM. (2010). Class 1 integrons in benthic bacterial communities: abundance, association with Tn402-like transposition modules and evidence for coselection with heavy-metal resistance. *FEMS Microbiol Ecol* **72**: 35–46.

Sandegren L, Andersson DI. (2009). Bacterial gene amplification: implications for the evolution of antibiotic resistance. *Nat Rev Microbiol* **7**: 578–588.

Seiler C, Berendonk TU. (2012). Heavy metal driven co-selection of antibiotic resistance in soil and water bodies impacted by agriculture and aquaculture. *Front Microbiol* **3**: 399.

Skurnik D, Ruimy R, Ready D, Ruppe E, Bernede-Bauduin C, Djossou F *et al.* (2010). Is exposure to mercury a driving force for the carriage of antibiotic resistance genes? *J Med Microbiol* **59**: 804–807.

Soucy SM, Huang J, Gogarten JP. (2015). Horizontal gene transfer: building the web of life. *Nat Rev Genet* **16**: 472–482.

Stokes HW, Gillings MR. (2011). Gene flow, mobile genetic elements and the recruitment of antibiotic resistance genes into gram-negative pathogens. *FEMS Microbiol Rev* **35**: 790–819.

Summers AO. (2002). Generally overlooked fundamentals of bacterial genetics and ecology. *Clin Infect Dis* **34**: S85–S92.

Tezel U, Pavlostathis SG. (2015). Quaternary ammonium disinfectants: microbial adaptation, degradation and ecology. *Curr Opin Biotechnol* **33**: 296–304.

Wales AD, Davies RH. (2015). Co-selection of resistance to antibiotics, biocides and heavy metals, and its relevance to foodborne pathogens. *Antibiotics* **4**: 567–604.

Wellington EMH, Boxall ABA, Cross P, Feil EJ, Gaze WH, Hawkey PM *et al.* (2013). The role of the natural environment in the emergence of antibiotic resistance in gram-negative bacteria. *Lancet Infect Dis* **13**: 155–165.

WHO. (2014). *Antimicrobial Resistance: Global Report on Surveillance*. WHO: Geneva, Switzerland.

Wireman J, Liebert CA, Smith T, Summers AO. (1997). Association of mercury resistance with antibiotic resistance in the Gram-negative fecal bacteria of primates. *Appl Environ Microbiol* **63**: 4494–4503.

Woolhouse ME, Gowtage-Sequeria S. (2005). Host range and emerging and reemerging pathogens. *Emerg Infect Dis* **11**: 1842–1847.

Wright GD. (2010). Antibiotic resistance in the environment: a link to the clinic? *Curr Opin Microbiol* **13**: 589–594.

Yang Y, Jiang X, Chai B, Ma L, Li B, Zhang A *et al.* (2016). ARGs-OAP: online analysis pipeline for antibiotic resistance genes detection from metagenomic data using an integrated structured ARG-database. *Bioinformatics* **32**: 2346–2351.

Yang Y, Li B, Ju F, Zhang T. (2013). Exploring variation of antibiotic resistance genes in activated sludge over a four-year period through a metagenomic approach. *Environ Sci Technol* **47**: 10197–10205.

Zhang T, Zhang XX, Ye L. (2011). Plasmid metagenome reveals high levels of antibiotic resistance genes and mobile genetic elements in activated sludge. *PLoS One* **6**: e26041.

Zhu YG, Johnson TA, Su JQ, Qiao M, Guo GX, Stedtfeld RD *et al.* (2013). Diverse and abundant antibiotic resistance genes in Chinese swine farms. *Proc Natl Acad Sci USA* **110**: 3435–3440.

Zowawi HM, Harris PN, Roberts MJ, Tambyah PA, Schembri MA, Pezzani MD *et al.* (2015). The emerging threat of multidrug-resistant gram-negative bacteria in urology. *Nat Rev Urol* **12**: 570–584.

Supplementary Information accompanies this paper on The ISME Journal website (http://www.nature.com/ismej)