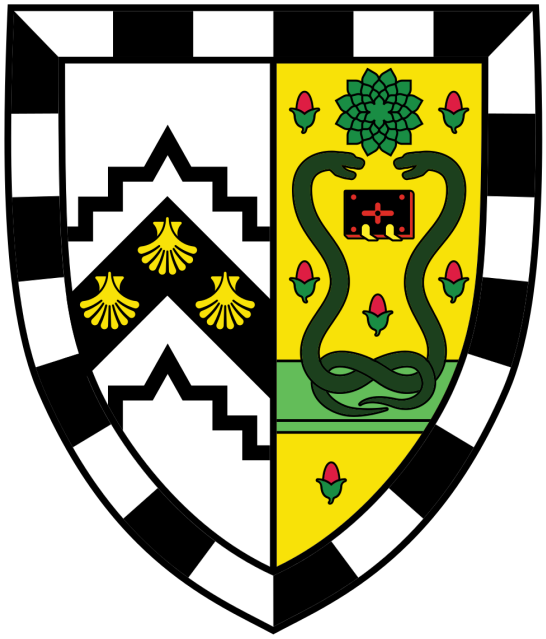
Characterising antibiotic susceptibility and resistance in human commensal gut bacteria



Lindsay Jacqueline Pike

Gonville and Caius College, University of Cambridge

Wellcome Sanger Institute

August 2019

This dissertation is submitted for the degree of Doctor of Philosophy

Supervised by Dr Trevor Lawley, Host-Microbiota Interactions Laboratory

Funded by the Medical Research Council and the Wellcome Sanger Institute

Declaration

This dissertation is the result of my own work and includes nothing that is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. It does not exceed the word limit of 60,000 words (excluding bibliography, figures, and appendixes) as prescribed by the Degree Committee for the Faculty of Biology at the University of Cambridge.

Mr Mark Stares assisted with phenotyping gut bacteria against antibiotics and extracting DNA for whole genome sequencing. Mr Matthew Dorman assisted with cloning of candidate novel antibiotic resistance genes. Dr Simon Clare and members of his team looked after the mice used in this study and collected mouse faecal pellets. Dr Sam Forster, Dr B. Anne Neville, Dr Ana Zhu, Dr Elisa Viciani, Dr Hilary Browne and Mr Mark Stares assisted with mouse sample processing and isolating individual bacterial colonies. Dr Kevin Vervier helped analyse diversity indices in metagenomic samples.

Lindsay Jacqueline Pike

August 2019

‘Life, uh… finds a way.’

Dr Ian Malcom

Jurassic Park (1993)

Dedication

For my Grampy, who always said I should be a weather girl on TV. Hopefully this isn't too much of a disappointment!

To my parents: it is only with your support and values you have instilled in me that I have achieved all that I have.

Unconventionally, as ever, I also dedicate this thesis to myself: it stands as a monument to my determination, tenacity and strength.

Abstract

The human commensal gut microbiota can act as a reservoir of antimicrobial resistance genes that can persist and spread to pathogens. However, the extent and diversity of antibiotic resistance encoded by human commensal bacteria remains to be determined. Due to immediate clinical relevance and our previous inability to culture these commensal bacteria, the majority of research into antibiotic resistance has focused on pathogenic organisms or well-characterized antibiotic resistance mechanisms. Here, I demonstrate the existence of unpredicted antibiotic resistance, not detected by several genome-based prediction methods, in diverse bacterial species from the human gastrointestinal tract.

178 antibiotic resistance genes and mutations were identified in a culture collection of 737 phylogenetically diverse gut bacteria from healthy humans. Recent developments in culturing anaerobic gut bacteria were used to determine antibiotic sensitivity phenotypes and observe the spectrum of clinically relevant antibiotics across the diversity of these isolates. These data were combined to assess the accuracy of genome-based predictions in human commensal gut bacteria, revealing multiple instances of unpredicted antibiotic resistance. This highlights the importance of combining computational genomic prediction with functional validation and increases our knowledge of antibiotic resistance in commensal human gut bacteria.

In addition, the impact of therapeutic amoxicillin treatment on antibiotic resistance in mice with human-derived gut microbiota was studied. These experiments model processes in humans and reveal community- and strain-level changes in antibiotic resistance following antibiotic therapy. These experiments further elucidate the role of the gut microbiota as a reservoir of antibiotic resistance and the influence of antibiotics on this reservoir.

Acknowledgements

Firstly, I must thank my supervisor, Dr Trevor Lawley, for his guidance throughout the four years of my PhD. He has very astute insight into project directions and the future of our field, which has been enormously helpful. In addition, he was very understanding when I experienced some personal issues over the last few years and supportive of my future aspirations, for which I am very grateful.

In addition, I would like to thank Professor Stephen Baker and Dr Estee Torok as members of my thesis committee. In addition, to Professor Nick Thompson, a member of the Committee of Graduate Studies at Sanger during my PhD. Their perspectives, advice and support – both academic and pastoral – were always gratefully received and helped to shape this thesis, as well as keeping me motivated at the times I found myself struggling.

I would also like to thank Professor Gordon Dougan and Professor Julian Parkhill – I have been extremely fortunate in receiving advice from these two eminent microbiologists due to the collaborative and supportive nature of the Parasites and Microbes programme at Sanger.

Dr Hilary Browne and Dr Sam Forster have demonstrated immense patience and diligence whilst reading and providing feedback on the various drafts of the chapters contained within this thesis. It is thanks to them and Trevor that my scientific writing has developed from when I arrived straight out of my undergraduate degree. It is also mainly thanks to Sam and Dr Nitin Kumar that my bioinformatics skills have advanced. I also owe Hilary, Nitin and Sam thanks for answering my endless questions – they are fountains of knowledge and expert scientists and I have learnt much from them. This is extended to all other current members of the Host-Microbiota Interactions Laboratory (Team 162): Mr Nick Dawson (the brave culture collection curator), Dr Junyan Liu (the plasmid expert), Yan Shao (the metagenomics magician), Mr Mark

Stares (the lab wizard), Dr Kevin Vervier (the statistics king), Dr Ana Zhu (the bioinformatics and now mouse experiment queen). In addition, to previous Team 162 members: Mr Matt Dunn (the fermenter guru), Dr B. Anne Neville (the most enthusiastic microbiologist I have ever met), Dr Giulia Falivelli (a role model for how to balance science and family), Dr Fernanda Schrieber (the cell culture pro), and Dr Elisa Viciani (an extremely diligent scientist). Thank you for your wisdom, patience, camaraderie and friendship.

Special thanks in particular must go to Sam, who assisted with the long days of mouse experiments, and anyone else who helped pick colonies in those times (at least Anne, Ana, Elisa, Mark). Also to Mark, for assisting with phenotyping gut bacteria against antibiotics and extracting DNA for whole genome sequencing – I would not have been able to perform as many tests or sequence as many bacteria without your help. Plus, to Mr Matthew Dorman, who lent time, knowledge and hands to plasmid cloning experiments. In addition, to Dr Simon Clare, Ms Cordelia Brandt, and other members of Simon's team for collecting mouse stool samples for this study. Finally, to Sam, Nitin and Kevin who helped me write scripts to analyse my data.

Thank you also to the Graduate Office at Sanger for their support: Dr Annabel Smith, Dr Christina Hedberg-Delouka and Dr Carl Anderson are always there to answer questions via email or when cornered unexpectedly in their office or the DiNA. Also, to Carl and other members of CoGS for giving me the opportunity to do the PhD at Sanger in the first place. It hasn't been easy, but I have always tried to remind myself how lucky I have been to study here and I am so glad to have done it.

On a more personal level, I would like to thank my fellow Sanger PhD buddies, Fiona Calvert and Tapoka Mkandawire. This triangle of mutual support has kept me going more times than

you will know and I can't wait to see where we go after our PhDs. I would like to extend this to my friends I have met through Gonville and Caius College, where I had the fortune of being part of a wonderful graduate community. The College porters as well deserve their own mention for making Caius such a special, friendly place to be. It truly has been my home for the last four years.

My friends from "home home" have also been instrumental in helping me stay grounded and in touch with reality (a PhD should not be your whole life!) – Katie, Natalie, Bonita, Christina, Robin, Adam, thank you for still being here despite my frequent moaning about PhD life when truthfully it's been pretty great overall. Also to my friends I met as an undergraduate – Suzy, Becky, Megan, Sophie, we've all been continuing further education at the same time and it's been enormously helpful to have that support network as well. I only hope I've been as supportive to you as you have been to me.

To a recent addition in my support network, Will ("Dr Dr" Hamilton) – thank you for reminding me how to enjoy science. With your positivity and encouragement I have enjoyed my final year and writing this thesis much more than I expected!

Finally, to all of my family – especially my mum Cathy, dad Julian and grandma Frances (my biggest fan) – thank you for your unconditional support not just over the last four years but the last 27. My success is all thanks to you.

Publications

Arising elsewhere:

Pike, L. J., Viciani, E. & Kumar, N. Genome watch: Microbial diversity knows no borders. *Nat Rev Microbiol* **16** (2), 66 (2018).

Pike, L. J. & Forster, S. C. Genome watch: A new piece in the microbiome puzzle. *Nat Rev Microbiol* **16** (4), 186 (2018).

Forster, S. C., Kumar, N., Anonye, B. A., Almeida, A., Viciani, E., Stares, M. D., Dunn, M., Mkandawire, T. T., Zhu, A., Shao, Y, **Pike, L. J.,** Louie, T., Browne, H. P., Mitchell, A. L., Neville, B. A., Finn, R. D. & Lawley, T. D. A human gut bacterial genome and culture collection for improved metagenomic analyses. *Nat Biotechnol* **37**, 186-192 (2019).

Contents

List of Figures

List of Tables

Glossary

AMR: antimicrobial resistance

AMX: amoxicillin

ANI: average nucleotide identity

ARIBA: Antibiotic Resistance Identification By Assembly

ARG: antibiotic resistance gene

AST: antibiotic susceptibility testing

BNF: British National Formulary

Bp: base pairs

BSAC: British Society of Antimicrobial Chemotherapy

CARD: Comprehensive Antibiotic Resistance Database

CARD-RGI: Comprehensive Antibiotic Resistance Database – Resistance Gene Identifier

CFU: colony forming units

CIA: critically important antimicrobials

CLSI: Clinical Laboratory Standards Institute

Confirmed Resistance: genetic resistance and phenotypic resistance to a particular antibiotic both observed in an isolate

Confirmed Susceptibility: no genetic or phenotypic resistance to a particular antibiotic observed in an isolate

CRE: carbapenem-resistant Enterobacteriaceae

D7AMX1: the first Donor 7 mouse experiment performed in this study

D-Ala-D-Ala: D-Alanyl-D-Alanine

DDD: defined daily doses

DRI: drug-resistant infection

ESBL: extended-spectrum beta-lactamase

ESKAPE: Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa and Enterobacter spp.

EUCAST: European Committee on Antibiotic Susceptibility Testing

FMT: faecal microbiota transplant

GF: germ-free

GI: gastrointestinal

HAI: hospital acquired infection

HBC: Human Gastrointestinal Bacteria Culture Collection

HGT: horizontal gene transfer

iTOL: Interactive Tree of Life

KAPE: Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa and Enterobacter spp.

LJP01: the second Donor 7 mouse experiment performed in this study

LJP02: the Donor 2 mouse experiment performed in this study

MAG: metagenome assembled genome

MDR: multi-drug resistant

MGE: mobile genetic element

MIC: minimum inhibitory concentration

MLPS: Macrolide, Lincosamide, Pleuromutilin and Streptogramin antibiotics

MRSA: methicillin-resistant *Staphylococcus aureus*

NGS: next generation sequencing

ORF: open reading frame

OTU: operational taxonomic unit

PATRIC: Pathosystems Resource Integration Center

PBPs: penicillin-binding proteins

R&D: research and design

SEM: standard error of the mean

SNP: single nucleotide polymorphism

STC: sodium taurocholate

TB: tuberculosis

Unpredicted Susceptibility: also referred to as “False Positive”; genetic resistance predicted to a particular antibiotic in an isolate, but the isolate was phenotypically susceptible

Unpredicted Resistance: also referred to as “False Negative”; no genetic resistance to a particular antibiotic predicted in an isolate, but the isolate was phenotypically resistant

UTI: urinary tract infection

VRE: vancomycin resistant enterococci

WGS: whole genome sequencing

WGS-AST: whole genome sequencing based antibiotic susceptibility testing

WHO: World Health Organisation

WSI: Wellcome Sanger Institute

Chapter 1: Antibiotics and commensal human gut bacteria

1.1 Antibiotics

Antibiotics are a cornerstone of modern medicine and are used extensively throughout the developed and developing worlds, in humans and for veterinary and agricultural purposes. They are drugs that either kill or inhibit the growth and replication of a bacterium[1] and are used to treat bacterial infections in, or on, the body. They are distinct from antiseptics (used to sterilise living tissue and reduce the risk of infection, rather than treat an infection) and disinfectants (non-selective antimicrobials that kill a range of microorganisms, not just bacteria, and are used on non-living surfaces)[1]. They are also not toxins, as toxins are defined as a poisonous substance produced by a microorganism, plant or animal that causes illness in the body; antibiotics do not cause direct harm to human cells. Antibiotics can be administered intravenously (into veins via syringe or catheter), intramuscular (into muscle via syringe), orally (in tablet, capsule or liquid form), or topically (e.g. creams, lotion, sprays or drops). Intravenous is considered the most effective route as it creates an immediate therapeutic blood level of the antibiotic, but oral routes are often preferred as they are less intrusive, do not require a hospital stay, and often achieve supra-inhibitory blood levels[2].

Antibiotics can now be synthesised in a laboratory, but are often based on compounds produced naturally by microorganisms to harm or kill bacteria in their environment. Not all microorganisms produce antibiotic compounds – in fact, there are only around 20 species that produce antibiotics that are now mass-produced and used in medicine[3]. Typically these are soil-dwelling microorganisms[3], although the search for new antibiotics is now beginning to shift to marine microbes[4]. Both of these environments feature high diversity and density of microorganism; antibiotic production is thought to be a mechanism of attack against

neighbouring bacteria or defence, enabling survival of the antibiotic-producer[5]. Antibiotic compounds can also act as signalling molecules between bacterial cells, regulating bacterial behaviour such as biofilm formation[6]. Therefore, they have important roles in nature that humans have co-opted for our benefit and integrated antibiotics into human and animal medicine. It is difficult to quantify exactly the overall impact of antibiotics on infection dynamics and distinguish between the impacts of other factors such as the introduction of vaccines[7] or improved sanitation[8]. Despite this, antibiotic use is associated with a decrease in deaths caused by communicable diseases per year[9] and increase in average life expectancy[9].

1.2 The history of antibiotics

It is only recently, and over a relatively short timeline, that antibiotics have become entrenched in our society to treat bacterial infections (Fig. 1.1). Just 110 years ago, Paul Ehrlich (the German biochemist, 1854-1915) coined the term "magic bullet" when theorising a chemical that could selectively target and kill disease-causing agents in the body without harming the patient[10]. The first "true" magic bullet antibiotic against bacteria that was used clinically against bacterial infections was in fact used a number of years before this term was used. Pyocyanase was an extract prepared from the Gram-negative bacterium *Pseudomonas aeruginosa* by Emmerich and Low in 1899[11]. It was active against a number of pathogenic bacteria but since its effectiveness was inconsistent and it was mildly toxic to humans it was abandoned[11]. Ehrlich, now considered the founding father of chemotherapy and large-scale, systematic drug screens, developed his own magic bullet ('Compound 606') against the causative bacterium of syphilis, *Treponema palladium*[12]. Compound 606, or Salvarsan, was the most frequently prescribed drug after its discovery in 1909[13]. Prontosil was another early antibiotic, a compound synthesised as part of a screen by Bayer chemists Josef Klarer and Fritz

Mietzsch and shown to have antibacterial properties against a number of diseases by Gerhard Domagk[14]. The active component sulphanilamide had been previously used in the dye industry and had already come off-patent, enabling sulphonamide derivatives to be produced by various companies[15]. Sulphonamides are the oldest class of synthetic antibiotics and new versions are still produced and used today.

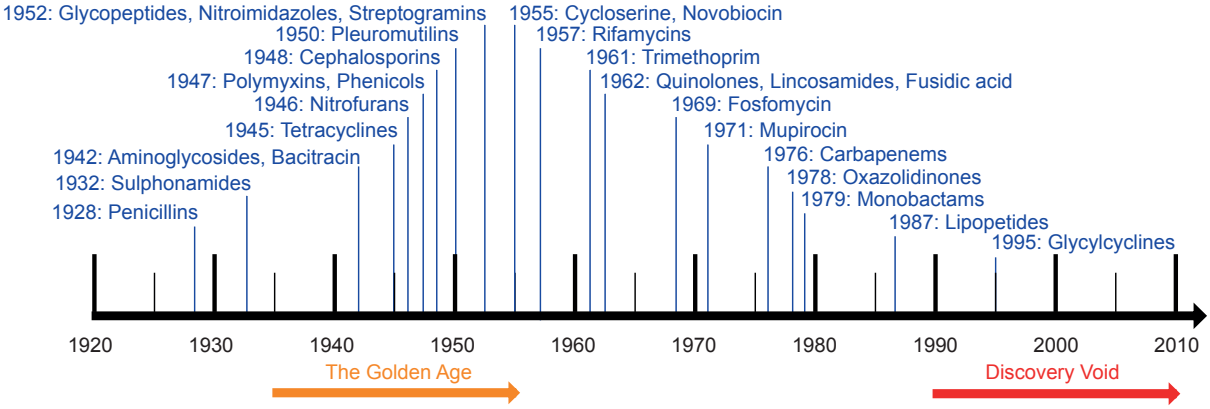


Figure 1.1. Timeline of antibiotic discovery. Adapted from Silver 2011[16].

Probably the most well-known antibiotic, and one of the oldest mass-produced antibiotics, is penicillin, "discovered" in 1928 in the famous tale of Alexander Fleming (Scottish bacteriologist, 1881-1955) and an open window. A mysterious fungus had blown through an open window and contaminated plates being used to study *Staphylococcus*, which was known to cause infections in humans, and the fungus had managed to halt the growth of this bacterium[17]. The antimicrobial properties of moulds were already known, but Fleming was remarkably dedicated in his efforts to purify the exact compound responsible for this effect. Penicillin was eventually brought into mass production in 1945, overtaking Salvarsan in usage[15]. Penicillin antibiotics are still the most prescribed drug globally to this day[18]. The mass production of penicillin coincided with the discovery of streptomycin in 1943, an antibiotic isolated from the soil bacterium *Streptomyces griseus* that could treat *Mycobacterium tuberculosis*, the cause of tuberculosis (TB)[19]. These events heralded the start of the "golden

age of antibiotics" (Fig. 1.1): eighteen new types of antibiotics were discovered between 1944 and 1970[16]. The rate of discovery then began to decline, with only four new classes in the 1970s and one in the 1980s. Since then, no new classes of antibiotics have been discovered[16] – only "re-discoveries" have occurred, involving modifications of already known antibiotics[16] (Table 1.1). The decline in discovery of novel classes of antibiotics is problematic due to the emergence and spread of antibiotic resistance, where bacterial infections can no longer be treated with the same drugs that were once so effective. Since antibiotics are inseparable from today's world, there are concerns a "post-antibiotic era" may be approaching[20]. In light of this, extraordinary research is taking place across the world to discover new antibiotics, develop alternatives to antibiotics, and understand the evolution and spread of antibiotic resistance.

