FEMS Microbiology Ecology, 92, 2016, fiv168

doi: 10.1093/femsec/fiv168
Advance Access Publication Date: 6 January 2016
Research Article

Downloaded from https://academic.oup.com/femsec/article-abstract/92/2/fiv168/2484737 by Maynooth University user on 15 May 2019

RESEARCH ARTICLE

Antibiotic resistance genes across a wide variety of metagenomes

David Fitzpatrick[†] and Fiona Walsh[*†]

Department of Biology, Maynooth University, Maynooth, Co. Kildare, W23 F2H6, Ireland

[*]**Corresponding author:** Department of Biology, Maynooth University, Maynooth, Co. Kildare, W23 F2H6, Ireland. Tel: +353-14747246; E-mail: Fiona.walsh@nuim.ie
[†]Both contributed equally to this work.

One sentence summary: Distribution and relative abundances of antibiotic resistance genes across the world from the ocean to the human gut and everywhere in between.

Editor: Kornelia Smalla

Fiona Walsh, http://orcid.org/0000-0003-0789-1689

ABSTRACT

The distribution of potential clinically relevant antibiotic resistance (AR) genes across soil, water, animal, plant and human microbiomes is not well understood. We aimed to investigate if there were differences in the distribution and relative abundances of resistance genes across a variety of ecological niches. All sequence reads (human, animal, water, soil, plant and insect metagenomes) from the MG-RAST database were downloaded and assembled into a local sequence database. We show that there are many reservoirs of the basic form of resistance genes e.g. bla_{TEM}, but the human and mammalian gut microbiomes contain the widest diversity of clinically relevant resistance genes using metagenomic analysis. The human microbiomes contained a high relative abundance of resistance genes, while the relative abundances varied greatly in the marine and soil metagenomes, when datasets with greater than one million genes were compared. While these results reflect a bias in the distribution of AR genes across the metagenomes, we note this interpretation with caution. Metagenomics analysis includes limits in terms of detection and identification of AR genes in complex and diverse microbiome population. Therefore, if we do not detect the AR gene is it in fact not there or just below the limits of our techniques?

Keywords: resistome; microbiome; metagenome; soil; animal; gut

INTRODUCTION

Antibiotic resistance (AR) genes are widespread and have been identified in almost every microbiome throughout the planet from Antarctic ice to the phyllosphere, soil, insects, animals and humans (Feller, Sonnet and Gerday 1995; de Kraker and van de Sande-Bruinsma 2007; Allen *et al.* 2009; Cantas *et al.* 2013; Walsh and Duffy 2013; Duffy, Holliger and Walsh 2014; Garcia-Migura *et al.* 2014). Through the identification of AR genes, or putative resistance genes, many scientific studies have come to the following conclusions:

(i) Soil is a major reservoir of AR genes (Finley *et al.* 2013; Martínez 2008; Forsberg *et al.* 2014; Nesme *et al.* 2014; Su *et al.* 2014b).
(ii) Animals, especially food animals, are sources of AR genes due to antibiotic use in agriculture (Wang *et al.* 2012; Andersson and Hughes 2014).
(iii) The spreading of animal manure on land increases the abundance and diversity of transferable resistance genes present (Heuer, Schmitt and Smalla 2011; Jechalke *et al.* 2013).
(iv) Fecal pollution of water causes AR gene pollution of freshwater and marine biomes (Alves *et al.* 2014; Jia *et al.* 2014).
(v) Antibiotic use in clinical settings has created selective environments for the proliferation of antibiotic-resistant pathogens (Bacquero *et al.* 1998).

Advances in next-generation sequencing have led to an explosion of data from a wide variety of microbiomes. These

Received: 11 May 2015; **Accepted:** 18 December 2015
© FEMS 2016. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

advances have also enabled research in the non-human and non-animal microbiomes to extend beyond the culturable bacteria. Thus, enabling microbiologists to investigate the roles of soil, water and plant microbiomes as reservoirs of AR genes. However, the data analysed must be placed in the relevant context. While no environment is completely free from AR genes, it is vital to differentiate between the clinically relevant resistance genes, in particular transferable resistance genes and resistance genes, which are intrinsic to specific bacterial species. This will then aid in the identification of the ecological roles of clinically relevant AR genes, how they have been selected in different biomes and the risks they pose to the treatment of infections. In order to minimize the risks to human health posed by AR genes in the environment and the risk to the environment by waste containing AR genes, we must identify the critical points of control, which are the AR gene hotspots (Berendonk *et al.* 2015). Our study utilized the freely available metagenome data to investigate the ubiquitous resistance genes, and those, which are microbiome specific in order to identify potential hotspots of resistance.

MATERIALS AND METHODS

Metagenomics sequence data

Environmental shotgun sequencing metagenomic datasets were downloaded from the MG-RAST metagenomic analysis server (http://metagenomics.anl.gov/) (Meyer *et al.* 2008). Sequence reads were annotated locally and combined into a local sequence database. Redundant sequences were removed from appropriate datasets. Overall, our database contains data from 10 181 metagenomic surveys encompassed within 432 metagenome datasets (Dataset 1, Supporting Information). The total size of the database is approximately 3160 Gbp.

BLAST searches for AR genes

A local file of AR genes was created by downloading all sequence data from the Antibiotic Resistance Database (ARDB) (Liu and Pop 2009). The current version of ARDB contains 23 137 sequences classified into 380 AR gene types encoding resistance to 249 antibiotics (Liu and Pop 2009). To avoid redundancy, a single representative AR gene was selected for each AR gene type, in addition the bla_{NDM} gene sequence was included as this was absent from the ARDB (Dataset 2, Supporting Information). All 381 AR genes were treated as query sequences and compared to our local metagenomic database using BLASTp with a cut-off expectation (E) value of $<10^{-10}$ (Altschul *et al.* 1997). A read was annotated as a potential AR gene if its top high-scoring segment pair (HSP) had a percent identity >90% and the HSP length was greater than 60% of the query AR gene. The total number of reads that pass these criteria for each AR gene was recorded and labeled as AR reads. The relative abundance for each AR gene in the metagenome was calculated by dividing the specific number of AR reads by the total number of reads in that metagenome.

RESULTS

Overall, our database contains data from 10 181 metagenomic surveys encompassed within 432 metagenome datasets from a wide variety of human, animal, plant and environmental sources (Dataset 1, Supporting Information). The metadata associated with some of the MG-RAST datasets are poorly described e.g. describing both the plant and rhizosphere in one project. Therefore, we have described the overall patterns regarding the biomes rather than compare within the biomes. In order to minimize stringency and expression bias, a read was annotated as an AR gene if its top HSP had a percent identity >90% and the HSP length was greater than 60% of the query AR gene.

Efflux resistance genes are present in commensal and pathogenic bacteria as well as phytobacteria and environmental bacteria and contribute to the resistance armory of clinical pathogens and defence mechanisms of non-pathogenic bacteria. However, it is the upregulation of expression of the efflux pumps, most frequently associated with a mutation in a repressor gene that results in the resistance phenotype, rather than the presence of the gene. The wild-type version of these genes should not confer resistance following transfer to susceptible bacteria (Martínez, Coque and Baquero 2015). Therefore, they should not be considered a real risk when identified in metagenomic datasets (Martínez, Coque and Baquero 2015). Therefore, we analysed the data in terms of non-efflux-mediated resistance to identify AR genes, which have the greatest probabilities of transferring to pathogenic bacteria. While some genes encoding an AR efflux pump e.g. *tetA*, have been detected on mobile elements, the vast majority of genes encoding efflux pumps are not transferable. Other intrinsic resistance genes e.g. chromosomally encoded beta-lactamases remain in the dataset as these have been shown to transfer from the chromosome to mobile elements. When these mobile elements are present in pathogenic bacteria, they confer a resistance phenotype.

The microbiomes containing the highest proportion of efflux resistance mechanisms were soils, the Giant African snail and the phyllosphere/rhizosphere (Table 1). There was agreement between the diversity and abundance data in almost all projects. However, in a small number of projects the data did not correlate. In the polluted soil project (MG RAST ID 3266) and the Western English Channel (MG RAST ID 109), the contribution of efflux to the resistance profile in terms of gene abundance were 10.5% and 0.7%, respectively, whereas in terms of diversity of resistance genes it was 56% and 33%, respectively. This suggests that while the varieties of efflux resistance genes were high, the abundance of the efflux genes was low. The converse was also identified e.g. in the canine intestinal microbiome (MG RAST ID 103) and one human gut microbiome project (MG RAST ID 98) efflux constituted 24% and 23% of the variety of resistance genes in these projects, but 43% and 39% of the total resistance gene abundances, respectively.

This study investigated if the diversity and abundance of potential AR genes, which confer a risk to clinical pathogen treatment, vary across metagenomes from a wide variety of biomes. The first section investigated the relative abundance of non-efflux resistance genes and the second the variety of non-efflux resistance genes. In order to minimize the potential bias posed by the variations in the total gene content of the different metagenomic projects, we analysed all metagenome projects for the relative abundance of resistance genes. The projects with greater than one million genes were chosen (Fig. 1, Table 2), and these comprised projects from a wide variety of biomes. The relative abundances of resistance genes in the human projects were the least variable, which could be due to the fact that the human gut microbiome contains the least variability in terms of bacterial species present. Those of the human and mammalian projects were also the highest when taken together to represent the biomes. There were large variations in the relative abundances of the resistance genes present in the marine water and soil projects. The highest peaks in relative abundance in the marine water (MG RAST ID 109) and the WWTP PCE-dechlorinating

Downloaded from https://academic.oup.com/femsec/article-abstract/92/2/fiv168/2484737 by Maynooth University user on 15 May 2019

Table 1. The contribution of efflux resistance genes to the total resistance diversity and abundances.