Table 1.1. Sources of antibiotics. The year of discovery and source of current antibiotic classes are described. Adapted from Silver, 2011[16] and Chandra and Kumar, 2016[3].

Year	Antibiotic	Source	Kingdom; Division (Fungi) or Phylum (Bacteria); Class; Order; Family)
1928	Penicillins	*Penicillium*	Fungi; Ascomycota; Eurotiomycetes; Eurotiales; Trichocomaceae
1932	Sulphonamides	Synthetic	
1942	Aminoglycosides	*Streptomyces* spp.	Bacteria; Actinobacteria; Actinomycetes; Actinomycetales; Actinomycetaceae
1942	Bacitracin	*Bacillus subtilis*	Bacteria
1945	Tetracyclines	*Streptomyces* spp.	Bacteria; Actinobacteria; Actinomycetes; Actinomycetales; Actinomycetaceae
1946	Nitrofurans	Synthetic	
1947	Polymyxins	*Paenibacillus polymyxa*	Bacteria
1947	Phenicols	*Streptomyces* spp.	Bacteria; Actinobacteria; Actinomycetes; Actinomycetales; Actinomycetaceae
1948	Cephalosporins	*Acremonium*	Fungi; Ascomycota; Hypocreales; Hypocreaceae
1950	Pleuromutilins	*Clitopilus passecherianus*	Fungi; Basidiomycota; Agaricomyctes; Agaricales; Entolomataceae
1952	Glycopeptides	*Streptomyces* spp.	Bacteria; Actinobacteria; Actinomycetes; Actinomycetales; Actinomycetaceae
1952	Nitroimidazoles	*Streptomyces* spp.	Bacteria; Actinobacteria; Actinomycetes; Actinomycetales; Actinomycetaceae
1952	Streptogramins	*Streptomyces* spp.	Bacteria; Actinobacteria; Actinomycetes; Actinomycetales; Actinomycetaceae
1955	Cycloserine	*Streptomyces* spp.	Bacteria; Actinobacteria; Actinomycetes; Actinomycetales; Actinomycetaceae
1955	Novobiocin	*Streptomyces* spp.	Bacteria; Actinobacteria; Actinomycetes; Actinomycetales; Actinomycetaceae
1957	Rifamycins	*Amycolatepsis rifamycinica*	Bacteria; Actinobacteria; Actinomycetes; Actinomycetales; Pseudonocardiaceae
1961	Trimethoprim	Synthetic	
1962	Quinolones	Synthetic	
1962	Lincosamides	*Streptomyces* spp.	Bacteria; Actinobacteria; Actinomycetes; Actinomycetales; Actinomycetaceae
1962	Fusidic acid	*Fusidium coccineum*	Fungi; Ascomycota; Hypocreales; Nectriaceae
1969	Fosfomycin	*Streptomyces* spp.	Bacteria; Actinobacteria; Actinomycetes; Actinomycetales; Actinomycetaceae
1971	Mupirocin	*Pseudomonas fluorescens*	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae
1976	Carbapenems	*Streptomyces* spp.	Bacteria; Actinobacteria; Actinomycetes; Actinomycetales; Actinomycetaceae
1978	Oxazolidinones	Synthetic (derived from cycloserine)	
1979	Monobactams	Synthetic (derived from *Chromobacterium violaceum*)	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales; Neisseriaceae
1987	Lipopeptides	*Streptomyces* spp.	Bacteria; Actinobacteria; Actinomycetes; Actinomycetales; Actinomycetaceae
1995	Glycylcyclines	Synthetic	Synthetic (tetracycline derivative)

1.3 Clinically relevant antibiotics

Today, seventeen classes of antibiotics are considered essential by the World Health Organisation[21] (Table 1.2). There are three categories of essential antibiotics: Access (first or second line of defence options for common infections); Watch (higher resistance potential but still recommended as first or second line treatments); and Reserve ('last resort' options for serious antibiotic-resistant infections). Each antibiotic has a different mechanism of action (Fig. 1.2), mechanism of resistance and can target different types of bacterial pathogens (Table 1.2). This may include Gram-negative and/or Gram-positive bacteria, as determined by the Gram staining technique developed by Hans Christian Gram in 1884[22]. In addition, bacteria can be strictly aerobic or anaerobic, meaning that the absence or presence of oxygen respectively is toxic, or somewhere in between[23]. Antibiotics such as aminoglycosides rely on components of the aerobic respiration pathway, and as such their spectrum of activity is thought to only include aerobic bacteria and not anaerobes[24]. Essential antibiotics are discussed in more detail below, including their reported spectrum of activity.

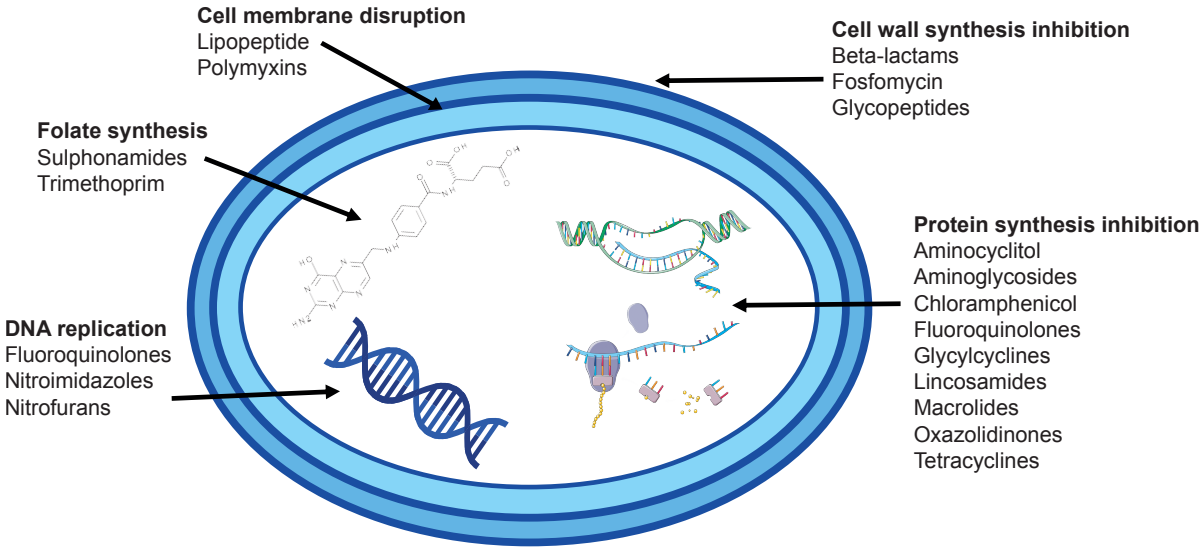


Figure 1.2. Mechanism of action for antibiotics on the WHO List of Essential Medicines 2017[21]. Adapted from Shaikh *et al.* 2015[25]. Vector images of DNA, folate and protein synthesis from Freepik.com and smart.servier.com.

Table 1.2. The antibiotics included in the 20th Edition of the WHO List of Essential Medicines (2017)[21]. These antibiotics are considered essential by the WHO to treat bacterial infections. There are three categories: Access (first or second line options for common infections); Watch (higher resistance potential but still recommended as first or second line treatments); and Reserve ('last resort' options for serious antibiotic-resistant infections). Antibiotics can be considered both Access and Watch; here, they have been listed under the higher concern category of Watch for conciseness.

WHO Category	Antibiotic Class	Sub-class	Drug name	Mechanism of action	Mechanism of resistance	Used against
Access	Aminocyclitol		Spectinomycin	Protein synthesis inhibition: bind 30S ribosomal subunit	Ribosomal mutations; enzymes that modify either the antibiotic or target	Aerobic Gram-negative bacteria e.g. *Pseudomonas aeruginosa* and staphylococci
Access	Aminoglycoside		Amikacin Gentamicin	Protein synthesis inhibition: bind 30S ribosomal subunit	Ribosomal mutations; antibiotic- or target-modifying enzymes	Aerobic Gram-negative bacteria e.g. *Pseudomonas aeruginosa* and staphylococci
Access	Amphenicol		Chloramphenicol	Protein synthesis: interacts with 23S rRNA of 50S ribosomal subunit	Ribosomal mutations; antibiotic-modifying enzymes; efflux	Gram-negative bacteria e.g. *Haemophilus influenzae* and Gram-positives e.g. *Streptococcus pneumoniae*
Access	Beta-lactam	1st generation cephalosporin	Cefalexin Cefazolin	Cell wall synthesis inhibition: interact with cell wall catalysis enzymes	Antibiotic-degrading enzymes, beta-lactamases; mutations in cell wall proteins	Gram-negative and Gram-positives e.g. penicillin-resistant bacteria such as *Pseudomonas aeruginosa*
Access	Beta-lactam	Penicillin	Amoxicillin Amoxicillin + clavulanic acid Ampicillin Benzathine benzylpenicillin Benzylpenicillin Cloxacillin Phenylmethylpenicillin Procain benzylpenicillin	Cell wall synthesis inhibition: interact with cell wall catalysis enzymes	Antibiotic-degrading enzymes, beta-lactamases; mutations in cell wall proteins	Gram-negative and Gram-positives

WHO Category	Antibiotic Class	Sub-class	Drug name	Mechanism of action	Mechanism of resistance	Used against
Access	Folic acid metabolism inhibitors		Sulphamethoxazole + trimethoprim	Folate synthesis inhibition: inhibit enzymes in different stages of the folic acid biosynthesis pathway	Mutations to increase expression of trimethoprim targets, or in the targets themselves, or acquisition of genes encoding less sensitive targets	Gram-negative and Gram-positives including MRSA
Access	Nitroimidazole		Metronidazole	DNA replication inhibition: the prodrug is reduced and converted to the toxic drug that inhibits DNA synthesis	Decreased uptake; altered reduction efficiency; efflux; drug inactivation; acquisition of genes that encode alternative reductases and convert prodrug to non-toxic alternative	Anaerobic Gram-negative or Gram-positive bacteria e.g. *Prevotella*
Access	Nitrofuran		Nitrofurantoin	DNA replication inhibition: the prodrug is reduced and the intermediate metabolites bind enzymes involved in DNA and RNA synthesis	Mutations in reductases	Gram-negative and Gram-positives excluding intrinsically-resistant *Klebsiella* and *Pseudomonas* spp.
Access	Tetracycline		Doxycycline	Protein synthesis inhibitor: prevent translation by binding 16S rRNA in 30S subunit	Efflux; target-protection proteins that dislodge tetracycline	Gram-negative and Gram-positives including MRSA
Watch	Beta-lactam	Anti-pseudomonal penicillin	Piperacillin + tazobactam	Cell wall synthesis inhibition: interact with cell wall catalysis enzymes	Antibiotic-degrading enzymes, beta-lactamases; mutations in cell wall proteins	Gram-negative and Gram-positives including *Pseudomonas aeruginosa*
Watch	Beta-lactam	Carbapenem	Imipenem + cilostatin Meropenem	Cell wall synthesis inhibition: interact with cell wall catalysis enzymes	Antibiotic-degrading enzymes, beta-lactamases; mutations in cell wall proteins	Gram-negative and Gram-positives, especially multi-drug resistant infections

WHO Category	Antibiotic Class	Sub-class	Drug name	Mechanism of action	Mechanism of resistance	Used against
Watch	Beta-lactam	3rd generation cephalosporin	Cefixime Ceftriaxone Cefotaxime Ceftazidime	Cell wall synthesis inhibition: interact with cell wall catalysis enzymes	Antibiotic-degrading enzymes, beta-lactamases; mutations in cell wall proteins	Gram-negative and Gram-positives, especially for bacterial meningitis and sepsis
Watch	Beta-lactam	Penem	Faropenem	Cell wall synthesis inhibition: interact with cell wall catalysis enzymes	Antibiotic-degrading enzymes, beta-lactamases; mutations in cell wall proteins	Gram-negative and Gram-positives, especially multi-drug resistant infections
Watch	Fluoroquinolone		Ciprofloxacin Levofloxacin Moxifloxacin Norfloxacin	DNA replication inhibition: inhibit DNA gyrase	DNA gyrase mutations; efflux; Qnr production (protein that competes for fluoroquinolone binding site)	Gram-negative and Gram-positives, including anaerobes
Watch	Glycopeptide		Teicoplanin Vancomycin	Cell wall synthesis inhibition: bind D-Ala-D-Ala cell wall molecules	Acquisition of van gene cluster (replaces final D-Alanine in cell wall molecules to prevent glycopeptide binding and destroys normal D-Ala-D-Ala molecules)	Gram-positives e.g. *Enterococcus faecalis* and *Clostridioides difficile*
Watch	Lincosamide		Clindamycin	Protein synthesis inhibition: inhibit translocation by binding 23S rRNA in 50S ribosomal subunit	Ribosomal mutations or modifications; efflux	Gram-positives and anaerobes, but not typically Gram-negative aerobes
Watch	Macrolide		Azithromycin Clarithromycin	Protein synthesis inhibition: inhibit translocation by binding 23S rRNA in 50S ribosomal subunit	Ribosomal mutations or modifications; efflux	Gram-positives and anaerobes, but not typically Gram-negative aerobes
Reserve	Beta-lactam	Monobactam	Aztreonam	Cell wall synthesis inhibition: interact with cell wall catalysis enzymes	Antibiotic-degrading enzymes, beta-lactamases; mutations in cell wall proteins	Gram-negative aerobes, including *Pseudomonas aeruginosa*

WHO Category	Antibiotic Class	Sub-class	Drug name	Mechanism of action	Mechanism of resistance	Used against
Reserve	Beta-lactam	4th generation cephalosporin	Cefepime	Cell wall synthesis inhibition: interact with cell wall catalysis enzymes	Antibiotic-degrading enzymes, beta-lactamases; mutations in cell wall proteins	Gram-negative and Gram-positives
Reserve	Beta-lactam	5th generation cephalosporin	Ceftaroline	Cell wall synthesis inhibition: interact with cell wall catalysis enzymes	Antibiotic-degrading enzymes, beta-lactamases; mutations in cell wall proteins	Gram-negative and Gram-positives including MRSA
Reserve	Peptide: cyclic lipopeptide		Daptomycin	Membrane function disruption: inserts into cytoplasmic membrane, disrupts integrity and triggers rapid cell death	Mutations in genes that alter the target or entry of daptomycin into cytoplasmic membrane	Gram-positives e.g. vancomycin-resistant enterococci and MRSA
Reserve	Fosfomycin		Fosfomycin	Cell wall synthesis inhibition: inhibits the first step of cell wall synthesis	Mutations in *murA*, the protein that initiates cell wall synthesis; mutations in membrane transporters, acquisition of fosfomycin-inactivating enzymes	Gram-negative and Gram-positive bacteria, especially for antibiotic-resistant UTIs
Reserve	Glycylcycline		Tigecycline	Protein synthesis inhibition: binds 30S ribosome, blocking tRNA entry	Efflux	Gram-negative and Gram-positive, including tetracycline-resistant bacteria
Reserve	Oxazolidinone		Linezolid	Protein synthesis inhibition: binds 23S rRNA in 50S ribosomal subunit and interacts with tRNAs to prevent protein synthesis initiation	Efflux (especially in Gram-negative aerobes); ribosomal mutations	Gram-positives e.g. vancomycin-resistant enterococci and MRSA, Gram-negative anaerobes
Reserve	Peptide: polymixins		Colistin Polymyxin B	Outer membrane disruption: bind lipopolysaccharide to destabilise OM and IM, induce osmotic imbalance and oxidative stress	Chromosomal mutations; acquisition of genes that alter the lipopolysaccharide	Multi-drug resistant Gram-negative aerobes

1.3.1 Access antibiotics: first or second line of defence

Access antibiotics are those that, according to the WHO, are first or second line drugs for common infectious diseases and should be: widely available, affordable, appropriately formulated and quality controlled[18]. First line antibiotics are typically narrow spectrum: they only target a limited range of bacteria, such as the intended pathogenic bacterium and possibly some close relatives[18]. They should also have low resistance potential and positive benefit-to-risk ratios[18]; i.e., the risk of harm to the patient is low compared to the potential treatment of their infection. Broad spectrum antibiotics, that target a wider range of bacteria, are generally used as second line options, as are drugs with higher resistance potentials or less positive benefit-to-risk ratios[21] (e.g. they might have more common or more serious side effects).

Aminocyclitol antibiotics, such as spectinomycin, bind the 30S subunit of the bacterial ribosome, disrupting protein synthesis. **Aminoglycosides**, such as amikacin and gentamicin, act via the same mechanism. To do this, they must enter the cell through its cytoplasmic membrane – a process that requires energy dependent active bacterial transport[24]. This in turn requires oxygen and an active proton motive force; therefore these drugs are only typically effective for bacteria capable of aerobic respiration[24]. Aminocyclitols and aminoglycosides are considered inactive against anaerobic bacteria, which have adapted to survive without oxygen and do not have the usual required components to transport these drugs into their cells[24]. Therefore, these drugs are considered broad-spectrum but only for a range of aerobic bacteria. Although these antibiotics are known to be toxic to humans and can cause hearing loss, they are considered first line treatments for respiratory infections in cystic fibrosis patients[26]. They are also used to treat multiple-drug resistant (MDR) infections, often as part of combination therapy with beta-lactams, since these classes of antibiotics have a

synergistic effect when used together[26]. Aminoglycoside resistance is mediated by ribosomal mutations, enzymes that modify the antibiotic, or enzymes that modify the target[27].

Chloramphenicol, an **amphenicol**, is considered a broad-spectrum antibiotic and interacts with the 23S rRNA of the 50S ribosomal subunit, preventing protein synthesis[24]. It is a first line option for eye infections such as conjunctivitis, which can be caused by bacteria including *Haemophilus influenzae*, *Streptococcus pneumoniae* and *Moraxella catarrhalis*[28]. Chloramphenicol resistance is mediated by ribosomal mutations or chloramphenicol-modifying enzymes. It can also be due to efflux-mediated resistance[29]. Mutations in components of efflux systems can increase the affinity for certain molecules including antibiotics, resulting in increased resistance[29].

Beta-lactams are a diverse group of antibiotics with several sub-classes. This includes penicillins, cephalosporins, carbapenems and monobactams. Beta-lactam antibiotics have a chemical structure that is similar to that of D-Alanyl-D-Alanine (D-Ala-D-Ala), a 'building block' of bacterial cell walls[30]. This means the beta-lactams can interact with the enzymes that catalyse the cell wall synthesis, preventing proper cell wall synthesis and resulting in osmotic instability and bacterial cell death[30]. The different sub-classes of beta-lactams are categorised by the modification to the standard beta-lactam structure that defines that particular group. **Penicillins** and **1st generation cephalosporins** are on the Access list of antibiotics. Both these sub-classes of beta-lactam are typically considered narrow spectrum. However, the spectrum varies from drug to drug: for example, amoxicillin is considered broad-spectrum or extended-spectrum compared to the original penicillin[31]. Penicillins, especially amoxicillin, are the first line drug of choice for common bacterial infections, such as dental, ear, respiratory and throat infections[32], hence their position as the most used drugs in the world[33]. Genes that encode beta-lactamase enzymes, which degrade beta-lactam antibiotics, are some of the most

common antibiotic resistance genes (ARGs): over 1000 different beta-lactamase genes are described[29]. Beta-lactams can be given alongside a beta-lactamase inhibitor, such as amoxicillin with clavulanate, to prevent a beta-lactamase enzyme from functioning properly[24]. This combination increases the efficiency of the amoxicillin. First generation cephalosporins, such as cefalexin and cefazolin, have often been used as second-line options as an alternative to penicillin, in the case of resistance or allergies[21]. Beta-lactam resistance can also be mediated by mutations or modifications in the antibiotic target (penicillin-binding proteins (PBPs) in the cell wall) or in porins, which allow the beta-lactam antibiotics to enter the cell[29].