MG-RAST Project ID	Metagenome project	Total number of different AR genes	Number of different non-efflux AR genes	Number of different efflux AR genes	Efflux gene diversity as % of total resistance gene diversity	Abundance of efflux genes as % of total resistance genes
5324	Achatina (Giant African) snail	27	7	20	74	74
611	Project for: 5mgHCH soil	21	6	15	71	76
729	Project for: 70mgHCH soil	23	7	16	70	79
2766	Phyllosphere/rhizosphere	22	8	14	64	77
3266	GenoRem_GH_MT (polluted soil)	18	8	10	56	10.5
2850	EarlhamMetagenomes2012 (farm soil)	18	9	9	50	70
3997	Sao Paulo Zoo Composting	45	23	22	49	33
4778	Healthy human gut microbial metagenomic study	35	19	16	46	43
6298	Weddell's Saddleback Tamarin Fecal Metagenome	20	11	9	45	51
2321	Comparative metagenomics of two microbial mats at Cuatro Cienegas basin	7	4	3	43	44
5130	Wild mice enterotypes	43	25	18	42	44
10	Twin Gut Microflora Study	49	29	20	41	29
29	Human Faeces - Kurokawa	40	24	16	40	31
109	Western English Channel Time Series	12	8	4	33	0.7
115	muegge_mammals_paper_mammal_samples_MG	62	42	20	32	48
154	Gut microbiota in the Irish Elderly and its links to health and diet	61	41	19	31	44
2558	Functional metagenomic profiling of intestinal microbiome in extreme ageing	49	35	14	29	29
730	Project for: 124HCHsoil	7	5	2	29	67
27	Global Ocean Sampling Expedition	19	14	5	26	16
582	Project for: DPI_rumen_454_bovine_and_ovine_combined	8	6	2	25	14
103	Canine Intestinal Microbiome	17	13	4	24	43
98	Human gut microbiome viewed across age and geography (WGS)	101	78	23	23	39
128	The oral metagenome in health and disease	19	15	4	21	19
60	Human Gut Microbial Metagenome	60	49	11	18	16
2982	SEPRL/PoultryGut2012	6	5	1	17	1.3
5687	Circadian_mouse_cecum	13	11	2	15	42
731	WWTP-activated sludge plasmid (and viral) metagenomes from Switzerland	19	17	2	11	15
1167	BIOS Microbial Oceanography Course	3	2	1	33	33
2706	Puget Sound	4	2	2	50	33
4970	Asian long horned beetle larval midgut metagenome	4	2	2	50	40
5736	Deep biosphere	3	3	0	0	0
20	Antarctica Aquatic Microbial Metagenome	3	3	0	0	0
3744	Freshwater_Lagoon_Albufera	1	1	0	0	0
1449	TNPRC Rhesus monkey	1	1	0	0	0
1452	NEPRC_Macaque_64_weeks	3	3	0	0	0
2374	GED Great Prairie Iowa Assembly	2	2	0	0	0
405	Rothamsted grassland soil	3	3	0	0	0
922	Metagenomics data WWTP	6	6	0	0	0

Only metagenome projects with greater than 500 000 genes were included in this table.

mixed culture (MG RAST ID 602) projects were due to the presence of the bla_{TEM} gene. However, the relative abundances of other projects were not as highly influenced by the bla_{TEM} resistance gene e.g. the human projects. The results of the relative diversity study concur with the variety of resistance genes study. The greatest variety and abundances were detected in human gut metagenomes. However, when analysing metagenomic data it is important to consider the limitations of sequencing in relation to the depth currently possible and the actual depth required to identify low diversity genes in the dataset.

Non-efflux-mediated resistance genes to all known classes of antibiotics were detected in at least one metagenome (Fig. 2 and Fig. S1, Supporting Information). The human gut microbiomes contained the widest variety of AR genes (n = 29–78, Table 2; Fig. S1, Supporting Information). The human oral metagenomes, however, contained at most 15 different AR genes

Downloaded from https://academic.oup.com/femsec/article-abstract/92/2/fiv168/2484737 by Maynooth University user on 15 May 2019