The **folic acid metabolism inhibitor** combination of sulphamethoxazole and trimethoprim is active against many different bacterial infections[24]. Bacteria must synthesise *de novo* folates that can act as cofactors for various other biosynthetic pathways; without these cofactors their growth is inhibited[34]. Sulphonamides, such as sulphamethoxazole, and trimethoprim, a diaminopyrimidine antibiotic, inhibit enzymes involved in different stages of the folic acid biosynthetic pathway and thus this combination is considered synergistic[24]. It is a common first-line treatment for infections including urinary tract infections (UTIs), traveller's diarrhoea, methicillin-resistant *Staphylococcus aureus* (MRSA) skin infections, respiratory tract infections, and cholera[35]. Resistance to trimethoprim is mediated by mutations that increase the expression of trimethoprim targets so that the antibiotic is outnumbered and its effect limited[29]. Mutations that alter the targets themselves or the acquisition of genes that encode less sensitive targets are also known to result in trimethoprim resistance[29]. Sulphonamide resistance is common and mainly due to plasmid-borne genes encoding alternative enzymes that are less sensitive to the antibiotic[36]. This is a method of horizontal gene transfer (HGT): the movement of genes between bacterial cells that bypasses vertical transmission.

The **nitroimidazole**, metronidazole, inhibits DNA synthesis and causes DNA damage in anaerobic bacterial cells[35]. Strict aerobic bacteria lack electron-transport proteins with sufficient negative redox potential to activate the prodrug form once it is inside the cell, thus metronidazole has the opposite spectrum to aminoglycosides[37]. It is the first-line option for infections caused by anaerobic bacteria, such as dental abscesses (e.g. by *Prevotella* or *Streptococcus)*[38] or bacterial vaginosis caused by *Gardnerella vaginalis*[37]. Resistance is mediated by decreased uptake or altered reduction efficiency (that also leads to decreased uptake)[37]. Efflux and drug inactivation are also possible, as is the acquisition of *nim* genes from other bacteria via HGT. These *nim* genes encode alternative reductase enzymes to convert the prodrug into a non-toxic alternative[37].

The **nitrofuran**, nitrofurantoin, has multiple antibacterial properties – none of which are fully understood yet[39]. Like nitroimidazoles, it is a prodrug, and is activated inside the bacterial cell by the action of nitroreductases[39]. The intermediate metabolites produced bind to bacterial ribosomes and enzymes involved in DNA and RNA synthesis, plus other metabolic processes[39]. It is broad-spectrum with activity against both Gram-negatives and Gram-positives, though some *Klebsiella* and *Pseudomonas* species are intrinsically resistant[39], and it is a first-line treatment for uncomplicated, lower UTIs[39]. Resistance is thought to be mediated by mutations in nitroreductases[39].

Tetracyclines such as doxycycline are another type of protein synthesis inhibitor that act upon conserved sections of the 16S rRNA in the 30S subunit, preventing translation[24]. It has broad spectrum and is commonly used to treat sexually transmitted infections (STIs), Lyme disease, skin infections and MRSA[40]. Resistance is mediated by tetracycline-specific efflux pumps or target-protection proteins that dislodge tetracycline when it is bound to the ribosome[29].

1.3.2 Watch antibiotics: first or second line of defence with high resistance potential

Watch antibiotics have higher resistance potential and may be of broader spectrum[21]. These antibiotics are also featured on the Highest Priority Critically Important Antimicrobials (CIA) list[41], which is intended to ensure that these antibiotics are prioritised for stewardship strategies and used prudently.

Piperacillin with tazobactam is another penicillin/beta-lactamase inhibitor combination, as described earlier. *Pseudomonas aeruginosa* is an opportunistic, Gram-negative pathogen that can cause infections in humans, but is intrinsically resistant to penicillin antibiotics caused by the production of a beta-lactamase enzyme[42]. Piperacillin with tazobactam has **anti-pseudomonal** activity and therefore is a common option for the treatment of *P. aeruginosa* infections. This bacterium has become a major contributor to infections that are hospital-acquired and/or in immunocompromised or critically ill patients[43]. It is important then that antibiotics capable of treating these infections are only used when appropriate to limit the development of resistance and conserve their utility.

There are two other sub-classes of beta-lactams that are on the Watch list: **third generation cephalosporins** and **penems**. Third generation cephalosporins have a specific type of modification to the basic beta-lactam structure that increases their specificity to binding cell wall proteins[44]. They have a broader spectrum than previous generations, although again it varies between drugs. For example, cefotaxime and ceftizoxime have the best Gram-positive coverage of third generation cephalosporins[44], though they are typically used to treat Gram-negative infections in hospitals associated with meningitis and sepsis[44]. Penems, including carbapenems, have the broadest spectrum of activity of the beta-lactams and are relatively resistant to most types of beta-lactamases[45]; hence why they are often used to treat beta-

lactam resistant or MDR infections[45]. Carbapenem resistance can be conferred by single nucleotide polymorphisms (SNPs) that affect PBPs in the cell wall[46] or in the porin that enables carbapenem import into the cell[46], or by increased activity of efflux pumps[46]. The most clinically important cause of carbapenem resistance is carbapenemases enzymes[29].

Fluoroquinolone antibiotics, such as ciprofloxacin, are highly active against both Gram-negatives and Gram-positives[47]. Newer fluoroquinolones in particular, such as moxifloxacin, have very broad spectrums of activity, including against anaerobes[47]. They work by inhibiting DNA gyrase, an enzyme involved in DNA replication and transcription[24]. Fluoroquinolones are very commonly prescribed drugs, including for ear, gastrointestinal (GI) and respiratory infections and UTIs[47]. Resistance is commonly mediated by mutations in the target enzymes, as well as over-expression of efflux pumps and production of Qnr, a protein encoded by plasmids that competes for the fluoroquinolone binding site[29].

Glycopeptides are similar to beta-lactams in that they inhibit cell wall synthesis, however they do this by binding to the D-Ala-D-Ala molecules rather than the enzymes that process synthesis[24]. These antibiotics target Gram-positive organisms only; due to their size, they are typically unable to cross the outer membrane (OM) of Gram-negatives[48]. Teicoplanin is more potent than vancomycin, but both are crucial antibiotics used to treat infections caused by Gram-positive bacteria such as *Enterococcus* or *Clostridioides difficile*[48]. Resistance to vancomycin is mediated by the acquisition of a *van* gene cluster, which act to replace the final D-Alanine in the cell wall molecules with D-lactate or D-serine, preventing vancomycin binding[29]. The normal D-Ala-D-Ala molecules are also destroyed by the function of the van gene operon[29].

Lincosamides, such as clindamycin, and **macrolides**, such as azithromycin, have similar mechanisms of action. They inhibit translocation, an early stage in protein synthesis, by targeting the conserved sequences of the 23S rRNA in the 50S ribosomal subunit[24]. Lincosamides target Gram-positives and anaerobes, but typically not Gram-negative or strictly aerobic organisms[49]. They are often used to treat head, neck, respiratory, bone, soft tissue, abdominal, and pelvic infections[50]. Macrolides have a similar broad spectrum of activity, plus some Gram-negative aerobes or facultative anaerobes, but Enterobacteriaceae tend to be resistant due to having relatively impermeable cell walls[51]. They are also used for respiratory tract, skin and soft tissue infections[51]. Resistance to lincosamides and macrolides can be mediated by mutations in the ribosome, modifications of the ribosome (e.g. methylation), or by efflux pumps[29].

1.3.3 Reserve antibiotics: last resort

Reserve antibiotics are those that should only be used as a last resort; that is, in the case of a serious or life-threatening multi-drug resistant infection that is not responding to first or second line treatments[41].

There are three types of beta-lactam included on the reserve list: **fourth and fifth generation cephalosporins** and **monobactams**. Fourth and fifth generation cephalosporins have a broader spectrum of, and higher, activity[52] and are more resistant to extended-spectrum beta-lactamases[53] than previous generation cephalosporins. The fifth generation cephalosporin ceftaroline also has activity against Methicillin-Resistant *Staphylococcus aureus* (MRSA) and is now a vital drug in the treatment of MRSA infections[54]. In contrast, monobactams such as aztreonam have a narrow spectrum and are only active against Gram-negatives capable of

aerobic respiration, including *P. aeruginosa*[55]. These synthetic antibiotics are highly resistant to beta-lactamases and are used to treat multi-drug resistant Gram-negative infections[55]. However, there are certain beta-lactamases, such as TEM-3, that are capable of degrading aztreonam[29].

Daptomycin is a synthetic **cyclic lipopeptide** antibiotic first reported in 2002[56] with a unique mechanism of action. It inserts into the bacterial cytoplasmic membrane, where it disrupts the membrane functional integrity to trigger the release of intracellular ions and cause rapid cell death[57]. It has antibiotic activity against Gram-positive bacteria, such as vancomycin-resistant enterococci (VRE), MRSA and penicillin-resistant streptococci – for which there are few alternative treatments[57]. Daptomycin resistance is complicated but thought to be mediated by mutations in genes that alter the target or the entry of daptomycin into the cytoplasmic membrane[58]. Daptomycin-degrading enzymes exist in environmental bacteria but are not currently considered clinically relevant[59].

Fosfomycin (a simple form of phosphonic acid) was discovered from *Streptomyces* in 1969 and is unique in its mechanism of inhibiting the first step in bacterial cell wall synthesis[60]. It is active against both Gram-negative and Gram-positive bacteria and though it was uncommonly used for several decades, in recent years it has been revived to treat antibiotic-resistant UTIs[60]. UTIs account for a significant burden of hospital admissions and are increasingly resistant to first or second line antibiotics[61], and so reserving fosfomycin for use only when absolutely necessary is vital to continue treating these cases. Fosfomycin resistance is rare as there appears to be an associated fitness cost, but mutations in *murA*, the gene encoding the protein that initiates cell wall synthesis, can occur[62]. In addition, mutations that alter membrane transporters can prevent fosfomycin entering the cell, and bacteria can acquire plasmid-encoded genes that inactivate the drug[62].

Tigecycline is the first of the synthetic class of **glycylcycline** antibiotics, developed in the 2000s[63]. Glycylcyclines feature a modified tetracycline structure; tetracycline resistance is now rife since this antibiotic is relatively old and very widely-used[64]. The glycylcycline structure binds to the bacterial 30S ribosome, blocking tRNA entry and preventing protein synthesis[63]. Tigecycline is broad spectrum, including activity against drug-resistant Gram-positive infections and even some resistant to tetracycline, despite any presence of tetracycline-specific resistance mechanisms[65]. However, efflux pumps involved in MDR and point mutations can sometimes cause tigecycline resistance[29].

The **oxazolidinone** linezolid is another synthetic antibiotic, from the late 1980s, which interferes with protein synthesis in two separate stages: by binding the 23S rRNA in the 50S ribosomal subunit and by interacting with tRNAs to prevent protein synthesis initiation[66]. Like daptomycin, it is used to treat antibiotic-resistant Gram-positive infections such as VRE and MRSA but also has activity against Gram-negative anaerobes[67]. Gram-negative aerobic or facultative anaerobic bacteria can contain efflux pumps that are able to pump out linezolid and resist its action[67]. Resistance is also mediated by ribosomal mutations[29].

Polymyxins are non-ribosomal cyclic lipopeptides, originally discovered in 1947 during "the golden age of antibiotic discovery" before falling out of use due to nephrotoxicity. They have been revisited more recently due to the increasing cases of MDR infections[68]. Polymyxins bind to the lipopolysaccharide in the OM of Gram-negative bacteria, causing outer and subsequently inner membrane destabilisation[68]. They are also reported to induce osmotic imbalance and oxidative stress, resulting in cell lysis[68]. Polymyxins are not susceptible to the activity of efflux pumps[24], hence why they are such good options to selectively treat serious Gram-negative aerobic or facultative anaerobic infections that are resistant to other

antibiotics[69]. Colistin resistance is mediated by chromosomal mutations or acquisition of genes that alter the lipopolysaccharide[27].

From this information discussed so far, we can make two important conclusions: firstly, beta-lactam antibiotics feature in all three categories of antibiotics, encompass a large number of drugs, are among the oldest clinically-used antibiotics and are still the most prescribed drugs in the world. This underlines their critical importance in modern human medicine. Secondly, that resistance to all these clinically relevant antibiotics among bacterial pathogens has emerged[5], including to last resort antibiotics. This means that cases are arising where an infection cannot be treated with any available antibiotics – sometimes resulting in the patient's death[70]. For example, a woman died in the United States in 2017 following a carbapenem-resistant Enterobacteriaceae infection: the strain of *K. pneumoniae* she was infected with was resistant to 26 antibiotics available in the country at the time[70]. This highlights the increasing severity of antibiotic resistance.

1.4 Bacterial genetics and antibiotic resistance

There are three main ways in which bacteria can be resistant to antibiotics that are explained by their genetics: intrinsic resistance, where the bacteria are naturally resistant; DNA mutations, where mutations in the genome lead to increased resistance; and antibiotic resistance genes, where entire genes can explain increased resistance. These will now each be discussed in more detail.

1.4.1 Intrinsic resistance

There are a number of biological differences between bacteria that mean not all antibiotics are effective against all bacteria; this natural resistance is also described as intrinsic resistance. For example, Gram-negative bacteria have an OM, which Gram-positives do not possess[71]. The OM prevents access of glycopeptide antibiotics such as vancomycin to Gram-negative bacteria such as Proteobacteria[71]. Moreover, metabolic differences (e.g. aerobic versus anaerobic) can explain intrinsic resistance. For example, efflux pumps, that can pump out antibiotics from bacterial cells, are common in aerobic and facultatively anaerobic bacteria such as Proteobacteria[29]. As they require oxygen and active transport to function, they are uncommon in anaerobic bacteria. The decision of which antibiotic to prescribe must take these differences into consideration, along with a number of other factors relating to the disease-causing bacterium, the antibiotic and the patient. However, these intrinsic resistances are not involved in the development or spread of antibiotic resistance as they occur naturally; instead, these processes are caused by changes in the bacterial genome[72].

1.4.2 DNA mutations

One key way in which antibiotic resistance can develop is by changes in the DNA sequence, or mutations. Mutations occur naturally: when cells replicate and divide, DNA must also be replicated, but the process is not error-free[73]. If incorrect nucleotides are incorporated into the new DNA strand being synthesised, this could have an impact on the function of that particular DNA sequence. For example, a bacterial cell with a mutation in a gene that encodes the target of an antibiotic may have a reduced affinity for how the antibiotic binds to said target[72]. In the presence of that antibiotic, that cell is less likely to be killed by the antibiotic

and can divide to produce more cells featuring the same DNA mutation. In this way, a more resistant strain can replace and take over the previous population of bacterial cells; thus, antibiotics select for bacteria that are more tolerant of or resistant to the antibiotic being used. The problem is magnified when we consider the short generation times of bacteria: *E. coli* can divide in approximately 20 minutes, leading to millions of cells in a matter of hours[74]. This allows antibiotic-resistant bacteria to be selected for in a short time period and also for mutations to accumulate and potentially become fixed in the population[73]. Mutations can impact antibiotic resistance in several ways: they can alter the target of the antibiotic so that the two can no longer interact; they can alter an efflux pump to have increased activity and reduce the concentration of antibiotic in the cell; they can increase the amount of the target protein in the bacterial cell so the impact of the antibiotic is reduced; or increase the expression and production of antibiotic-degrading enzymes. Using sufficiently high antibiotic dosing regimens can prevent a sub-population of cells with antibiotic resistance-conferring mutations from surviving[75,76].

1.4.3 Antibiotic resistance genes and horizontal gene transfer

In addition to DNA mutations, bacteria can harbour antibiotic resistance genes (ARGs): genes that encode proteins that are capable of modifying or degrading antibiotics or their targets[72]. One of the most common examples is a beta-lactamase, an enzyme capable of degrading beta-lactam antibiotics (the class that includes penicillins). There are both different sub-classes of beta-lactam antibiotics and beta-lactamase enzymes. Different beta-lactamases can have differing levels of activity against sub-classes of beta-lactam antibiotics[25]. This means that beta-lactam antibiotics belonging to a different sub-class (e.g. a cephalosporin) could be used

to treat infections that are resistant to penicillin via a penicillin-degrading beta-lactamase. Modifying and degrading enzymes exist for different types of antibiotics, as outlined in section 1.3 and Table 1.2.

The critical issue with ARGs is that they can be transferred by HGT; there are three main processes through which this can take place. Firstly, there is transformation, where bacteria that are "competent" can take in DNA from their environment that is self-replicable or is incorporated into their genome[72]. Competence can vary between species and even within a single bacterial cell depending on its replication cycle[72]. Secondly, there is transduction, which involves the packing of bacterial DNA into bacteriophages (viruses that infect bacteria)[72]. When the phage infects another bacterium it inserts the bacterial DNA inside it into the new host bacterium, where it can be incorporated into the bacterial chromosome[72]. Thirdly, there is conjugation, involving the direct cell-cell contact or bridge-like connections between bacteria, allowing for DNA to be transferred from the donor cell to the recipient cell[72].

The movement of ARGs is therefore a major contributing factor to the spread of antibiotic resistance. ARGs are not always transferred alone but can reside in mobile genetic elements (MGEs)[77]; regions of DNA that contain more than a single gene and even elements responsible for their own replication or movement. Such MGEs include:

- Integrative and Conjugative Elements (ICEs): also known as conjugative transposons – DNA sequences that are integrated into the host bacterial genome and encode their own functional conjugation systems
- Integrons: sections of DNA that easily recombine to capture new DNA
- Plasmids: self-replicating, circular extra-chromosomal DNA transferred by conjugation
- Phages: bacterial viruses that can integrate into the host bacterial genome

- Transposable elements: transposons and insertion sequences are DNA sequences that can move position within a genome

The conjugative MGEs are especially relevant to the spread of antibiotic resistance as they aid the transfer between unrelated bacteria. A single MGE can contain multiple ARGs, meaning that a bacterium can acquire resistance to multiple antibiotics at once, even in the absence of selective pressure from those antibiotics. Furthermore, not only does antibiotic use introduce a selective pressure that selects for mutations conferring antibiotic resistance, but it also promotes mutations and HGT[78]. Antibiotic exposure can induce the SOS DNA damage response in bacteria, which arrests the cell cycle and allows any DNA damage to be repaired[79]. SOS-induced DNA damage repair is especially error prone, resulting in an increased mutation rate[79]. In addition, the SOS response is linked to increased rates of conjugative transfer[79]. Thus, antibiotic therapy both promotes and selects for a variety of antibiotic-resistance conferring changes in bacterial genomes.