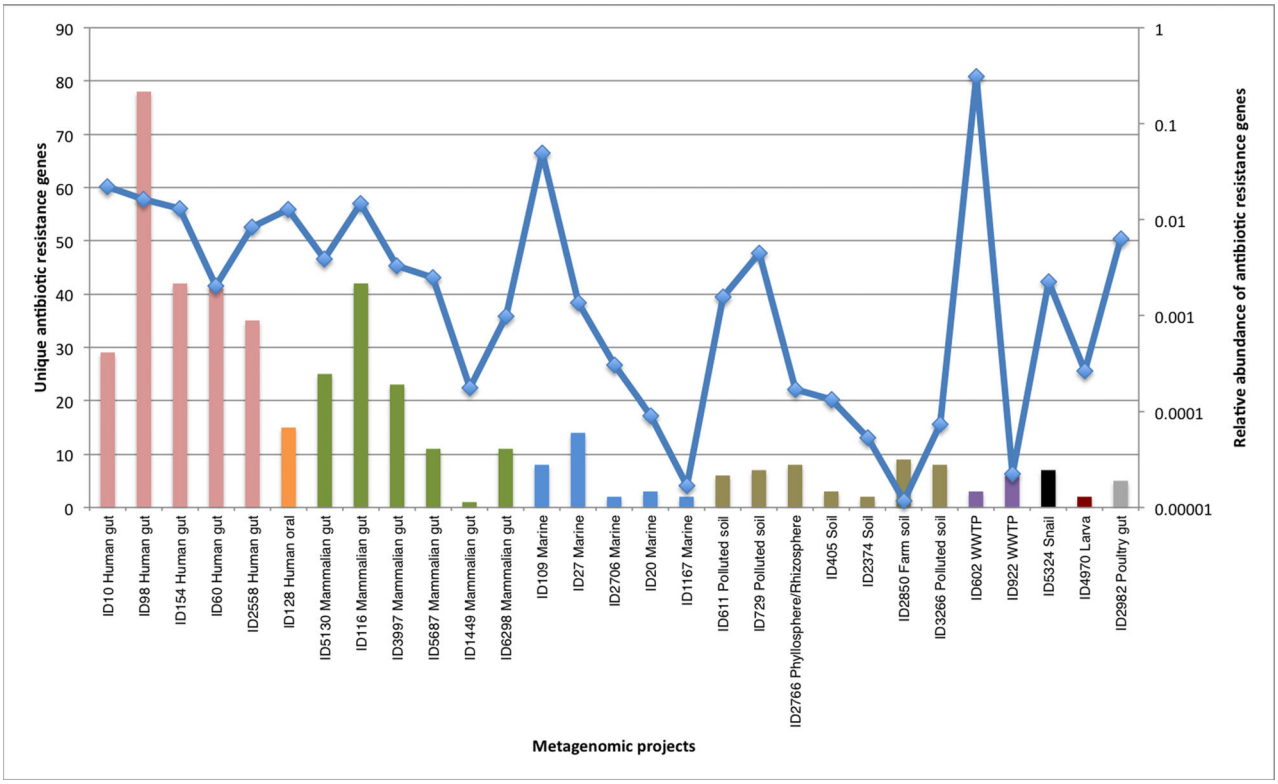

Figure 1. Relative abundance and variety of AR genes in metagenome projects. The bar chart represents the number of different genes identified per project. The line graph represents the relative abundance of resistance genes in each project. Each project had over 1 million gene reads.

(MG RAST ID 128). Of the wastewater treatment plants (WWTP) metagenomes, the hospital WWTP (MG RAST ID 1963) contained the highest diversity (n = 24) even though the coverage was relatively low at 177 214 total genes detected. However, the relative abundance of the resistance genes was very low at 0.025%. This project comprised both influent and WWTP sludge. In comparison, the non-hospital WWTP sludge samples contained between 9, 17 and 6 different AR genes in the three projects MG RAST IDs 43, 731 and 922. The numbers of genes in the datasets were 224 738, 560 998 and 54 080 251 genes, respectively. However, the relative abundances of the resistance genes were 0.01%, 0.047% and 2.2×10^{-5}%, respectively. Thus, while the relative abundance of resistance genes was influenced by the coverage, the relative diversities of resistance genes were not. Metagenomic datasets lacking AR genes were dominated by water, soil, plant or animal associated microbiomes. However, as we do not know the limit of detection of metagenomic analysis these results may be a function of the limits of metagenomic techniques rather than an indication that these environments do not contain AR genes of clinical relevance. Environmental microbiomes contain a great diversity of bacteria, most of which are uncharacterized. The results from metagenomic data may be influenced by bacterial diversity, their genome sizes and the relative abundance of AR genes to their host DNA or the AR gene hosts to the total DNA in the sample. Therefore, while PCR will amplify a gene present even in low relative abundance, metagenomic datasets will only identify genes at a certain abundance within the total DNA present.

Beta-lactamase genes were detected in the widest variety of microbiomes of which bla_{TEM} was most frequently identified (Fig. S1, Supporting Information). The bla_{TEM} genes were present in 48 projects, including at least one metagenome project from all of the different biomes, except the insect metagenomes (MG RAST IDs 5324, 4970, 36), which did not contain the bla_{TEM} genes. The relative abundances of the bla_{TEM} resistance genes were the highest for any gene detected. Their abundance dominated certain environmental metagenomes, but the bla_{TEM} resistance genes were not detected in an abundance of greater than 3% in mammalian or human gut metagenomes (Table 2). The cow virome (MG RAST ID 705) and phage (MG RAST ID 484) projects contained beta-lactam resistance genes. The bla_{TEM} and $bla_{\text{CTX-M}}$ genes were previously identified in the bacteriophage DNA isolated from the human faecal microbiome (Quirós *et al.* 2014). At a stringency of 90%, $bla_{\text{CTX-M}}$ was identified in one human gut metagenome (MG RAST ID 154) and bla_{IMP} in the hospital WWTP (MG RAST ID 1963) (Fig. S1, Supporting Information).

There was a relatively low variety or absence of tetracycline resistance genes in the non-animal-associated metagenomes. The tetracycline resistance genes (*tetW*, *tetQ*, *tetO*) were identified exclusively in animal and human metagenomes. However, tetracycline resistance genes have been identified in high abundance using qPCR in sewage sludge and environmental microbiomes (Duffy, Holliger and Walsh 2014; Xu *et al.* 2015). Trimethoprim is a synthetic antibiotic and not produced in nature. Trimethoprim resistance genes were identified in four human metagenome projects (MG RAST IDs 98, 60, 154, 2558), one mammalian gut metagenome project (MG RAST ID 116) and one farm soil metagenome project (MG RAST ID 2850). The human gut microbiomes contained the greatest diversity of trimethoprim resistance genes (n = 6). However, the abundance of the trimethoprim resistance genes was very low (less than four gene copies) when detected.

Aminoglycoside resistance genes were detected in at least one metagenome project from each of the different biomes.

Downloaded from https://academic.oup.com/femsec/article-abstract/92/2/fiv168/2484737 by Maynooth University user on 15 May 2019

Table 2. Metagenomic projects ranked according to the diversity and abundance of non-efflux resistance genes.