1.5 Antibiotic resistance is a major global issue

Antibiotic resistance is an ancient strategy used by bacteria to survive competition from other microorganisms that produce antibiotics[5]. It is, therefore, a perfectly natural biological occurrence. However, it is problematic when it occurs in clinical isolates of bacterial infections. Antibiotic resistance is a global threat to human health that is increasing in both prevalence and severity, and is considered one of the most pressing issues of the 21st Century[80]. Some estimates place the mortality rate of drug-resistant infections (DRIs) at 10 million deaths per year by 2050[80]. The highest burden of DRIs is in developing countries, where antibiotics are becoming increasingly more accessible and are often available without a prescription[81]. In

these countries, proper sanitation (e.g. clean water, basic toilets, good hygiene practises and waste disposal) and strict antibiotic guidelines are often lacking[82]. South-East Asia in particular is a global hotspot for emerging infectious diseases due to these reasons, plus small-scale animal production and high livestock densities[83]. As a result, infections are more easily transmitted and there is a strong selective pressure for the evolution of resistant bacteria[82]. For example, enteric pathogens isolated from Vietnamese children suffering from Acute Watery Diarrhoea (AWD) are often MDR[84]. The issue is also important in developed countries: in the USA, 23,000 people die of DRIs annually[85].

Of particular concern are hospital-acquired infections (HAIs), infections acquired whilst receiving other medical treatments that total over 1.7 million cases and almost 99,000 related deaths per year[86]. The leading causes of these are the ESKAPE pathogens: *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* and *Enterobacter* spp.. These contribute significantly to healthcare and economic burdens and are often reported as resistant to multiple drugs[87]. The latter four of these ("KAPE") are all members of the phylum Proteobacteria; of 538 species of bacterial pathogens[88], 43 % (236) belong to Proteobacteria – the biggest contribution of any phyla (Fig. 1.3). Antibiotic resistance in pathogenic Proteobacteria is therefore especially relevant and understanding the spread of antibiotic resistance, plus discovering alternative or novel treatments for MDR pathogenic bacteria, is of high priority[89]. Indeed, the top three microbial infections of critical priority for new treatment R&D according to the World Health Organisation (WHO) involve three of the KAPE Proteobacteria pathogens: *A. baumannii*, *P. aeruginosa*, and Extended-Spectrum Beta-lactamase (ESBL)-producing *Enterobacteriaceae* (and specifically isolates that are resistant to carbapenems)[90].

Carbapenems are typically reserved for infections that are resistant to multiple other antibiotics, including other beta-lactams. As mentioned, beta-lactam antibiotics are some of the most commonly used drugs in the world, and beta-lactamases are some of the most common and well-studied types of ARGs[91]. These often exist on MGEs, and have contributed significantly to the burden of DRIs. For example, the aminopenicillin amoxicillin is one of the

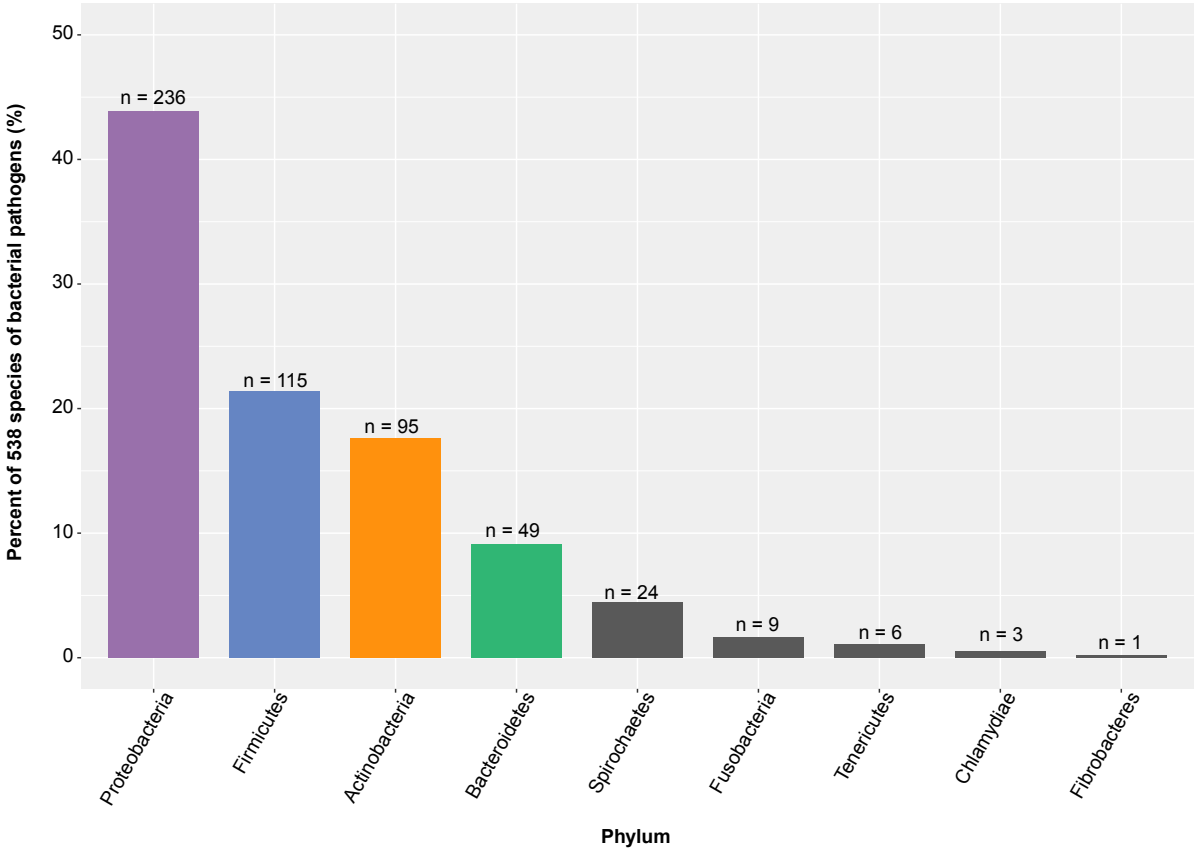


Figure 1.3. Proportions of 538 species of bacterial pathogens belonging to particular phyla. The list of bacteria species was taken from Taylor 2001[88] and the phyla they belong to identified with the online NCBI taxon identifier tool (https://www.ncbi.nlm.nih.gov/Taxonomy/TaxIdentifier/tax_identifier.cgi). The bacterial pathogens belong to nine different phyla, with over 40 % belonging to Proteobacteria and 20 % to Firmicutes. The numbers above the phyla names refer to the number of species observed in each phyla to be bacterial pathogens according to Taylor 2001[88]. The four coloured phyla highlight the four key phyla of the human gut microbiota.

most prescribed antibiotics globally due to its role as a first-line drug in the treatment of otitis media ear infections[33]. These infections are the most frequent reason doctors in the USA prescribe antibiotics[33] and are often caused by bacteria such as the Gram-negative,

facultatively anaerobic, Proteobacteria *E. coli*[92]. Nowadays, up to 95 % of *E. coli* isolates in national collections are resistant to amoxicillin (Fig. 1.4). Although there are a wide range of enzymes that confer extended-spectrum beta-lactam resistance[93], the highly mobilisable CTX-M genes are the most common and important types of ESBL due to their ease of dissemination and prevalence in human, animal and environmental isolates[94]. ESBLs were originally defined as enzymes that could hydrolyse the beta-lactam ring in most beta-lactam sub-classes except carbapenems[93]. However, some ESBLs, such as OXA-23, OXA-40 and OXA-48, have been reported as conferring resistance to carbapenems[95].

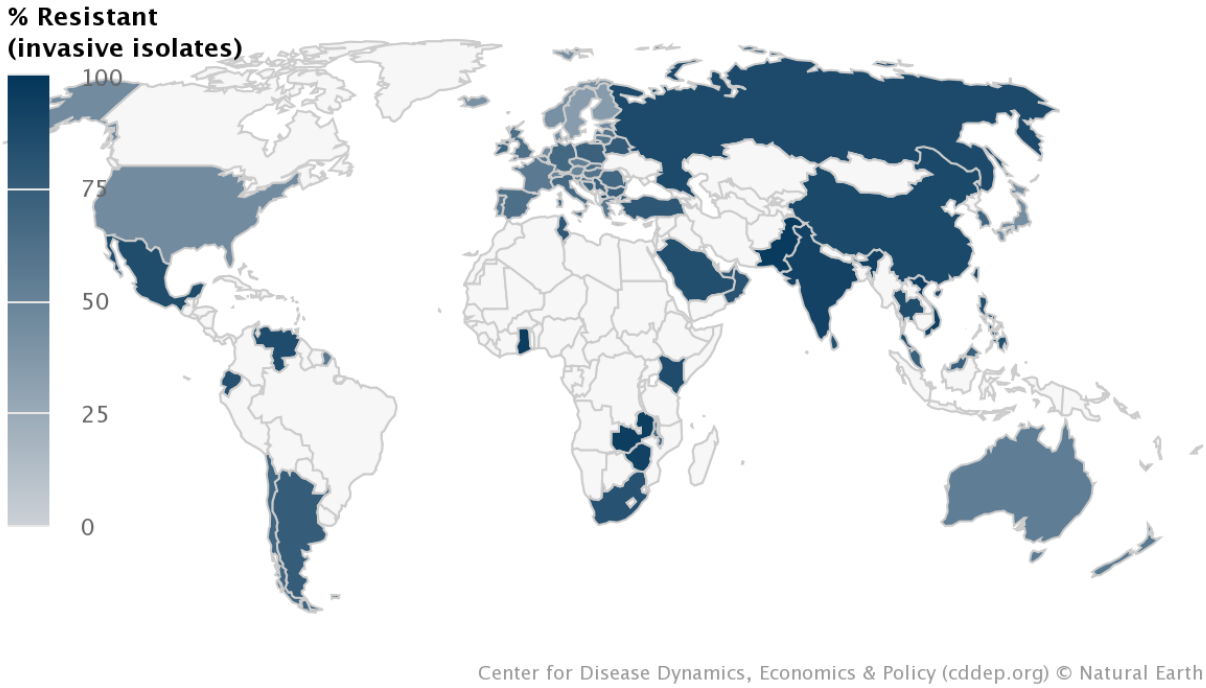


Figure 1.4. Global aminopenicillin (including amoxicillin) resistance in *Escherichia coli*. Map created at https://resistancemap.cddep.org/ on 02/03/2018. Data curated by the Center for Disease Dynamics, Economics and Policy and includes aggregated resistance rates for isolates (includes intermediate resistance) from blood and cerebrospinal fluid (i.e., invasive) from inpatients of all ages. Because of differences in scope of collections and testing methods, caution should be exercised in comparing across countries. Full data available online: The Center for Disease Dynamics Economics & Policy. ResistanceMap: Antibiotic Resistance. 2018. https://resistancemap.cddep.org/AntibioticResistance.php. Data accessed: March 02, 2018.

Carbapenem-resistant Enterobacteriaceae (CRE) infections have increased by 3 % in the USA between 2001 and 2010[90] and have high mortality rates of up to 48 %[96] due to the limited

treatment options. In addition, carbapenem-resistance conferred by mobile carbapenemases is especially important clinically due to their specific ability to hydrolyse the beta-lactam ring in carbapenems and propensity to spread[97]. One treatment option in these cases is colistin, but even resistance to this drug is rapidly spreading. It was previously thought that colistin resistance only involved chromosomal DNA mutations[98]. However, the first plasmid-mediated colistin resistance gene *mcr-1* was identified in 2015[99]. MCR-1 is an enzyme that modifies the colistin target and thus reduces the drug-target interaction, first discovered in *E. coli* isolates from China[99]. Eight *mcr* variants have now been identified[100], at least two of which are globally distributed[101]. Therefore, treatment options are severely limited and emerging colistin resistance in carbapenem-resistant infections is one of the most pressing concerns for global health. The emergence and global spread of mobile carbapenemases and *mcr* genes in clinical isolates of pathogenic bacteria highlights the extent and severity of horizontal antibiotic resistance gene transfer (HGT).

1.6 Antibiotic misuse and overuse: a One Health problem

Many antibiotic treatments are unnecessary, unregulated, or do not correctly follow guidelines[102], leading to increased opportunities for bacteria to develop or spread antibiotic resistance. Underdosing is a particular problem, where bacteria are exposed to a concentration of antibiotics too low to kill them. This can occur if the concentration required to kill the infectious bacterial cells is underestimated, or if the person prescribing, selling or purchasing antibiotics is doing so without consulting official guidance. Antibiotics are now available over-the-counter, without prescription, in countries across the world, meaning someone can choose to self-medicate with antibiotics who may not understand or receive instructions on the correct dosage to take. In addition, some people stop taking antibiotics

once they start to feel better, instead of finishing the advised course of therapy, which may lead to sub-optimal doses of antibiotic that could select for bacterial cells more tolerant to or better able to resist the antibiotic[75]. This 'underdosing' phenomenon was even warned about by Fleming in his Nobel Prize acceptance speech in 1945[103]:

'It is not difficult to make microbes resistant to penicillin in the laboratory by exposing them to concentrations not sufficient to kill them, and the same thing has occasionally happened in the body. The time may come when penicillin can be bought by anyone in the shops. Then there is the danger that the ignorant man may easily underdose himself and by exposing his microbes to non-lethal quantities of the drug make them resistant.' – Alexander Fleming, 1945.

However, there is uncertainty over the guidelines for prescription doses and length of treatments and how important it is to adhere to these rules[75]. Indeed, there is evidence that taking antibiotics for longer is more likely to result in increased resistance than shorter treatment durations[75,104]. Clinical trials investigating end points such as fever resolution[105] as a guide for when to stop antibiotic treatment are recommended to better understand and develop antibiotic prescribing guidelines that reduce the risk of bacteria developing resistance.

The overuse of antibiotics is also problematic – for example in the USA, up to 50 % of antibiotic use is reportedly unnecessary[106]. Reasons for overuse include: prescribing antibiotic therapy despite lack of evidence of bacterial infection or suspected non-bacterial infection; lack of prescription required to obtain antibiotics; easy access to antibiotics (e.g. over-the-counter)[107]. The contribution of these issues varies from country-to-country, leading to variable rates of antibiotic use between nations. Antibiotic use can be measured in defined

daily doses (DDD), which represents assumed average maintenance dose per day for a drug used for its main indication in adults[108]. For example, the DDD for amoxicillin administered orally is 1.5 g per day[109]. DDD can be compared across a number of factors such as inpatients in a hospital, or population per country, e.g. DDD per 1000 people per day. In the UK antibiotics can only legally be obtained with a prescription from a licensed medical practitioner, and there are national campaigns to reduce the unnecessary or inappropriate use of antibiotics. In 2015, the UK had a total 8696 DDDs per 1000 people per day across all antibiotics measured (Fig. 1.5) – this is equivalent to 8.696 DDDs of antibiotics per person per day.

In contrast, Turkey is one of the heaviest users of antibiotics in the world[110], and had 18,095 DDDs per 1000 people per day (18.095 DDDs of antibiotics per person per day). The reasons for such high usage are much the same as already described: poor medical education regarding antibiotics, pharmaceutical industry pressure and promotions, and lack of antibiotic policy or guidelines[111]. Thankfully Turkey, and most other countries, have brought in measures to reduce the overuse and misuse of antibiotics[112]. These measures, such as national campaigns to increase awareness and understanding of antibiotic resistance or more strictly regulate antibiotic use, are known as antibiotic stewardship[113]. This stewardship is designed to limit the development and spread of antibiotic resistance[113] and therefore conserve our antibiotics and prolong their lifespan as much as possible. Despite this, antibiotic use has generally increased over the last fifteen years (Fig. 1.5) and broad-spectrum penicillins such as amoxicillin continue to be the most used type of antibiotic.

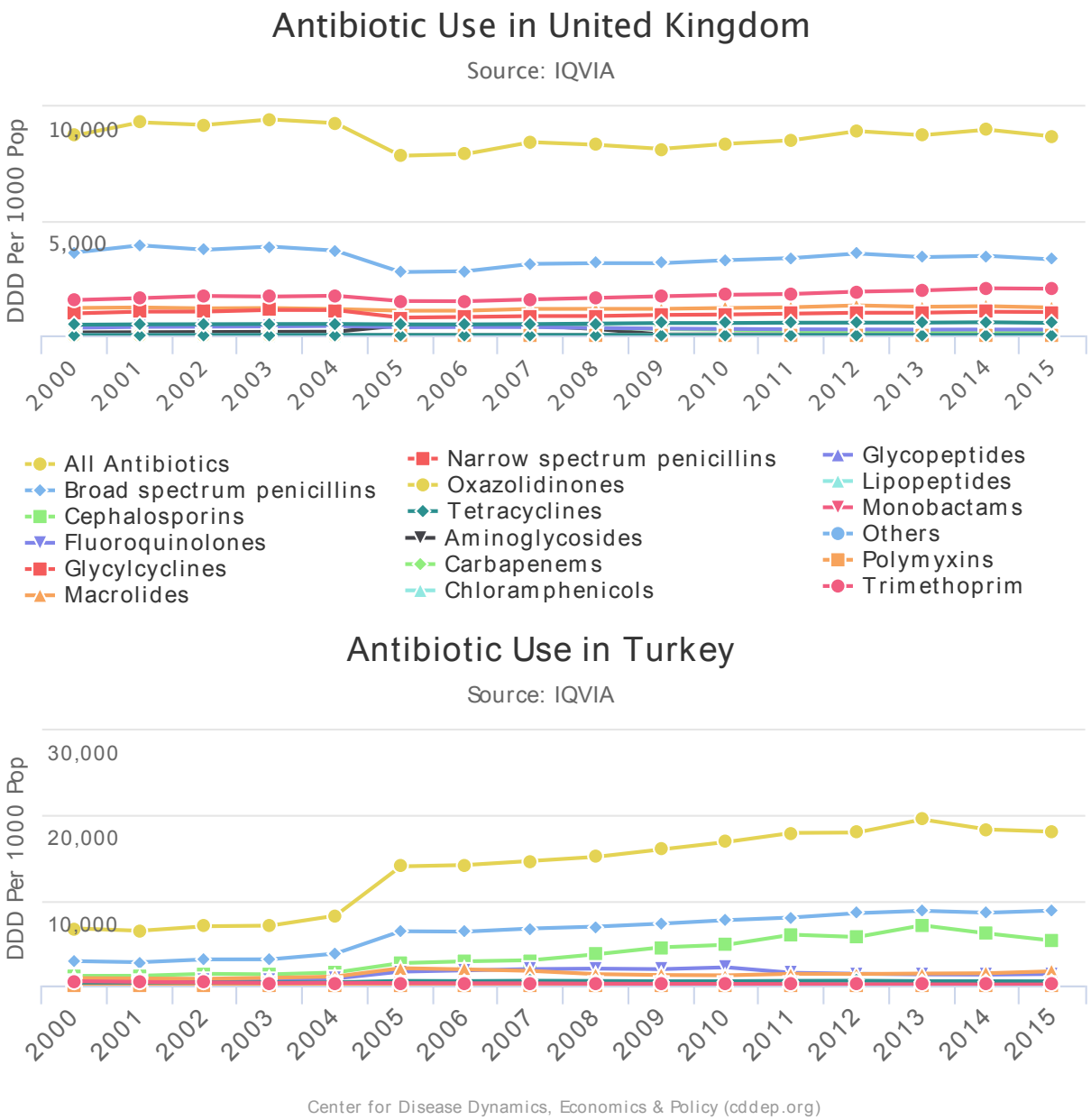


Figure 1.5. Usage of antibiotics in the United Kingdom and Turkey. The UK is the location of the current study and Turkey has the highest use of antibiotics of any country. Antibiotic use data are shown in defined daily doses (DDD) per 1,000 individuals (population, pop) per day. The Center for Disease Dynamics Economics & Policy. ResistanceMap: Antibiotic Use 2018. https://resistancemap.cddep.org/AntibioticUse.php. Date accessed: 01/03/2018.