#Genes in project	ID	Projects	#AB resistance genes without efflux	Total non-efflux resistance gene abundance as a function of total genes (%)	Contribution of bla_{TEM} gene abundance to total AR gene abundance (%)
16754433	98	Human gut	78	0.016347912	1.9
3179289	154	Human gut	42	0.013084687	1.2
2033449	116	Mammalian gut	42	0.014802437	3.0
3064560	60	Human gut	41	0.002023129	1.6
2626836	2558	Human gut	35	0.008375095	0.5
6717876	10	Human gut	29	0.021822374	0
6243927	5130	Mammalian gut	25	0.003843735	0
4879354	3997	Mammalian gut	23	0.003299617	0
1913381	128	Human oral	15	0.012700032	1.2
10729528	27	Marine	14	0.001342091	71.3
15125044	5687	Mammalian gut	11	0.002485943	0
2846938	6298	Mammalian gut	11	0.000983513	0
111255128	2850	Farm soil	9	1.16849E-05	0
4894785	109	Marine	8	0.049052206	85.3
5943390	2766	Phyllosphere/rhizosphere	8	0.000168254	10
287594033	3266	Polluted soil	8	7.40627E-05	96.4
1102372	729	Polluted soil	7	0.004444961	0
2886157	5324	Snail	7	0.00225213	0
1613523	611	Polluted soil	6	0.001549405	0
54080251	922	WWTP	6	2.21892E-05	0
1232946	2982	Poultry gut	6	0.006245205	68
12933299	405	Soil	3	0.000131444	90.2
2050142	602	WWTP	3	0.313002709	96.3
1312454	2706	Marine	2	0.000304773	0
11970451	1167	Marine	2	1.67078E-05	0
5652927	2374	Soil	2	5.30699E-05	0
1137603	4970	Larva	2	0.000263712	0
1099557	20	Marine	1	9.09457E-05	100
1131372	1449	Mammalian gut	1	0.000176777	0

With the exception of *aph3iiia*, all of the frequently identified aminoglycoside resistance genes confer resistance to streptomycin. Similar to the trends concerning the diversities of beta-lactam, tetracycline and aminoglycoside resistance genes, the greatest diversity of chloramphenicol resistance genes was also identified within the human gut microbiomes (MG RAST IDs 98, 60, 154, 2558, 10). Overall, the *catA* gene was the most frequently identified chloramphenicol resistance gene. Eleven different, non-efflux-mediated, macrolide resistance genes were identified, which were present in human gut (n = 11), mammalian gut (n = 6), WWTP (n = 3) or soil microbiomes (n = 3) (Fig. S1, Supporting Information). The *sul1* and *sul2* genes were identified in human (n = 6) and mammal (n = 4) metagenomes, WWTP, soil and marine and freshwater metagenomes, but the *sul3* was restricted to one mammalian metagenome project, which also contained the *sul1* and *sul2* genes (MG RAST ID 116). The vancomycin resistance genes were identified in few metagenome projects, which were confined to human gut (MG RAST IDs 98, 60, 154, 10) and mammalian (MG RAST IDs 116, 5130, 5687, 6244, 1452) microbiomes. The two plasmid-mediated quinolone resistance genes, *qnrB* and *qnrS*, were detected in four and one metagenome datasets, respectively (Fig. S1, Supporting Information). The *qnrB* gene was identified in human gut microbiomes (MG RAST ID 98, 2558), a marine microbiome (MG RAST ID 27) and the giant African snail microbiome (MG RAST ID 5324). The *qnrS* gene was identified in a WWTP associated with a hospital (MG RAST ID 1963). The *qnrB* and *qnrS* resistance genes have been previously identified in human samples (Folster *et al.* 2011), WWTPs (Su *et al.* 2014a) and aquatic environments (Zhang *et al.* 2012). As snail microbiomes have not been analysed for the presence of AR genes, these genes have not previously been identified in a snail microbiome.

DISCUSSION

The results discussed in this study were all derived from metagenomic studies. While metagenomes provide a wealth of data in comparison to culture-based or PCR-based studies from total community DNA, they have limitations. In complex microbiomes such as soil, the metagenomics technique is prone to the limitations and bias of all molecular techniques e.g. nucleic acid extraction, or specific limitations and bias such as the depth of coverage required to represent the minor or rare members of the community and genes. Therefore, complete coverage is normally not obtained. The deep coverage of the majority of a soil community is not achieved, even with 300 Gbp of sequence data (Howe *et al.* 2014). Thus, rare members of the soil or other complex microbiomes, which contain AR genes, may have been missed. With further advancements in sequencing technologies, we will hopefully soon reach the depth required to analyse both the rare and major bacteria species in complex microbiomes.

Studies have identified the presence of AR genes in a wide variety of ecological niches (Feller, Sonnet and Gerday 1995; de Kraker and van de Sande-Bruinsma 2007; Allen *et al.* 2009; Cantas *et al.* 2013; Walsh and Duffy 2013; Duffy, Holliger and Walsh 2014; Garcia-Migura *et al.* 2014). However, the presence of a

Downloaded from https://academic.oup.com/femsec/article-abstract/92/2/fiv168/2484737 by Maynooth University user on 15 May 2019