However, it is not just antibiotic use in people that contributes to the problem of antibiotic resistance. It is a One Health problem, linking human health to the health of animals and the environment, where ensuring optimal health for all these components requires a collaborative effort across scientific disciplines and geography (e.g. locally, nationally, internationally and globally)[114]. This means we need to consider health – and specifically the use of antibiotics

and the presence of antibiotic resistance – in animals and the environment, alongside humans. For example, antibiotic use in agricultural animals is widespread[18,113,115]. Whilst often for medical purposes, antibiotics are also commonly misused in agriculture as growth promoters[116] – the exact mechanism is not well understood, but antibiotics are thought to reduce the amount of bacteria in the gut microbiome of livestock animals, reducing competition for energy derived from the animal's feed[116]. The doses used for these "medically unnecessary" purposes is typically not controlled[115] and so animals can receive sub-inhibitory concentrations[117], creating conditions that select for antibiotic-resistant bacteria.

Clinical cases of antibiotic-resistant infections have been acquired from agricultural animals: for example, drug resistant *Salmonella* infections from poultry date back to the 1960s[118]. This "farm-to-fork" hypothesis proposes that livestock carry antibiotic-resistant bacteria that spread to humans through e.g. direct contact or contaminated animal products[99,119]. If these resistant bacteria enter a human's GI tract, they may potentially cause an infection requiring antibiotic treatment, where the first-line drug of choice may not work. In addition, several important ARGs were first identified in animal associated bacterial strains and have since spread to human-associated strains – including the *mcr-1* colistin resistance gene in colistin-resistant *E. coli* from pigs[99]. Therefore, antibiotic use in agricultural animals should also be monitored and reduced to what is only necessary for proper veterinary medicine. In some countries, such as Sweden, Denmark, the Netherlands and the UK, the use of antibiotics as growth promoters in agriculture is banned[115]. However, other countries (for example, China) continue to use antibiotics critical to human medicine – including the last-resort antibiotic colistin – for this purpose[113]. Clearly, there is still work to be done in reducing the unnecessary use of antibiotics in both humans and animals.

Since antibiotic resistance is a natural phenomenon, antibiotic-resistant bacteria are found in the environment. However, humans have increased the selective pressure on environmental bacteria to promote the development or spread of antibiotic resistance. This can be linked to agricultural use of antibiotics: faeces from livestock may contain antibiotic-resistant bacteria or antibiotic compounds themselves. Areas exposed to agricultural run-off can therefore contain concentrations of antibiotics that promote the development of antibiotic resistance in bacteria that should not have received antibiotic exposure[117]. Again, these may cause antibiotic-resistant infections – if someone ingests vegetables grown in contaminated soil[99,119], for example. Additionally, hospital waste can contaminate local water sources with antibiotics or antibiotic-resistant bacteria[120,121] – if someone went swimming in this water, they may also acquire an antibiotic-resistant infection[122]. Environments containing antibiotic-resistant bacteria and ARGs – which could be harboured by pathogenic and non-pathogenic, environmental bacteria that do not typically cause disease – are referred to as reservoirs of antibiotic resistance. In reservoirs of antibiotic resistance, it is possible that ARGs can move between diverse and multiple bacteria via HGT. Since HGT has contributed to the global dissemination of antibiotic resistance, it is important to understand its mechanisms and the environments in which it occurs. For example – where do pathogenic bacteria acquire ARGs from, and what other bacteria do they donate ARGs to?

1.7 The gut microbiome as a reservoir of antibiotic resistance

The gut microbiome, the complement of commensal microorganisms and their genes that live in the GI tract, was first proposed as a reservoir for antibiotic resistance in 2004[123]. The intestinal microbiota refers specifically to the microorganisms, including bacteria that reside within the GI tract. They make up part of the total human microbiome, along with the skin

microbiome, vaginal microbiome, and microbiomes of other organs and tissues. Humans first become colonised by microbiota during birth, for example from the vagina and faeces if born via the birth canal. Caesarean-born babies are less exposed to these sources of bacteria and are more likely to first be colonised by through skin-skin contact and contact with the environment[124]. We continue to be colonised as we age, from our diets and our environments, and our microbiome matures from its initial form into a mature, relatively stable state[124]. In the healthy adult, there is estimated to be approximately 300 species of commensals in the GI tract[125], of which the majority and most abundant belong to the Firmicutes and Bacteroidetes respectively[126]. Proteobacteria and Actinobacteria are two other common intestinal microbiota phyla, though typically less abundant and diverse in the gut microbiome than Bacteroidetes and Firmicutes[126]. These phyla vary in their respiration physiology: Actinobacteria are typically considered aerobic, with two genera described as anaerobic or facultatively anaerobic[127]; Bacteroidetes are typically considered anaerobic, though two families are thought to be strict aerobes[128]; Firmicutes includes taxa described as aerobic and others described as anaerobic; Proteobacteria are described as facultative anaerobes[129]. Since the gut contains a very low concentration of oxygen and is generally considered an anaerobic environment, most gut bacteria taxa are adapted to these conditions and are considered strictly anaerobic[130]. Facultatively anaerobic or microaerophilic bacteria are much less abundant in the gut than strict anaerobes[130].

Our intestinal microbiota are very important to our health: they help train our immune system[131], digest food and produce energy, metabolise waste products, and protect from certain diseases[132]. One particular example is providing colonisation resistance against bacterial infections such as *C. difficile*[133]; the gut microbiota can prevent this pathogen from colonising the gut and producing toxins which causes severe diarrhoea in

immunocompromised patients. Unfortunately antibiotic use, to treat or prevent bacterial infections including *C. difficile*, can also negatively affect our gut microbiota. This is because antibiotics, especially those considered broad-spectrum and delivered orally, can have a direct impact on our indigenous gut microbiota. This impact may vary at different life stages: for example, in a new-born baby to prevent infection, antibiotics can disrupt the normal colonisation process[134]. This is associated with the later development of immunological conditions such as asthma[134]. As an adult, antibiotic therapy can alter the colonisation resistance of the gut microbiota and leave the person at risk of a *C. difficile* infection[135].

We are exposed to antibiotics throughout our lives: during birth to prevent infection; during childhood to treat common infections such as ear or throat infections; more rarely in adulthood (perhaps for a wisdom tooth infection or a more severe but relatively rare infection such as bacterial meningitis); to prevent infection during labour, in immunocompromised Intensive Care Unit patients or the elderly, amongst others. Due to the use and overuse of orally-administered broad-spectrum antibiotics, our gut microbiota are also often under selective pressure to evolve antibiotic resistance[136]. There is a strong positive correlation between antibiotic consumption and proportions of antibiotic resistant bacteria and antibiotic resistance gene abundance in gut microbiomes[137]: for example, southern Europe consumes more antibiotics than Denmark and accordingly human faecal samples from Denmark have lower carriage of antibiotic resistance genes than samples from Spain or France[137]. In addition, due to the density and diversity of commensal bacteria, the potential for exposure to food-borne pathogens such as members of Enterobacteriaceae is high. It is possible transient intestinal pathogens transfer ARGs to the resident gut microbiota, rendering these ARGs theoretically accessible to any other pathogen that passes through[123], or vice versa. As just one example, conjugation events of plasmids carrying antibiotic resistance genes between

commensal *E. coli* and pathogenic *Salmonella* can occur[138]. Therefore, not only can gut microbiota acquire ARGs from pathogens, but they also harbour their own ARGs that can be transferred to other bacteria[123].

Indeed, many studies have identified the presence of ARGs in commensal gut bacteria[137,139-144] including genes with over 90% similarity to known ARG sequences from pathogenic bacteria in human stool and saliva samples[139]. Even in healthy infants under one month of age who had not been treated with antibiotics, ARGs conferring resistance to up to 14 antibiotics have been identified[141]. Tetracycline genes, such as *tetO*, *tetQ* and *tetW*, are particularly prevalent in gut microbiome samples collected from across the world[139,142]; this is thought to be due to the historical high usage of tetracycline in agriculture (now banned in the EU and USA)[137] and therefore associated with the farm-to-fork hypothesis. Macrolide and sulphonamide resistance genes (e.g. *ermB and sul2* respectively) are also reportedly common in gut microbiomes[137,141,142]; macrolides have also been used as agricultural growth promoters[137], whereas sulphonamides are very old types of antibiotics and so selective pressure for sulphonamide resistance in the gut has existed for a comparatively long period of time[137]. Clearly then, antibiotic resistance can occur in the gut even in the absence of antibiotic treatment, although antibiotic usage is associated with increase abundance or diversity of resistance genes and mutations.

Moreover, gut bacteria have been found to contain novel ARGs that have not previously been detected in pathogenic bacteria[139,144]. For example, ten novel beta-lactamase genes with between 35 and 61 % similarity to known beta-lactamases at the time were identified from human stool and saliva[139,144]. More recently, novel beta-lactamases have been detected in sewage sludge containing human faecal matter[145], including one found on a mobile element that is considered likely to be highly mobilisable between bacteria[145]. Thus, despite the

increasing evidence that the human gut microbiome harbours diverse ARGs, it is unlikely we have discovered all ARGs or mutations that cause antibiotic resistance. In addition, with studies typically focussing on Western or industrialised communities, antibiotic resistance in other communities is less well understood. Finally, since gut microbiome studies have often depended on culture-independent techniques, rare or lowly abundant members of the gut microbiota are less well studied than more abundant bacteria. As a result, the incidence, distribution and dissemination of ARGs within the gut microbiota is still not fully understood.

1.8 Studying antibiotic resistance in bacteria

Studying antibiotic resistance involves determining the sensitivity of particular bacterial taxa (e.g. isolate, strain, species, genus, family, order, class and phylum) to an antibiotic. This can involve measuring the physical response through culture and phenotyping an isolate-antibiotic combination, or studying the presence and prevalence of antibiotic resistance-conferring mutations or genes (genetic determinants of antibiotic resistance) in bacterial genomes or metagenomes. How these methods have been used to study antibiotic resistance, particularly in gut microbiota, will now be discussed in turn.

1.8.1 Phenotyping

Microbiology has its foundations in culturing bacteria and studying their physical characteristics (i.e., phenotypes). Pathogenic bacteria have come under much scrutiny because of their propensity to cause disease. As described earlier, the majority of clinical pathogens tend to be Gram-negative, facultatively anaerobic bacteria that are amenable to

culture, allowing their antibiotic resistance phenotypes to be measured and recorded on a large scale. This has resulted in huge quantities of data being collected regarding the variation in antibiotic resistance and susceptibility to various antibiotics in common pathogenic bacteria. This data can be used to can define cut off points of susceptibility or resistance, such as those by the European Committee on Antibiotic Susceptibility Testing (EUCAST), the Clinical Laboratory Standards Institute (CLSI) or the British Society of Antimicrobial Chemotherapy (BSAC). These consider bacteria as resistant if they require over a certain concentration of antibiotic to be killed, and bacteria that are killed with lower concentrations are considered susceptible. For example, the EUCAST guidelines (version 7.1, valid from March 2017) list the amoxicillin breakpoint MIC for Enterobacteriaceae as 8 mg/L: Enterobacteriaceae isolates with an amoxicillin MIC equal to or less than 8 mg/L are considered susceptible and those with MICs over 8 mg/L are considered resistant. This type of data also allows for antibiotic resistance over time and space and between related isolates to be monitored. This is critical for prescribing practices – if there is an increasing trend in resistance to an antibiotic by a bacterial pathogen in a particular area, then the local doctors can include this in their consideration when choosing which antibiotics to prescribe.

The situation is very different for bacteria that are typically considered non-pathogenic, such as gut bacteria, which until recently have been comparatively under-studied. Due to their adaption to conditions in the gut, including the requirement for anaerobic conditions[146], they have also been hard to culture and study physically. The spectrum of antibiotics is therefore often only based upon the testing of Gram-negative, facultatively anaerobic or aerobic pathogens plus a select few Gram-positive or anaerobic species[147], such as *Enterococcus faecalis* or *Bacteroides fragilis* (both opportunistic pathogens)[147]. Therefore, antibiotic breakpoints of resistance/susceptibility in commensal gut bacteria are not defined. Moreover,

it means we do not fully understand the impact of different antibiotics across the diversity of our gut microbiota. However, since gut bacteria can sometimes act as opportunistic pathogens (like *B. fragilis,* a common cause of post-operative infections[148]) and in combination with their role as an ARG reservoir, it is important to understand their variation in antibiotic sensitivity. In particular, it would be useful to know which antibiotics might be useful for treating infections caused by opportunistic pathogenesis by our commensal gut bacteria, which antibiotics we should avoid if we wish to limit harm to out gut microbiota, and what types of antibiotic resistance these organisms might contribute to spreading. Fortunately, with recent developments in culturing of gut bacteria[146,149-151], we can now culture over 90 % of species of gut bacteria found in an individual[146]. This offers an exciting opportunity to investigate phenotypic antibiotic resistance in gut bacteria at an unprecedented scale.

1.8.2 Genome sequencing

In addition to phenotypic antibiotic resistance, antibiotic resistance genotypes can also be studied. Genomic-based predictions of resistance is typically relied upon for the surveillance of antibiotic resistance in non-pathogenic bacteria, including intestinal microbiota[152]. Databases and tools designed for this purpose (summarised in Table 1.3) have been used to study antibiotic resistance genotypes of whole genomes, either from raw sequence reads or assembled contigs. These types of methods have become popular for monitoring or tracking antibiotic resistance genes in bacterial isolates over time and/or space. This is largely used for clinical isolates to see trends in increasing or more widespread antibiotic resistance; such as the Global Pneumococcus Project studying 20,000 isolates of *Streptococcus pneumoniae* that found resistance genotypes for five antibiotics of different classes were strongly associated

with the country of isolation[153]. Relatively few studies of this type have been performed in gut microbiota, due to a proportionate lack of multiple cultured isolates and thus whole genomes for a particular species: for example, a study of the genus *Bifidobacterium*, a dominant component of the early-life microbiome, assessed ten classes of antibiotic resistance genotypes and phenotypes of 91 isolates from 54 subspecies[154].

In addition, these described tools have been used in combination with whole genome shotgun metagenomics (sequencing all the DNA in a sample) to study the resistome[155] (the collection of antibiotic resistance genes, mutations and their precursors) in diverse microbiome environments – such as soils, wastewater or sludge, smog and sediments[156]. These whole genome shotgun metagenomic studies have demonstrated that antibiotic resistance genes are widespread throughout natural and human-associated environments[156]. However, the databases are often based upon knowledge from clinically relevant pathogens – which as demonstrated earlier in Figure 1.2 are predominantly Gram-negative, facultatively anaerobic Proteobacteria. Since the majority of commensal gut bacteria are strictly anaerobic Gram-positives[126,157], they are distantly related to clinically relevant pathogens and may harbour currently unknown antibiotic resistances[139]. These have the potential to become clinically relevant either through direct opportunistic pathogenesis of commensal species[158] or through the transfer of ARGs to pathogens[152].

Table 1.3. Summary of databases and tools for predicting antibiotic resistance genotypes from sequence data. A brief summary of each database and/or tool is provided alongside an overview of the advantages and disadvantages.

Resource	Type	Advantages	Disadvantages
Antibiotic Resistance Database (ARDB)[159]	• Database of resistance gene nucleotide sequences	• Large number of sequences • Extensive metadata	• Last updated 2009 • Redundancy
Antibiotic Resistance Identification by Assembly (ARIBA)[160]	• Tool	• Tool uses sequence reads • Detailed, customisable, easily interpretable output • Removes redundancy from databases used • Can identify resistance mutations including new variants	• Relies on input databases • Not really suited for metagenomic data
ARG-ANNOT[161]	• Database	• Extensive list of resistance gene sequences combined from multiple sources • Can be used to identify mutations	• Redundancy • Last updated 2017 • Data no longer appears to be available on website
Bacterial Antimicrobial Resistance Reference Gene Database	• Database of antibiotic resistance genes	• Combines data from several sources • Regularly updated	• Redundancy • Does not include resistance mutations
Bush-Jacoby Database[162]	• Database of beta-lactamases	• Highly curated database of sequences and metadata associated with beta-lactamases	• Website no longer appears to work
Comprehensive Antibiotic Resistance Database (CARD)[163,164]	• Database of resistance genes • Tools to identify resistance genes/mutations	• Regularly updated • Includes both genes (including all ARDB sequences) and mutations • Online tool to apply methods and command-line tool available • Extensive metadata • Can account for gene or mutation conferring resistance to multiple antibiotics • Tool has built in option to look for candidate novel resistance genes	• Redundancy • Ontology complex • Tool requires assembled genomes

Resource	Type	Advantages	Disadvantages
DeepARG[165]	• Database of resistance genes • Tool using machine learning to characterise and annotate resistance genes	• Online and command-line tool • Can be applied to sequence reads (e.g. for metagenomic samples) and assembled genomes • Highly curated database • Does not just rely on best-hit	• Database last updated 2017 • Does not account for mutations
Lactamase Engineering Database[166]	• Database of beta-lactamases	• Highly curated database of sequences and metadata associated with beta-lactamases	• Only two families of beta-lactamases included • Unsure when last updated, website does not appear up to date or completely functional
MegaRes[167]	• Database of resistance gene nucleotide sequences • Tool to identify resistance genes	• Simple ontology and metadata • Good for population level analysis (e.g. count-based analyses in metagenomic samples) • Easy to interpret results • Tool uses sequence reads • Non-redundant	• Does not include resistance mutations • Cannot account for gene or mutation conferring resistance to multiple antibiotics • Last updated 2016
PointFinder[168]	• Database of chromosomal resistance mutations • Tool to identify mutations	• Good for chromosomal point mutations • Online or command-line tool, can be run together with ResFinder • Regularly updated	• Focuses on small range of species and genes • Difficult to analyse many samples at once
ResFams[169]	• Database of profile HMMs of resistance genes	• Combines resistance genes from multiple sources • Includes HMMs based on resistance genes identified in functional metagenomic screens • Can be used to identify candidate novel resistance genes	• Last updated 2015 • Mainly intended for use for functional metagenomic screens, rather than surveillance of resistance genes

Resource	Type	Advantages	Disadvantages
ResFinder[170]	• Curated database of resistance genes • Tool to identify resistance genes	• Regularly updated • Online tool, can be run together with Pointfinder • Add-on function for functional metagenomic screens	• Difficult to analyse many samples at once
Resqu https://1928diagnostics.com/resdb/	• Database of resistance genes	• Highly curated database • Non-redundant • DNA and protein sequences	• Does not account for mutations • Focuses only on genes that can be transferred horizontally • Database does not appear to be publicly available • Last updated 2013
SRST2_ARG-ANNOT[171]	• Curated database of resistance genes • Tool to identify resistance genes or mutations in genomes	• Uses sequence reads • Removes redundancy from ARGANNOT • Can identify mutations	• Relies on ARGANNOT database • Last updated 2017 • Can only identify pre-defined mutations or variants
Structured ARG database[172]	• Database of resistance genes • Tool to identify resistance genes	• Nucleotide sequences and profile HMMs • Online analysis pipeline and command-line tools • Good for metagenomic samples • Claims to frequently update	• Redundancy • Does not account for mutations • Last updated 2018 • Not designed for whole genomes

Functional metagenomic studies, such as Sommer *et al*. 2009[139], have shown that ARGs unlike those seen in pathogens exist in gut bacteria. This technology clones DNA fragments that have been extracted directly from an environmental sample, such as a stool sample, into another bacterium, often *E. coli,* and screened for antibiotic resistance phenotypes by plating on agar containing antibiotics. This bypasses the issue of culturing gut microbiota, making functional metagenomics a very powerful tool for studying a community in a relatively unbiased way. Indeed, this method has been important in understanding antibiotic resistance genes in a range of microbiomes and environments, including but not limited to: chicken guts[173]; uncontacted Amerindians[174]; faeces from domesticated animals as well as soil, water and sanitation facilities of rural villages and "shanty" towns in Peru[143]; seawater[175]; and Alaskan soil[176]. From functional metagenomic studies, we have learned that antibiotic resistance genes not currently found in pathogenic bacteria are both diverse, abundant and widely distributed. In addition, bacteria do not have to have been exposed to clinical antibiotics to harbour antibiotic resistance genes, even ones that have been found in pathogenic bacteria. However, these studies tend to rely on cloning of genes into *E. coli*[143] and may miss genetic determinants that cannot be expressed in this organism; plus, there is extensive bacterial diversity across the planet that has not been studied in depth[177]. This leaves the possibility that there are still more unknown antibiotic resistance determinants waiting to be discovered – not just in the gut, but in the Earth's total microbiome.