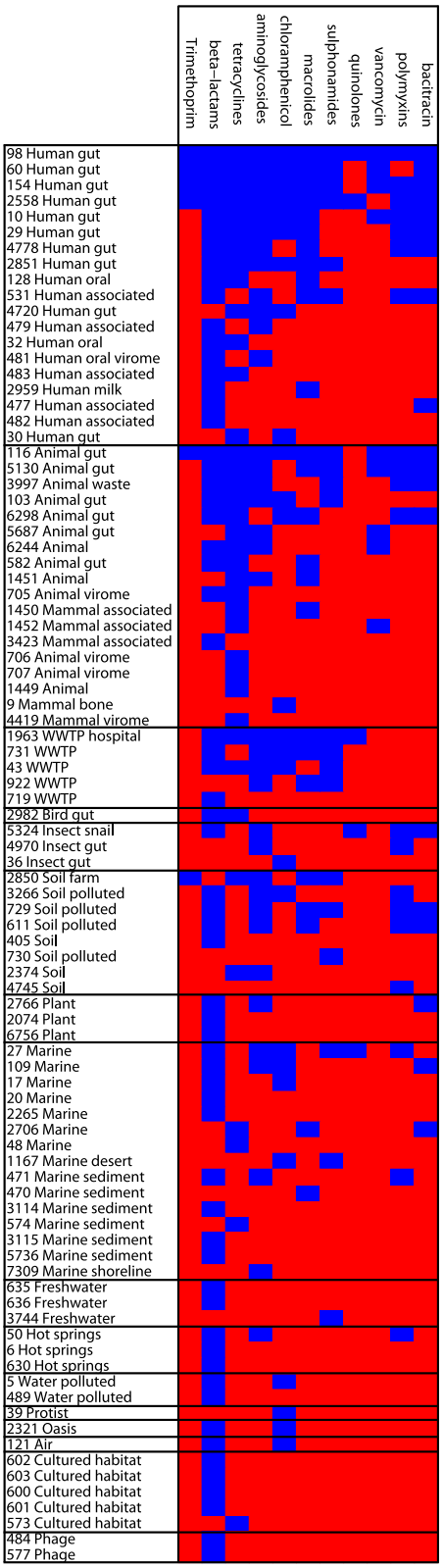

Figure 2. Total AR gene distributions throughout the entire metagenome project database. Blue = presence of non-efflux AR gene conferring resistance to this class. Red = absence of non-efflux AR gene conferring resistance to this class.

resistance gene does not indicate if this is the source or sink of the resistance gene. Thus, while metagenomic data may be utilized to identify the presence of resistance, they have not to date been used to answer questions relating to ecological connectivity of resistance genes. Studies have identified that AR genes may have an environmental reservoir. Soil is thought to be a substantial source of AR genes, and it is thought that antibiotic-producing microorganisms in soil are the source of resistance genes found in clinical pathogens (Canton 2009). However, direct evidence of gene transfer between the environmental resistome and the clinic is rare as resistance genes may undergo many rounds of evolution between the soil and the clinic (Aminov and Mackie 2007). Fresh vegetables may provide the potential link between the environment and clinic through the ingestion of antibiotic-resistant bacteria. However, research to date has focused on antibiotic-resistant pathogenic bacteria, such as *Salmonella* sp. and *Enterobacteriaceae*, which do not provide information about the extent to which commensal and non-pathogenic bacteria act as AR reservoirs (van Hoek et al 2015).

Our results suggest that soil, plant, insect and marine microbiomes use the non-specific efflux-mediated mechanisms to protect themselves from antibiotics. This finding has been identified in a study of metagenomic data from soil and culturable antibiotic-resistant bacteria from soil (Walsh and Duffy 2013; Nesme *et al.* 2014). Among the projects with the highest diversity of efflux resistance mechanisms were two contaminated soils (MG RAST IDs 611, 729), which concur with the hypothesis that these resistance mechanisms were not necessarily selected due to antibiotic selection, but rather toxins, which the bacteria pump out of the cell. Efflux resistance mechanisms are not antibiotic specific, but generically pump out toxins. Our study focused predominantly on the non-efflux resistance genes to minimize the bias due to such non-antibiotic influences.

We have demonstrated that soil and the environment contain potential AR genes of clinical importance and may be sources of the basic forms of AR genes e.g. bla_{TEM}. However, the human and mammalian gut microbiomes contained the greatest diversity and relative abundance of AR genes. This has been borne out in the high diversity of resistance genes to almost all classes of antibiotics identified in these microbiomes in comparison to human oral or non-mammalian microbiomes. Trimethoprim is a synthetic antibiotic and therefore not produced in nature. The trimethoprim resistance genes were identified in human and mammalian gut microbiomes and in one farm soil microbiome. The presence in the absence of direct selective pressure may be due to their frequent localization on integrons and coselection due to other genes present on the same mobile element.

The data presented in this study must be viewed in context. The vast majority of research into AR has focused on resistance present in human pathogenic bacteria. Therefore, the databases of genes are biased towards those associated with human pathogens. However, our aims were restricted to clinically important resistance genes. Thus, the bias afforded by the high proportion of clinically relevant resistance genes in the databases aids to answer our questions regarding the biomes containing the greatest diversity and levels of clinically important resistance genes.

Martínez, Coque and Baquero (2015) suggested that a series of bottlenecks occur in the transfer of AR genes between soil, water, animals, sewerage and humans.

The first bottleneck is ecological connectivity: a gene transfer event only occurs when donor and recipient populations come into contact. The second bottleneck is based on the

Downloaded from https://academic.oup.com/femsec/article-abstract/92/2/fiv168/2484737 by Maynooth University user on 15 May 2019

founder effect: when one resistance determinant is broadly distributed among bacterial populations, it is rare that another resistance determinant that has the same or a closely related substrate profile is acquired; bacteria are already resistant and, consequently, antibiotics do not exert any selective pressure on an organism that already harbours a gene with the same function.

The third bottleneck that determines the spread of a given resistance gene is fitness cost.