Genomic methods can be combined for more in-depth analyses of antibiotic resistomes. An important study of reservoirs of antibiotic resistance used 16S rRNA gene sequencing (amplification and sequencing of variable regions in the 16S rRNA gene, used as markers for bacterial species or genera), functional metagenomics and whole genome shotgun metagenomic sequencing to explore the similarity of resistomes between different

environments[143]. This included the human gut, domesticated animal gut microbiomes, soil, water, and sanitation facilities[143]. This multi-genomics approach allows antibiotic resistance genes to be quantified in terms of their relative abundance and associated with particular taxa, rather than just observing their presence. The resistomes of the different samples correlated with the phylogenetic diversity of each sample across an ecological gradient but certain antibiotic resistance genes were able to move between more diverse habitats and were linked to mobile genetic elements[143]. For example, the sulphonamide resistance gene *sul2* was found in 50 % of samples from six of seven environments studied and appeared to be localised in integrons – indicating it has the potential to transfer between bacteria[143]. Combining next-generation sequencing methods is therefore a powerful tool for understanding resistomes.

In addition to using genomics to discover or monitor ARGs, it has been used to assess the impact of antibiotics on communities of bacteria. For example, 16S rRNA gene sequencing of gut bacteria following antibiotic treatment in humans has shown that diverse gut taxa are impacted by antibiotic therapy; however, the extent varies between individuals[178]. Moreover, 16S rRNA gene-based studies have revealed the long-lasting impact of antibiotics on the gut microbiota, where often the gut microbiome does not fully return to its pre-antibiotic treatment state[179]. However, 16S rRNA gene sequencing only allows for species- or genus-level resolution, meaning detailed analysis of species or strains is not possible. Whole genome shotgun metagenomics adds more resolution, and has been used to identify that the initial state of the gut microbiome determines the impact antibiotics will have[180]. However, both 16S rRNA gene sequencing and whole genome shotgun metagenomics depend on reference 16S rRNA gene or genome sequences[181]. Therefore, the effect of antibiotics on bacteria for which reference genomes are not available cannot be readily detected. As previously discussed,

there are uncharacterised organisms without reference sequences, meaning there is still much to be learned about the development of antibiotic resistance following antibiotic therapy in humans.

1.9 Thesis aims

Clearly antibiotic resistance is a global issue and whilst the gut has been described as an antibiotic resistance reservoir, the full capacity for antibiotic resistance and which antibiotics are effective against gut bacteria are not well defined. This thesis therefore sets out to characterise the antibiotic resistance potential of human gut bacteria. To do this, I will exploit recent developments in culturing of gut bacteria and characterise the genotypic and phenotypic resistance profiles of intestinal microbiota. Moreover, I will seek direct experimental evidence of the selection of antibiotic resistance within communities of commensal human microbiota to help understand the dynamics of antibiotic resistance in the gut microbiota. The thesis can be broken down into three key parts:

- **Characterisation of genomic antibiotic resistance in commensal gut bacteria:** Determine a comprehensive overview of antibiotic resistance genes and mutations in commensal human gut bacteria representing the phylogenetic diversity of the human gut microbiome.
- **Determination of phenotypic antibiotic resistance in commensal gut bacteria and the accuracy of genotypes:** Measure antibiotic sensitivity phenotypes in commensal human gut bacteria representing the phylogenetic diversity of the human gut microbiome and compare this to antibiotic genotypes.
- **Modelling the development of antibiotic resistance *in vivo*:** Assess the impact of amoxicillin therapy on amoxicillin resistance in mice with human-derived gut microbiota at both community- and strain-level.

Chapter 2: Materials and Methods

All analysis was performed by me, unless otherwise stated.

2.1 Anaerobic gut bacteria culture collection

The human commensal gut microbiota used in this thesis belong to the Human Gastrointestinal Bacteria Culture Collection (HBC)[151] that I helped to develop (published in 2019). The HBC contains 737 gut bacteria isolated from 20 humans, 8 from the UK and 12 from North America[151]. Any 16S rRNA gene sequences of less than 98.7 % similarity to known sequences are considered to belong to novel species, < 94.5 % to novel genera, < 86.5 % to novel families, < 82.0 % novel orders, < 78.5 % novel classes and 75.0 % novel phyla[182].

Genomic DNA was extracted from the 737 HBC isolates using a phenol:chloroform procedure[146,151] by Dr Hilary Browne. DNA was sequenced by the Sanger Sequencing Pipelines (Wellcome Sanger Institute, WSI) team using the Illumina HiSeq platform, generating paired-end reads of 125 or 150 bp; these reads were assembled using the pipeline described by Page *et al.* 2016[146,151,183]. The amino acid sequences of forty core genes were extracted from each genome using FetchMG[184], concatenated and aligned using MAFFT v7.0[185,186]. A phylogeny was inferred using FastTree v2.1.3 SSE3[187,188] and the JTT+CAT model of amino acid evolution.

2.2 Genome-based predictions of antibiotic resistance in the HBC

Raw paired-end sequencing reads for each genome were used as input for the Antibiotic Resistance Identification By Assembly (ARIBA) algorithm[160], which performs local assemblies and maps them against a database of antibiotic resistance genes and mutations (genetic

resistance determinants). In this study, the Comprehensive Antibiotic Resistance Database[163,164] (CARD; version 2.0.2 June 2018) was used with default ARIBA parameters: 90% nucleotide identity for clustering sequences (ARIBA prepareref task), plus minimum 90 % alignment identity, minimum 20 % alignment length, minimum 50 reads assembly coverage and 95 % of gene cluster sequence must be assembled to call a gene cluster present (ARIBA run task). The observed determinants were grouped by the class of antibiotics they are reported to confer resistance to, using the CARD ontology. If a determinant was described as conferring resistance to more than one antibiotic, it was classified as 'nonspecific' antibiotic resistance. The exception is for resistances to Macrolide, Lincosamide, Pleuromutilin and Streptogramin (MLPS) antibiotics, as resistance determinants against these antibiotics can have cross-resistance to each other and are grouped together in a single, separate category. The grouping of these determinants was visualised using Krona[189]. The proportion of identified genetic antibiotic resistance determinants that belonged to a particular antibiotic class were calculated and visualized using Krona[189]. The proportion of isolates with at least one resistance determinant was calculated. The interquartile range (determined using the Tukey method[190]) and mean number of predicted antibiotic resistance determinants in an individual isolate was plotted according to the bacterial phyla of each isolate. This was repeated for the number of antibiotic classes isolates were predicted to be resistant to.

In addition, the presence of predicted resistance to a particular antibiotic class was visualised against a core genome phylogeny (generated as described in section 2.1) using the online interactive Tree of Life (iTOL) tool[191]. To identify which antibiotic resistances were enriched in certain phyla, the proportion of isolates with at least one resistance determinant in each phyla was compared to the proportion of the overall HBC with Fisher exact tests[192] and corrected using two-stage linear step-up procedures of Benjamini, Krieger and Yekutieli[193]. These

analyses were repeated in bacterial families containing more than five HBC isolates and in novel versus characterised isolates, plus the proportion of predicted resistances were determined for each antibiotic in novel versus characterised genomes.

ARIBA was also implemented with the MegaRes[167], ResFinder[170] and SRST2-ARGANNOT[161] databases, plus the CARD-RGI[164] tool was applied, with default parameters to predict antibiotic resistance in the HBC genomes.

2.3 Genome-based predictions of antibiotic resistance in pathogenic genomes

PATRIC[194] was searched for genomes for the ESKAPE pathogens (*E. faecium*, *S. aureus*, *K. pneumoniae*, *A. baumannii*, *P. aeruginosa* and *Enterobacter* species). *E. coli* and *C. difficile* were also included as both are important causes of gut-related bacterial infections, and there are commensal isolates of these species in the HBC. I filtered for bacteria isolated from humans in clinical settings to ensure they were definitely pathogenic isolates. In addition, I chose bacteria isolated from 2010 or after, and from Canada, the US or the UK, to be consistent with the HBC isolates. From this list, I selected isolates that were considered to have 'good' quality genomes and with a high level of completeness (CheckM[195] completeness score equal to or greater than 98 %). PATRIC considers genomes good or poor quality based on summary annotation statistics and from comparisons with other PATRIC genomes after they have been through the PATRIC comprehensive genome analysis service[194]. There were no *E. faecium* isolates with post-2010 dates in the official isolation date meta-data column, but looking at other meta-data columns I identified 37 isolates from 2012 or later, two from 2001, one from 2000 and one from 1997; these were all included in the subsequent analyses. From the resultant list of genomes, I downloaded paired-end short read sequences and

predicted the presence of genetic antibiotic resistance determinants in CARD using ARIBA as in section 2.2. I then repeated the analyses of the proportion of isolates with predicted resistance, the interquartile range and mean number of resistance determinants per isolate, and the proportion of predicted resistances per antibiotic as in section 2.3 for each of the eight pathogenic species and between the pathogenic isolates and HBC isolates.

2.4 Phenotypic antibiotic sensitivity in commensal gut bacteria

HBC isolates were streaked out from glycerol stocks onto modified YCFA[146,151,196] agar (Table 2.1) plates in anaerobic conditions in a Whitley DG250 workstation at 37°C and left for 48 hours. Single colonies were sub-cultured and left to grow for another 48 hours; this was repeated once more, then a single colony for each isolate was used to inoculate 1ml of YCFA broth in a 96 well plate. Three separate inoculations were performed per isolate to allow three biological replicates to be tested. After 48 hours in broth, a cotton swab was dipped in each culture and streaked on a YCFA agar plate three times, turning 60° each time. An Oxoid antibiotic disk dispenser was used to place single-concentration antibiotic disks onto the inoculated agar plates. Mr Mark Stares assisted with these phenotypic tests.

Table 2.1 Modified YCFA media. Volumes are to make 500 ml of YCFA broth or agar. The solutions and mixes are prepared separately and added in the required volume when the media is being prepared. d.H_2O = distilled water.

Ingredient (part number)	Amount		Components of solutions and mixes (part number)	Amount
Before autoclaving			**Resazurin Solution**	
Agar (optional)	8 g		Resazurin (10269990)	0.1 g
Tryptone (84610.0500)	5.0 g		d.H_2O	100 ml
Yeast extract (LP021B)	1.25 g			
$NaHCO_3$ (10583381)	2.0 g		**Mineral Solution I**	
(D)+Glucose (G8270)	1.0 g		K_2HPO_4 (10677623)	3 g
(D)+Maltose (CHE1900)	1.0 g		d.H2O	1 L
(D)+Cellobiose (10207603)	1.0 g			
L-cysteine (30089)	0.5 g		**Mineral Solution II**	
Mineral Solution I	75 ml		KH_2PO^4 (P9791)	3 g
Mineral Solution II	75 ml		$(NH_4)2SO_4$ (A4418)	6 g
Resazurin Solution	0.5 ml		NaCl (10616082)	6 g
Haemin Solution	5 ml		$MgSO_4$ (CHE2458)	0.6 g
Vitamin Solution I	0.5 ml		$CaCl_2$ (dry) (10704492)	0.6 g
d.H_2O	Up to 500 ml		d.H2O	1 L
VFA mix	3.1 ml			
NaOH (CHSO0041)	pH to 7.45		**VFA mix**	
			Acetic acid (10304980)	17 ml
After autoclaving			Propionic acid (15658000)	6 ml
Vitamin Solution II	0.5 ml		n-Valeric acid (10686584)	1 ml
			Iso-valeric acid (129542)	1 ml
			Isobutyric acid (11366766)	1 ml
			Haemin Solution	
			KOH (8143530100)	0.28 g
			Ethanol 95 %	25 ml
			Haemin (10506591)	0.1 g
			d.H_2O	Up to 100 ml
			Vitamin Solution I	
			Biotin (B4501)	5 mg
			Cobalamin (Vitamin B12) (47869)	5 mg
			PABA (4-Aminobenzoic Acid) (A9878)	15 mg
			Folic acid (F7876)	25 mg
			Pyridoxine (P5669)	75 mg
			d.H_2O	Up to 500ml
			Vitamin Solution II	
			Thiamine hydrochloride (T4625)	25 mg
			Riboflavin (R4500)	25 mg
			d.H_2O	Up to 500 ml

Disk concentrations were selected based on advice for *Enterococcus* from CLSI, EUCAST and BSAC as a Gram-positive, facultative anaerobic Firmicute (Table 2.1). Though Bacteroidetes are Gram-negative, the same concentration disks were used for consistency. Zone of inhibition diameters were measured using a digital caliper after 48 hours and averaged across the biological replicates. The identity of each culture was confirmed using full-length PCR of the 16S rRNA gene (7F forward primer (5'-AGAGTTTGATYMTGGCTCAG-3') and 1510R reverse primer (5'-ACGGYTACCTTGTTACGACTT-3')) with the following program: 95 °C 15 mins; 35 cycles of 95 °C 30 s, 58 °C 30 s, 72 °C 2 mins; 72 °C 8 mins. DNA purification and capillary sequencing was performed by Eurofins Genomics (Germany). The forward and reverse sequences were trimmed and those shorter than 400bp discarded. The remaining forward and reverse reads were merged using Merger (Emboss: 6.3.1[197]) and BLASTn was used to identify a closest taxonomic match against the 16S rRNA gene sequences of the HBC[151].

Table 2.2. Single-concentration antibiotic disks used for phenotypic susceptibility testing. Antibiotics were chosen based on clinical relevance and presence of genetic determinants of antibiotic resistance in the HBC genomes.

Antibiotic (part number)	Antibiotic class	Oxoid Disk Concentration (µg)	Based on guidelines
Amoxicillin (11952962)	Beta-lactams: penicillin	10	*Enterococcus*; CLSI 2015
Ceftriaxone (11963812)	Beta-lactams: cephalosporin	10	*Enterococcus*; EUCAST 2018
Ciprofloxacin (11499838)	Fluoroquinolones	10	*Enterococcus*; CLSI 2015
Erythromycin (10280243)	Macrolides	15	*Enterococcus*; CLSI 2015
Gentamicin (10299772)	Aminoglycosides	30	*Enterococcus*; EUCAST 2018
Metronidazole (11913972)	Nitroimidazoles	5	Miscellaneous; BSAC 2015
Tetracycline (11963872)	Tetracyclines	30	*Enterococcus*; CLSI 2015
Trimethoprim (10597083)	Dihydrofolate reductase inhibitor	5	*Enterococcus*; EUCAST 2018
Vancomycin (11974012)	Glycopeptide	30	*Enterococcus*; CLSI 2015

Density curves of average zone of inhibition diameters for each phylum and antibiotic was plotted using R[198] and ggplot2[199]. In addition, the interquartile range of average zone sizes in isolates with and without genetic determinants of resistance were determined using the Tukey method[190].

2.5 Defining a scale for categorising resistant/susceptible phenotypes

Guidelines for defining resistance and susceptibility exist for a range of pathogens, though are less well-characterised in anaerobes and Gram-positive bacteria. Moreover, they are used for categorising closely related taxa, usually at the species level. In this thesis I wanted to compare at higher taxonomic levels, therefore I used the phenotypic data that I generated to create a scale, considering isolates 'resistant' if the zone size was in the lower quartile or smallest 25 % of all zone sizes for a particular antibiotic. Similarly, isolates were considered 'susceptible' if the zone size was in the upper quartile or largest 25 % of all zone sizes for a particular antibiotic. Isolates with zone sizes in the middle 50 % require further testing to determine whether or not they should be considered susceptible or resistant (see Fig. 4.4, Table 4.1 and Table 4.3 in Chapter 4). In combination with the genotypic data, there are four key genotype/phenotype combinations of confirmed or unpredicted susceptibility or resistance. These genotype/phenotype combinations were then converted to a four-number system and visualised as a heatmap in iTOL[191] against the phylogeny of the 73 isolates (generated as described in section 2.1). In addition, the proportions of these categories were determined for both the antibiotics tested and in each phylum. The enrichment of each combination in antibiotics or phyla was determined using Fisher exact tests[192] and corrected using the Benjamini, Hochberg, and Yekutieli method[193] (q significant < 0.05). The proportion and

enrichment of Unpredicted Resistances per phylum was determined for each antibiotic in the same way. Proportions of genotype/phenotype combinations were compared between databases or prediction methods as described in section 2.2. to the proportions first calculated using ARIBA[160] with CARD[163,164] with two-proportion z tests plus Yates correction (q value significant < 0.05).

2.6 Further investigations of Unpredicted Resistance

All phenotyped HBC isolates were ranked from most ceftriaxone-resistant to least ceftriaxone-resistant (i.e., from smallest to largest mean ceftriaxone zone of inhibition). Each average ceftriaxone zone of inhibition size was then plotted next to each isolate in this ranked order. Isolates in the top five most ceftriaxone-resistant isolates without beta-lactam resistance genes or mutations and with unexpectedly large average ceftriaxone zones of diameter (determined as described in section 2.4 using the Tukey method[190]) were investigated further (*Bacteroides faecis* 18048_2#66 and *Lachnospiraceae nov.* 20287_6#18). The HBC core gene phylogeny visualised in iTOL[191] from section 2.1 was used to identify the most closely related isolates in the HBC to *Bacteroides faecis* 18048_2#66 and *Lachnospiraceae nov.* 20287_6#18. Average Nucleotide Identity (ANI) between the isolates of interest and close relatives was performed using FastANI[200] with default parameters. The Minimum Inhibitory Concentration (MIC) for ceftriaxone was determined using Biomerieux Etest strips. Isolates were purified, grown in YCFA broth culture and used to inoculate YCFA agar plates as described for disk testing in section 2.4. Etest strips (ceftriaxone gradient 0.016-256 µg/mL; Biomerieux part number 506618) were placed on the plates and MIC measured after 48 hours growth in anaerobic conditions as for disk tests described in section 2.4. The most closely related isolate

to *Lachnospiraceae nov.* 20287_6#18 that was susceptible to ceftriaxone was only 81 % similar by ANI and so these isolates were excluded from subsequent analyses.

Roary[201] was used to identify genes unique to the resistant *B. faecis* isolates (MIC > 256 µg/mL) and genes unique to the susceptible *B. faecis* isolates (MIC < 48 µg/mL) as well as genes shared by both sets. The NCBI Protein database was searched for "beta-lactamase" and the amino acid sequence of all results was downloaded. ShortBRED[194] was used to group these sequences into unique amino acid markers. ShortBRED was also used to search the genes unique to the resistant isolates for the amino acid markers of beta-lactamase related sequences. A range of similarity cut offs from 90 % to 25 % were tested. The highest cut off, 90 %, represents a high level of similarity that can be used to infer function. A single beta-lactamase was identified in the *B. faecis* isolate ("Group 2384") with 90 % identity. The presence of this gene in the 737 HBC isolates was predicted using ARIBA[160] as in section 2.2 and was found in 16 additional HBC isolates. Five of those isolates also contained other beta-lactam resistance genes or mutations and so were excluded from downstream analyses. The MIC of ceftriaxone for the remaining 11 additional Group 2384-positive isolates and their most closely related Group 2384-negative isolates in the HBC was determined using Etests as before. ANI analysis was also performed for these isolates as above. The nucleotide sequence for the Group 2384 gene was extracted from each phenotyped Group 2384-positive isolate, aligned using Muscle[202,203] (visualised in SeaView[204]) and a phylogenetic tree inferred using the General Time Reversible model with FastTree[187,188].

The Group 2384 candidate beta-lactamase was synthesised in a plasmid vector containing a chloramphenicol resistance gene using GeneArt (ThermoFisherScientific, plasmid pACYC184[205], construct ID 18ADVNOP). The construct was transformed into electrocompetent ElectroMAX DH10B T1 Phage-Resistant Competent *E. coli* Cells

(ThermoFisherScientific, part number 12033015, via electroporation according to the manufacturer's instructions and with assistance from Mr Matthew Dorman). The ceftriaxone MIC of the recipient *E. coli* strain was determined using Etests as earlier described. The transformed cells were grown on LB agar plates containing chloramphenicol (12.5 µg/mL) and ceftriaxone at a concentration of 256 µg/mL (representing the observed *B. faecis* phenotype) or 4 µg/mL (slightly above *the E. coli*'s initial MIC).

Bacteroides faecis 18048_2#66 and *Lachnospiraceae nov.* 20287_6#18 isolates with unpredicted ceftriaxone resistance were purified and Mr Mark Stares isolated genomic DNA using phenol:chloroform extractions. The genomic DNA was separated by size using a 1 % low-melting point agarose gel and fragments of approximately 40kb extracted using the Copy Control Fosmid Cloning Kit (Lucigen part number CCFOS059) according to manufacturer's instructions. These fragments were cloned into ceftriaxone susceptible *E. coli* using the Copy Control Fosmid Cloning Kit (Lucigen part number CCFOS059) according to manufacturer's instructions. The vector contains a chloramphenicol resistance gene; clones were selected for on LB agar containing chloramphenicol (12.5 µg/mL). The ceftriaxone MIC of the recipient *E. coli* strain was determined using Etests as earlier described. The transformed *E. coli* were screened for gain of ceftriaxone resistance by growth on LB agar containing ceftriaxone at a concentration of 256 µg/mL (representing the observed *B. faecis* phenotype) and 4 µg/mL (slightly above the *E. coli*'s initial MIC).

2.7 Humanised microbiota mouse experiments

A mouse line with microbiota derived from a healthy human had been established previously by Dr Simon Clare, Dr Sam Forster, Dr B. Anne Neville and colleagues in the Wellcome Sanger Institute Research Support Facility, by oral gavage of homogenised stool into germ-free (GF) mice ("Donor 2 humanised microbiota mice"). A second mouse line with microbiota derived from a different healthy human had also previously been established using the same techniques ("Donor 7 humanised microbiota mice") by the same persons. Full details are described in section 5.2.1.

Mice were given a theoretically therapeutic dose (approximately 45 mg/kg/day), based on the concentration required to adequately exceed the MIC of sensitive organisms in otitis media infections[206] and assuming that the average mouse weighs 30 g and consumes approximately 5 ml of water per day (according to John Hopkins University, http://web.jhu.edu/animalcare/procedures/mouse.html, accessed June 2015 and June 2019, and as advised by Dr Simon Clare). Amoxicillin sodium (TOKU-E part number A059) was dissolved in water, sterilized using a 0.2 µm filter and given to the mice via drinking water for seven days by Dr Simon Clare and his team. Faecal pellets were collected from each mouse at various time points before and after treatment by Dr Simon Clare and his team. At each time point, I weighed each individual faecal pellet and homogenised them in 100 mg/ml in sterile PBS; faecal homogenates were pooled per cage. An aliquot of the undiluted pooled homogenate was treated with ethanol for 30 minutes (1:3 volumes of 70 % v/v ethanol) to select for ethanol-resistant and spore-forming organisms[146], and washed by centrifugation for 13200 g for five minutes at room temperature, before removal of the supernatant and resuspension in four volumes PBS. The wash was repeated twice more and after the third wash and removal of supernatant, the sample was resuspended in the original volume of PBS.

Untreated and ethanol-treated pooled homogenates were serially diluted 1 in 10 from 10^{-1} to 10^{-7}. Untreated dilutions were plated on modified complex, broad-range YCFA[188] (Table 2.1), with or without amoxicillin added, under aerobic conditions (37 °C in a New Brunswick Scientific Innova 42 incubator) and anaerobic conditions (37 °C in a Whitley DG250 workstation. Amoxicillin was included in the agar at a concentration representative of clinical resistance according to EUCAST and CLSI guidelines for anaerobic bacteria (8 mg/L). Ethanol-treated dilutions were plated as above except that the agar plates also contained the bile salt sodium taurocholate (STC; Fisher Scientific UK Ltd part number 10629452) to promote germination of spores[146].

2.8 Colony count data

At each time point, the number of colonies growing in each condition from each cage was counted and converted to colony forming units (CFU) per gram of stool. CFU/g values were averaged across both experiments for each culture condition: aerobic plates without amoxicillin, anaerobic vegetative plates without amoxicillin, anaerobic spore-forming plates without amoxicillin, aerobic plates with amoxicillin, anaerobic vegetative plates with amoxicillin and anaerobic spore-forming plates with amoxicillin. Standard deviation was determined for each of these six conditions. The colony count data was then tested for normal or log-normal distribution using the Anderson-Darling, D'Agostino and Pearson, Shapiro-Wilk, and Kolmogorov-Smirnov tests. As many of the conditions did not pass the normality tests, a non-normal distribution was assumed for performing Kruskal-Wallis (One-way ANOVA of non-parametric data) tests to determine statistically significant changes in bacterial load over the time course of the mouse experiments (comparisons were performed between time points).

The two-stage step-up method of Benjamini, Krieger and Yekutieli[193] was used to correct for multiple comparisons.

2.9 Isolation of individual isolates and analysis

Non-confluent colonies were picked from culture plates, after 24-hour and 48-hours of growth in aerobic and anaerobic conditions respectively, directly into 300 µl YCFA broth. This was done with the assistance of Dr Hilary Browne, Dr Sam Forster, Dr Elisa Viciani and Dr Ana Zhu alongside other members of the Host-Microbiota Interactions Laboratory. The aerobic colonies were left to grow in broth for another 24 hours; anaerobic colonies were left for 48 hours. Half a millilitre of each broth culture was mixed with 0.5ml of 50 % glycerol to create glycerol stocks and stored at -80 °C. The full-length 16S rRNA gene was amplified from each isolate using PCR and sequenced by Eurofins as described in section 2.4. Reads were aligned using ssu-align v0.1.1[207] and the STK alignment converted to MSA. Consensus OTU sequences were aligned with Mafft[185,186] and a phylogeny inferred using FastTree[187,188] for visualisation with ITOL[191]. The tree leaves were annotated based on the BLASTn results and Operational Taxonomic Units (OTUs) were assigned with Mothur v1.35.1[208] with a defined cutoff of 0.01 (i.e., 99 % similarity). The number of isolates in each OTU was visualised using Krona[189]. Isolates of interest were purified and genomic DNA extracted using phenol:chloroform by Mr Mark Stares. Sequencing was performed by the Sanger Sequencing Pipelines (WSI) team using Illumina X10 technology, generating 150 bp paired-end reads.

Paired-end sequence reads were filtered and assembled using the pipeline as earlier mentioned[183]. To perform phylogenetic analyses, the amino acid sequences of 40 core genes were extracted, aligned and visualised as for the HBC[184-188,191] in section 2.1. Each genome was

searched for the presence of known antibiotic resistance determinants described in CARD[163,164] using ARIBA[160] with default parameters, as for the HBC in section 2.2. The ARIBA results were compared between genomes from isolates cultured on agar with amoxicillin after the mice received amoxicillin and genomes of the same strain (>99 % 16S identity) from agar without amoxicillin before treatment. The sequence of the *cfxA* beta-lactamase was identified in isolates of one OTU, *Odoribacter splanchnicus* 99.556 %, after amoxicillin therapy but was not present in isolates of the same OTU from before therapy. The nucleotide sequences of *cfxA* were extracted from all the whole genomes generated in this study (where present) and aligned using Muscle[202,203]. The amoxicillin MIC was determined for four isolates of *Odoribacter splanchnicus* 99.556 %, as described in section 2.6 (amoxicillin gradient 0.016-256 µg/mL; Biomerieux part number 500918).

For each candidate OTU, Roary[201] was used to identify genes found in the isolates from culture plates containing amoxicillin after mice were treated with amoxicillin but absent from isolates in the same candidate OTU from culture plates without amoxicillin before the mice received treatment. Genes meeting this criteria were searched for sequences with 90 % similarity to the ShortBRED[209] beta-lactamase markers developed in section 2.6. In addition, paired-end sequencing reads from bacteria isolated after amoxicillin treatment were mapped against assembled contigs from an isolate of the same OTU cultured before amoxicillin treatment using Smalt (https://www.sanger.ac.uk/science/tools/smalt-0). Variant bcf files were produced and used to identify SNPs in coding sequences (open reading frames annotated as such if over 100 bases) in Artemis[210]. Coding sequences containing SNPs were compared to known protein sequences using BLASTx.

2.10 Metascrape and metagenomic analysis

The total growth on a culture plate was mixed with 1 ml sterile PBS and scraped off ("metascrape") for total DNA extraction using the FastDNA Spin Kit for Soil (MP Biomedicals part number 1149200) according to manufacturer's instructions. Metagenomic sequencing was performed by the Sanger Sequencing Pipelines (WSI) team using Illumina HiSeq 4000 technology, generating paired-end reads of 150 bp. Metagenomic samples were multiplexed at 32 samples per lane. Metagenomic data was analysed using Kraken[211] with default parameters to compare the sequence reads to databases of reference genomes and assign taxonomy. Raw read counts per species were used to determine alpha and beta diversity using R scripts developed by Dr Kevin Vervier at each experimental time point for three culture conditions (aerobic, anaerobic vegetative and anaerobic spore-forming), on agar plates with or without 8 mg/L amoxicillin added. Statistical significance was determined by Mann-Whitney U tests[212], adjusted for multiple-testing with the Benjamini, Hochberg, and Yekutieli method[193]; q value significant < 0.05. Read counts were normalised per sample and the normalised values were averaged across samples for each of the three conditions (aerobic, anaerobic vegetative and anaerobic spore-forming) from plates containing amoxicillin. The species were then ranked by their mean relative abundance in each culture condition and the mean relative abundance of the top 10 most abundant species plotted as stacked bar charts, with the relative abundances of the remaining species grouped as "Other". Relative abundance of species was compared between consecutive time points and the start and end of the experiment using two-proportion z tests with Yates correction (q value significant when $q < 0.05$). Finally, the whole genome sequences generated in this study were searched for in the metascrape data to confirm the presence or absence of candidate OTUs of increased amoxicillin resistance on plates containing amoxicillin using Mash[213] with default parameters.

An identity cut off of 0.99 (99 %) was used to conclude that a particular strain was present in a metascrape sample.

Chapter 3: Characterisation of genomic antibiotic resistance in commensal gut bacteria

3.1 Introduction

The gut microbiome is considered a reservoir of antibiotic resistance[214] through the sharing of antibiotic resistance genes (ARGs) amongst autochthonous, commensal gut bacteria. In addition, these ARGs may be shared with allochthonous, transient bacteria passing through the gastrointestinal tract, which can include bacterial pathogens[123]. This is evidenced by studies that have identified both known, clinically relevant antibiotic resistance genes in samples collected from the gut microbiome as well as 'novel' antibiotic resistance genes (not previously identified or seen in pathogenic bacteria). For example, the functional metagenomics study by Sommer, Dantas and Church in 2009[139] identified ARGs identical at the nucleotide level to genes previously identified in clinical isolates of disease-causing bacteria, proving that the gut microbiota do harbour clinically relevant ARGs. However, the majority of antibiotic resistance-conferring inserts (82 %) were distantly related to known antibiotic resistance genes from pathogens (nucleotide identity 60.7 % on average)[139]. These results highlight both the diversity of genetic antibiotic resistance determinants in the gut but also their potential to be shared between commensal and pathogenic bacteria.

Another interesting study of the gut as a reservoir of antibiotic resistance used whole genome shotgun metagenomic sequencing to examine the resistomes of 180 healthy individuals from 11 different countries across Europe, Asia, North and South America representing industrialised, low-income and remote societies[215]. In total, 507 different ARGs were identified, including eight shared by all 180 individuals[215]. This further highlights how antibiotic resistance appears to be widely distributed across gut microbiomes. If those ARGs

are acquired by pathogenic bacteria passing through the gut, or if the gut bacterium hosting the ARG moves outside of its usual location in the gut and causes an infection elsewhere in the body, then gut microbiota are part of the antibiotic resistance problem.

These types of metagenomic studies investigate the entire community but not individual bacteria, therefore it is essentially impossible to say exactly which bacterium a particular antibiotic resistance determinant might have come from. Antibiotic resistance genes are often located on mobile genetic elements; since mobile genetic elements can be shared between different bacteria, it is hard to place the nucleotide sequence of a mobile genetic element identified from a mixed sample into its original genomic context. It can also be difficult to validate the results, as without knowing the host or having an isolate of the suspected host a phenotype cannot be measured and correlated with the presence or absence of the ARG in question. Therefore, these studies are good for observing the presence and abundance of known or putative ARGs in an environment, but not determining exactly which bacteria they belong to. This means that we do not necessarily fully understand the taxonomic placement of antibiotic resistance determinants among the bacterial community in a mixed sample, such as a stool sample that represents the gut microbiome. Combined with the fact that we are still discovering new members of the gut microbiota through metagenome-assembled genomes and high-throughput culturing[146,151,216] the full potential of the gut microbiome as an antibiotic resistance reservoir remains to be understood.

With recent developments in culturing of anaerobic commensal gut bacteria[146,151], there is now the opportunity to use these methods to determine a comprehensive map of antibiotic resistance determinants in these diverse and relatively uncharacterised microorganisms. It is important to determine which commensal gut bacteria harbour antibiotic resistance genes as this information will help understand which ones are of concern for acting as donors to spread

antibiotic resistance. In this chapter, I use a unique collection of reference genomes to study the presence of known antibiotic resistance determinants in diverse commensal human gut bacteria.

3.2 Results

3.2.1 Summary of resources used in this chapter

To determine the presence of antibiotic-resistant determinants in diverse commensal human gut bacteria, I have used the genomes from the Human Gastrointestinal Bacteria Culture Collection (HBC)[151]. Each isolate has been whole genome sequenced using Illumina short-read paired-end sequencing, assembled and annotated[183], and the physical isolates are held in glycerol stocks at -80 °C allowing for phenotypic validation and characterisation (Chapter 4). The HBC contains 737 gut bacteria isolated from healthy adult humans, who had not taken antibiotics in the six months prior to sampling, using broad range culturing and targeted phenotype culturing for spore-forming bacteria[146,151] (Fig. 3.1). The collection contains 273 species in total, 105 of which are considered novel (Table 3.1).

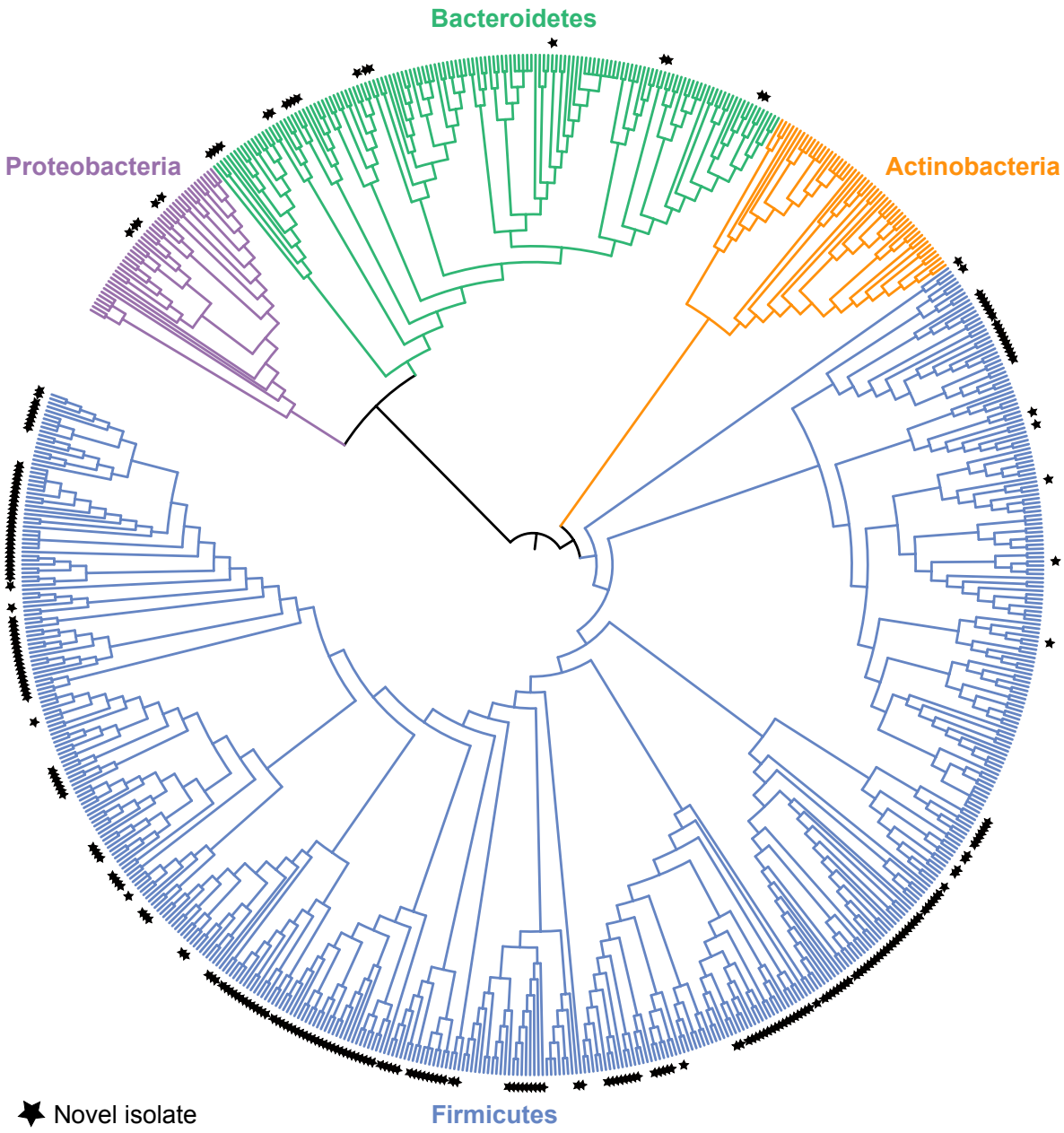


Figure 3.1 Phylogeny of the HBC commensal gut bacteria. The Human Gut Bacteria Culture Collection contains 737 isolates of human gut bacteria[151]. The amino acid sequences of 40 core genes were extracted from these genomes and used to infer a phylogeny to illustrate the taxonomic diversity of this culture collection. The stars mark which genomes are considered novel based on the similarity of their 16S rRNA gene sequence to known 16S rRNA gene sequences in RefSeq.