We applied these bottleneck theories to our data using one highly abundant resistance gene, which is widespread throughout different metagenomes, bla_{TEM} and another, which is not e.g. $bla_{CTX\text{-}M}$. These theories as applied to bla_{TEM} suggest that all metagenomes containing this gene are connected and are capable of transferring this gene between them, as it is present in almost all metagenomes. As this gene was also in relatively high abundance across many metagenome projects, the second bottleneck suggests that another resistance gene with similar substrate would be rare. This is the case in metagenome projects containing a low diversity of resistance genes. However, several different beta-lactamases were identified in many human and mammalian gut and WWTP metagenomes. Thus, the bla_{TEM} gene can exist in the presence of similar resistance genes, which suggests that either this theory is incorrect. The third bottleneck, which the bla_{TEM} gene has overcome, is the fitness costs associated with this gene. In order to have successfully transferred to such a large number of microbiomes, it must not bear a fitness burden.

We have defined the resistance genes in this study as those with high similarity to resistance genes present in clinical pathogens as we wished to identify if specific microbiomes posed a greater risk to human health than others. Recent debates have developed on how to prioritize risks associated with the presence of resistance genes in metagenomic datasets (Bengtsson-Palme and Larsson 2015; Martínez, Coque and Baquero 2015). The question emerging from this debate is whether to assign the greatest risk to AR genes already present in human pathogens or to those which are as yet unknown, but could potentially transfer to human pathogens. Bengtsson-Palme and Larsson argue that the transfer of known resistance genes is probably rare and therefore, as they are already present in human pathogens are unlikely to dramatically alter patient outcomes. Whereas Martínez, Coque and Baquero (2015) suggests that the highest risk to patients is from mobile resistance genes for which a role in resistance in human pathogens has already been demonstrated.

Using the data generated in this study, we have identified a wide range of AR genes, which are associated with mobile elements and have been shown to generate resistance phenotypes in human pathogens. Our results suggest that not all resistance genes can be grouped together to determine their risk, but should be analysed individually and in terms of their microbiome or bacterial context. The bottlenecks described earlier prevent certain resistance genes ever coming in contact with pathogens. If the resistance genes are capable of overcoming these bottlenecks then they will be ubiquitous, regardless of their function as shown by bla_{TEM}. However, if the bottlenecks prevent the gene from transferring to or being maintained within a pathogen, then its presence in high or low abundance outside the pathogen poses no risk to the human health. The risk arises when the resistance gene e.g. $bla_{CTX\text{-}M}$ is present and when transferred to a human pathogen can then be maintained within the pathogen and spread between different pathogens. In order for these rare genes to pose a risk to human health, they must be provided with the opportunity for proliferation, most frequently by the use of antibiotics. The AR mechanisms, which pose the greatest risk to human health, are those that are capable of transferring to other bacteria, especially human pathogens (Martínez, Coque and Baquero 2015).

When using metagenomics data to identify risk the context of the results and the limitations of the technique must be emphasized. Our findings indicated that the human and mammalian metagenomic datasets contained the greatest diversity and relative abundance of clinically relevant AR genes. However, we do not know the minimum coverage required in order to detect the rare AR genes or AR genes, which are present in a very diverse metagenome. Therefore, in order to identify AR gene risk using metagenomic data we must first identify the limits of detection of this technique, and second identify how we can normalize datasets containing vastly different diversities of microbial species and genome sizes. If we do not detect the AR gene is it in fact not there or just below the limits of our techniques? Once we have identified these caveats, we then need to understand if the context of the AR gene, is it on a promiscuous plasmid, capable of transferring to many species or is it present on the chromosome? Finally, we need to understand whether one copy of a gene in a total metagenomic dataset of over one million genes poses a risk to the treatment of a patient.

SUPPLEMENTARY DATA

Supplementary data are available at FEMSEC online.

ACKNOWLEDGEMENTS

We wish to thank the funders for enabling this research.

AUTHOR CONTRIBUTIONS

DF generated the datasets and the analysis pipelines and created the heatmap figures. FW conceived the idea for the work, analysed the data generated and compared the data to current knowledge. Both authors contributed to the manuscript writing.

FUNDING

This work was funded by Maynooth University starting funds. Dr Walsh's research group has received funding from the EU Water JPI Stare project, the COST Action ES1403 and the EU ERA-Net ANIHWA.

Conflict of interest. None declared.

REFERENCES

Allen HK, Cloud-Hansen KA, Wolinski JM *et al.* Resident microbiota of the gypsy moth midgut harbors antibiotic resistance determinants, *DNA Cell Biol* 2009;**28**:109–17.

Altschul SF, Madden TL, Schäffer AA *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997; **25**:3389–402.

Alves MS, Pereira A, Araújo SM *et al.* Seawater is a reservoir of multi-resistant *Escherichia coli*, including strains hosting plasmid-mediated quinolones resistance and extended-spectrum beta-lactamases genes. *Front Microbiol* 2014; **20**:426.

Downloaded from https://academic.oup.com/femsec/article-abstract/92/2/fiv168/2484737 by Maynooth University user on 15 May 2019

Aminov RI, Mackie RI. Evolution and ecology of antibiotic resistance genes. *FEMS Microbiol Lett* 2007;**271**:147–61.

Andersson DI, Hughes D. Microbiological effects of sublethal levels of antibiotics. *Nat Rev Microbiol* 2014;**12**:465–78.

Baquero F, Negri MC, Morosini MI *et al*. Antibiotic-selective environments. *Clin Infect Dis* 1998;**27**(Suppl 1):S5–11.

Bengtsson-Palme J, Larsson DG. Antibiotic resistance genes in the environment: prioritizing risks. *Nat Rev Microbiol* 2015;**13**:396.

Berendonk TU, Manaia CM, Merlin C *et al*. Tackling antibiotic resistance: the environmental framework. *Nat Rev Microbiol*. 2015;**13**:310–7.

Cantas L, Shah SQ, Cavaco LM *et al*. A brief multi-disciplinary review on antimicrobial resistance in medicine and its linkage to the global environmental microbiota. *Front Microbiol* 2013;**14**:96.

Cantón R. Antibiotic resistance genes from the environment: a perspective through newly identified antibiotic resistance mechanisms in the clinical setting. *Clin Microbiol Infect* 2009;**15**(Suppl 1):20–5.

de Kraker M, van de Sande-Bruinsma N. Trends in antimicrobial resistance in Europe: update of EARSS results. *Euro Surveill* 2007;**12**:E070315.3.

Duffy B, Holliger E, Walsh F. Streptomycin use in apple orchards did not increase abundance of mobile resistance genes. *FEMS Microbiol Lett* 2014;**350**:180–9.

Feller G, Sonnet P, Gerday C. The beta-lactamase secreted by the antarctic psychrophile *Psychrobacter immobilis A8*. *Appl Environ Microb* 1995;**61**:4474–6.

Finley RL, Collignon P, Larsson DG *et al*. The scourge of antibiotic resistance: the important role of the environment. *Clin Infect Dis* 2013;**57**:704–10.

Folster JP, Pecic G, Bowen A *et al*. Decreased susceptibility to ciprofloxacin among *Shigella* isolates in the United States, 2006 to 2009. *Antimicrob Agents Ch* 2011;**55**:1758–60.

Forsberg KJ, Patel S, Gibson MK *et al*. Bacterial phylogeny structures soil resistomes across habitats. *Nature* 2014;**509**:612–6.

Garcia-Migura L, Hendriksen RS, Fraile L *et al*. Antimicrobial resistance of zoonotic and commensal bacteria in Europe: the missing link between consumption and resistance in veterinary medicine. *Vet Microbiol* 2014;**170**:1–9.

Heuer H, Schmitt H, Smalla K. Antibiotic resistance gene spread due to manure application on agricultural fields. *Curr Opin Microbiol* 2011;**14**:236–43.

Howe AC, Jansson JK, Malfatti SA *et al*. Tackling soil diversity with the assembly of large, complex metagenomes. *P Natl Acad Sci USA* 2014;**111**:4904–9.

Jechalke S, Kopmann C, Rosendahl I *et al*. Increased abundance and transferability of resistance genes after field application of manure from sulfadiazine-treated pigs. *Appl Environ Microb* 2013;**79**:1704–11.

Jia S, He X, Bu Y *et al*. Environmental fate of tetracycline resistance genes originating from swine feedlots in river water. *J Environ Sci Health B* 2014;**49**:624–31.

Liu B, Pop M. ARDB–Antibiotic Resistance Genes Database. *Nucleic Acids Res* 2009;**37**:D443–7.

Martínez JL. Antibiotics and antibiotic resistance genes in natural environments. *Science* 2008;**321**:365–7.

Martínez JL, Coque TM, Baquero F. What is a resistance gene? Ranking risk in resistomes. *Nat Rev Microbiol* 2015;**13**:116–23.

Meyer F, Paarmann D, D'Souza M *et al*. The metagenomics RAST server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* 2008;**9**:386.

Nesme J, Cécillon S, Delmont TO *et al*. Large-scale metagenomic-based study of antibiotic resistance in the environment. *Curr Biol* 2014;**24**:1096–100.

Quirós P, Colomer-Lluch M, Martínez-Castillo A *et al*. Antibiotic resistance genes in the bacteriophage DNA fraction of human fecal samples. *Antimicrob Agents Ch* 2014;**58**:606–9.

Su HC, Ying GG, He LY *et al*. Antibiotic resistance, plasmid-mediated quinolone resistance (PMQR) genes and *ampC* gene in two typical municipal wastewater treatment plants. *Environ Sci Process Impacts* 2014a;**16**:324–32.

Su JQ, Wei B, Xu CY *et al*. Functional metagenomic characterization of antibiotic resistance genes in agricultural soils from China. *Environ Int* 2014b;**65**:9–15.

van Hoek AH, Veenman C, van Overbeek WM *et al*. Prevalence and characterization of ESBL- and AmpC-producing Enterobacteriaceae on retail vegetables. *Int J Food Microbiol* 2015;**204**:1–8.

Walsh F, Duffy B. The culturable soil antibiotic resistome: a community of multi-drug resistant bacteria. *PLoS One* 2013;**8**:e65567.

Wang H, McEntire JC, Zhang L *et al*. The transfer of antibiotic resistance from food to humans: facts, implications and future directions. *Rev Sci Tech* 2012;**31**:249–60.

Xu J, Xu Y, Wang H *et al*. Occurrence of antibiotics and antibiotic resistance genes in a sewage treatment plant and its effluent-receiving river. *Chemosphere* 2015;**119**:1379–85.

Zhang R, Ichijo T, Huang YL *et al*. High prevalence of *qnr* and *aac(6′)-Ib-cr* genes in both water-borne environmental bacteria and clinical isolates of *Citrobacter freundii* in China. *Microbes Environ* 2012;**27**:158–63.

Downloaded from https://academic.oup.com/femsec/article-abstract/92/2/fiv168/2484737 by Maynooth University user on 15 May 2019