Table 3.1. Taxonomic information for the HBC. The number of total or novel isolates, species and families is summarized for the four phyla of the HBC and the HBC overall.

	Actinobacteria	Bacteroidetes	Firmicutes	Proteobacteria	All HBC
Total isolates	53	143	496	45	**737**
Novel isolates	0	18	253	5	**276**
Total species	16	40	203	14	**273**
Novel species	0	13	91	1	**105**
Total families	6	4	18	3	**31**
Novel families	0	1	0	0	**1**

Species were defined by considering the sequence identity of each genome's full length 16S rRNA gene. 16S rRNA gene sequences at least 97.8 % identical to each other are considered the same species[182]. In addition, the HBC 16S rRNA gene sequences were compared to RefSeq 16S rRNA sequences to assign taxonomic classification. Any 16S rRNA gene sequences of less than 94.5 % similarity to known sequences are considered to belong to novel genera, < 86.5 % to novel families, < 82.0 % novel orders, < 78.5 % novel classes and 75.0 % novel phyla[182]. In total, there are 276 isolates that belong to novel taxonomic groups in the HBC. This unique collection offers extensive and novel phylogenetic diversity, compared to the six ESKAPE pathogenic species (*E. faecium, S. aureus*, *K. pneumoniae*, *A. baumannii*, *P. aeruginosa* and *Enterobacter* spp.) that represent just six species from two orders within the Firmicutes and Proteobacteria. ESKAPE pathogens are included later as a comparator as they have been the focus of most antibiotic resistance gene and mutation research.

Sequence reads for each genome were searched against the Comprehensive Antibiotic Resistance Database (CARD)[163,164] as it is one of the largest resistance determinant databases, contains both ARGs and resistance-associated mutations, and is updated regularly. It is also commonly used as a reference database for studying antibiotic resistance genes and

mutations (e.g.[143,215,217,218]). Version 2.0.2 of CARD (from June 2018, analysis performed in July 2018), contains 2426 antibiotic resistance determinants. The CARD ontology describes which antibiotics each determinant confers resistance to; these descriptions were further grouped by the major antibiotic class – for example all sub-classes of beta-lactams were combined (Fig. 3.2; see Appendix 1 for Table A1.1 describing groupings). In total, there are 29 different antibiotic classes or categories represented in CARD. Determinants conferring resistance to Macrolide, Lincosamide, Pleuromutilin and Streptogramin antibiotics were grouped together (MLPS) as resistance determinants against these antibiotics can have cross-resistance each other[219] and are grouped together in a single, separate category. Any other determinants described as conferring resistance to more than one class of antibiotics were grouped under "non-specific antibiotic resistance" for the purpose of this study.

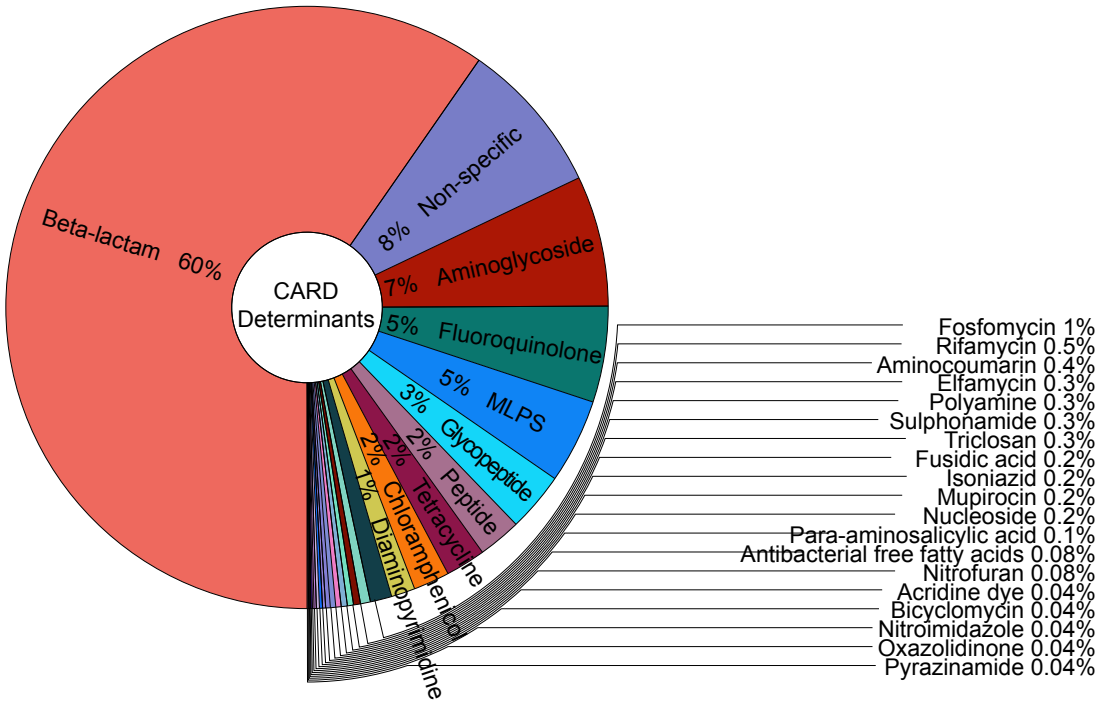


Figure 3.2. The proportions of antibiotic categories in the Comprehensive Antibiotic Resistance Database. 2426 antibiotic resistance determinants are described in CARD. The CARD ontology describes which antibiotics each determinant confers resistance to. If determinants were described as conferring resistance to multiple classes of antibiotics, these were grouped under "non-specific" antibiotic resistance. The exception is for resistances to Macrolide, Lincosamide, Pleuromutilin and Streptogramin (MLPS) antibiotics, as resistance determinants against these antibiotics can have cross-resistance to each other and are grouped together in a single, separate category.

3.2.2 Computational predictions of antibiotic resistance in 737 whole genome sequences of anaerobic gut bacteria

Approximately 60 % of the determinants in CARD are associated with beta-lactam resistance (Fig. 3.2). The CARD database includes redundancy, therefore I used ARIBA[160] (Antibiotic Resistance Identification By Assembly) to predict the presence of CARD determinants in the HBC genomes. ARIBA clusters gene sequences from CARD by similarity, so that only the presence of "unique" non-redundant antibiotic resistance determinants are considered. Following the ARIBA 'prepareref' command, which performs the clustering of determinants, 1024 clusters were created. ARIBA then performs local assemblies of sequence reads, in this case for each individual genome, against the reference sequence for each antibiotic resistance determinant cluster. If an assembled gene was 90 % similar to a reference antibiotic resistance determinant cluster sequence at the nucleotide level, that cluster was reported as present in that genome.

In total, 178 unique clusters of genetic antibiotic resistance determinants were observed in the HBC (Fig. 3.3, see Appendix 2, Table A2.1 for full description and https://docs.google.com/spreadsheets/d/1zwmhUicOW3JVW_9y6P6LssbavW47EFiRq4_wwnS9CMg/edit?usp=sharing for Table A2.2), with a combined total of 1434 occurrences across the 737 genomes. The determinants were grouped as described before: 19 of the 29 possible categories of antibiotics were represented by the HBC genomic resistances. The largest proportion of identified resistance determinants (29 %) were those considered "non-specific" determinants (Fig. 3.3). Beta-lactam antibiotic resistance determinants were the second most common type (13 %) to be observed. All antibiotic classes from the WHO List of Essential Medicines[21] are represented in these observations, with the exception of oxazolidinone; it is

possible oxazolidinone resistance may be present via non-specific determinants if they alter the 23S and 30S rRNA subunits targeted by this antibiotic.

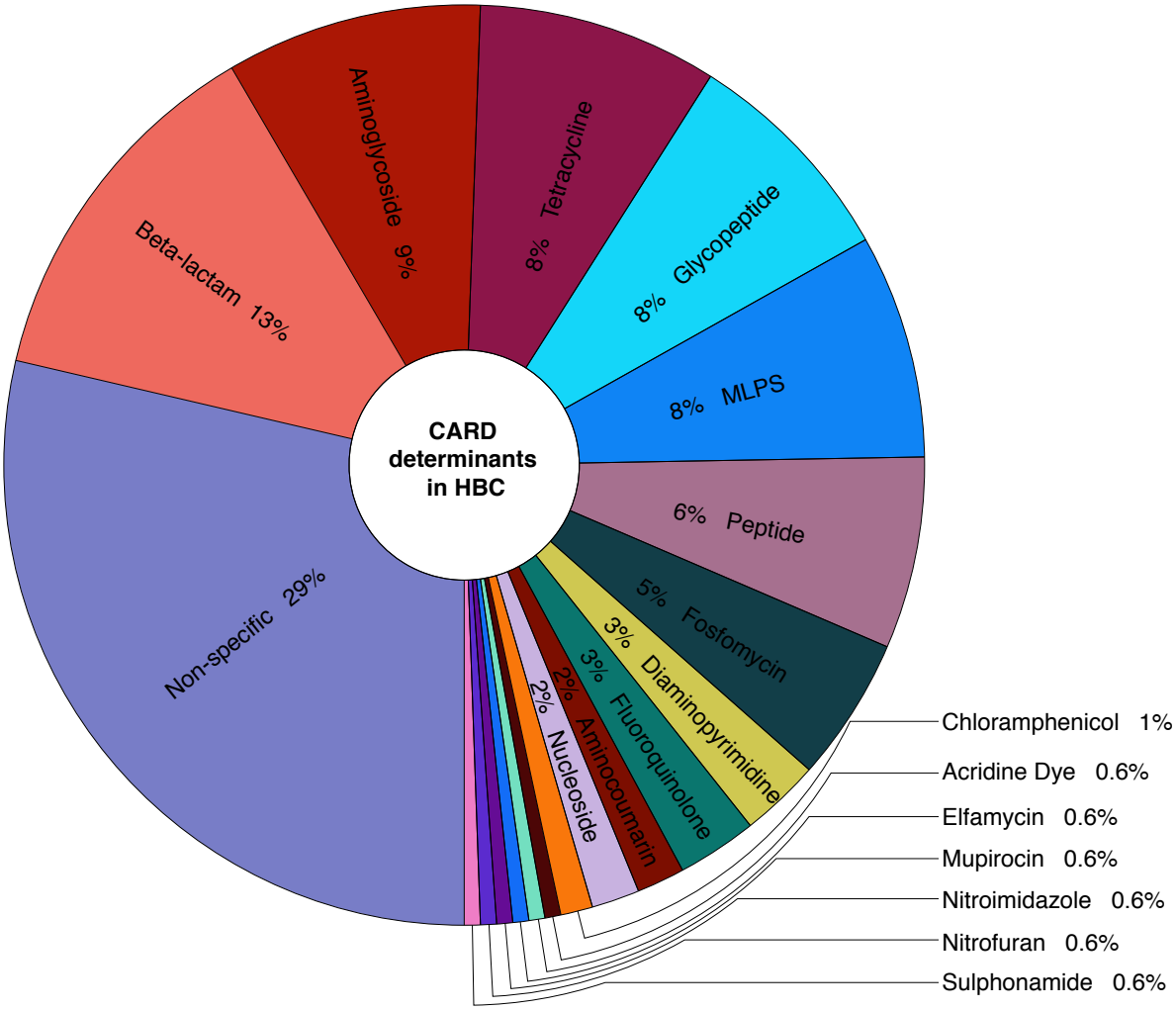


Figure 3.3. Proportions of 178 antibiotic resistance determinants identified in 737 isolates of human gut bacteria. The determinants were grouped by the antibiotic class that they are described as conferring resistance to according to the CARD ontology. MLPS = Macrolide, Lincosamide, Pleuromutillin and Streptogramin A.

Approximately half of the HBC genomes (390/737; 52.9 %) were predicted to contain at least one antibiotic resistance determinant (Fig. 3.4A). The number of unique determinants in a single genome ranged from 1 to 56 (including non-specific resistance determinants; Fig. 3.4B), although approximately three-quarters of the genomes were observed to contain just one or two antibiotic resistance determinants. Individual isolates carried resistances to 10 different classes of antibiotics (excluding the non-specific antibiotic resistance category; Fig. 3.4C),

though again approximately three quarters of the HBC only harboured resistance to two or fewer antibiotic categories. Overall, the majority of this phylogenetically diverse collection of human commensal gut microbiota contain antibiotic resistance determinants, despite not having been exposed to antibiotics for at least six months. In addition, the range in number of resistance determinants, and the number of antibiotic classes resistance is predicted to, per isolate indicates variability in antibiotic resistance genotypes across the HBC.

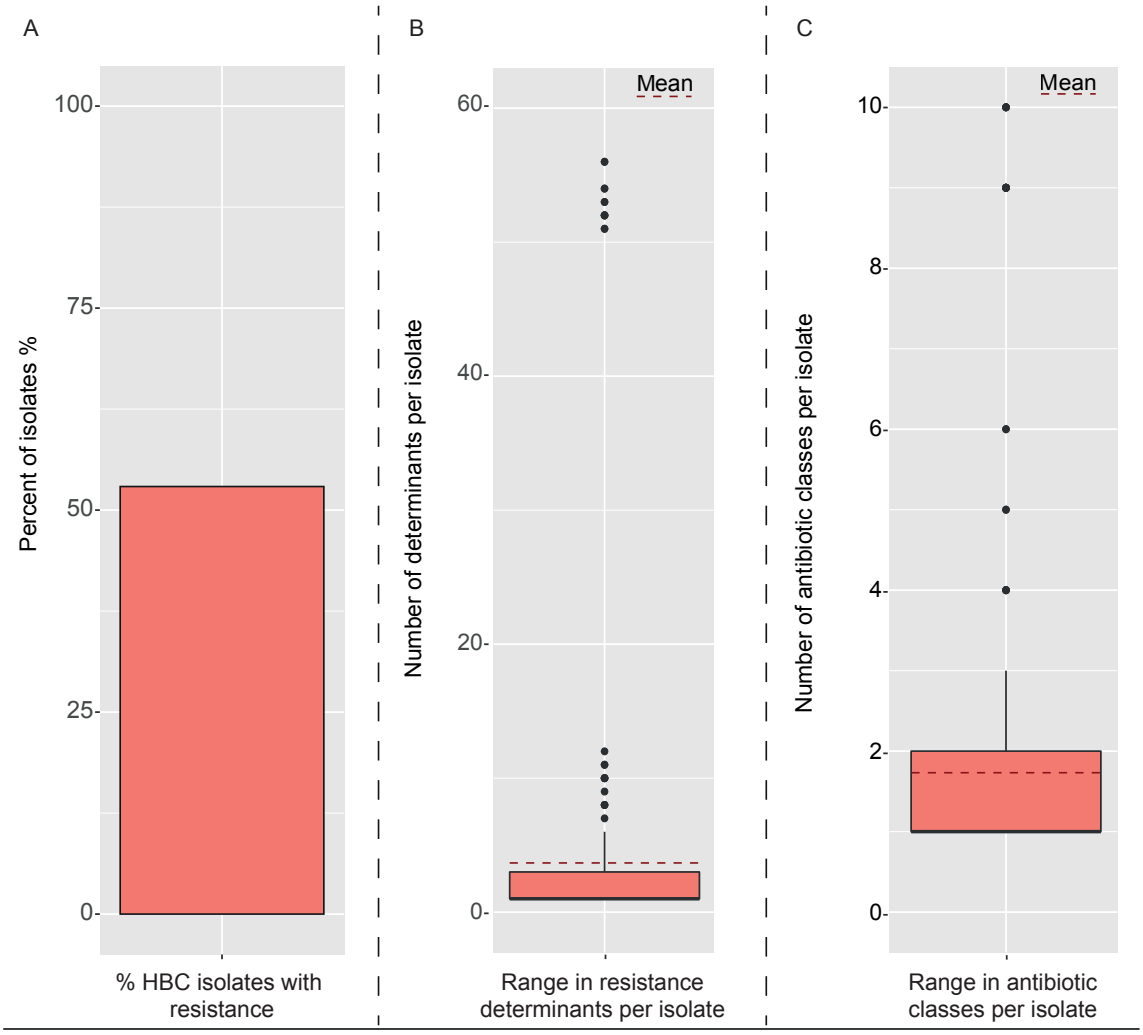


Figure 3.4. Summary of genetic determinants of antibiotic resistance in the HBC.
A) The proportion of the 737 genomes with genetic determinants of resistance identified using CARD and ARIBA.
B) The range in the number of genetic determinants of resistance predicted in an individual genome.
C) The range in different antibiotic classes an individual genome was predicted to be resistant against (excluding non-specific antibiotic resistance).
For B) and C), the boxes show the interquartile range determined using the Tukey method; the black circles indicate outlier observations and the thick black line represents the median. The red dashed line represents the mean.

3.2.3 Variation of predicted genomic resistance across the four key gut bacteria phyla

The presence of a resistance determinant to a particular antibiotic is assumed to confer phenotypic resistance to that antibiotic (e.g., the presence of a beta-lactam resistance gene predicts that isolate to be resistant to beta-lactam antibiotics). Therefore in this thesis these observations are considered "predicted resistance". The distribution of predicted resistances belonging to the 19 categories of antibiotic among the 737 genomes is demonstrated in Figure 3.5. This shows that antibiotic resistance is predicted throughout all four phyla in the HBC, and that predicted resistances vary, even between closely related isolates. However, Proteobacteria, and certain members of the Firmicutes, have more predicted antibiotic resistances than other isolates. These isolates belong to species known to be able to act as opportunistic pathogens such as *Enterobacter cloacae*[220], *Klebsiella oxytoca*[221], *K. pneumoniae*[221] (Proteobacteria) and *E. faecalis*[222] and *E. faecium*[222] (Firmicutes). These initial findings support the role of the gut microbiota as a reservoir of antibiotic resistance and its potential to contribute to antibiotic-resistant infections.

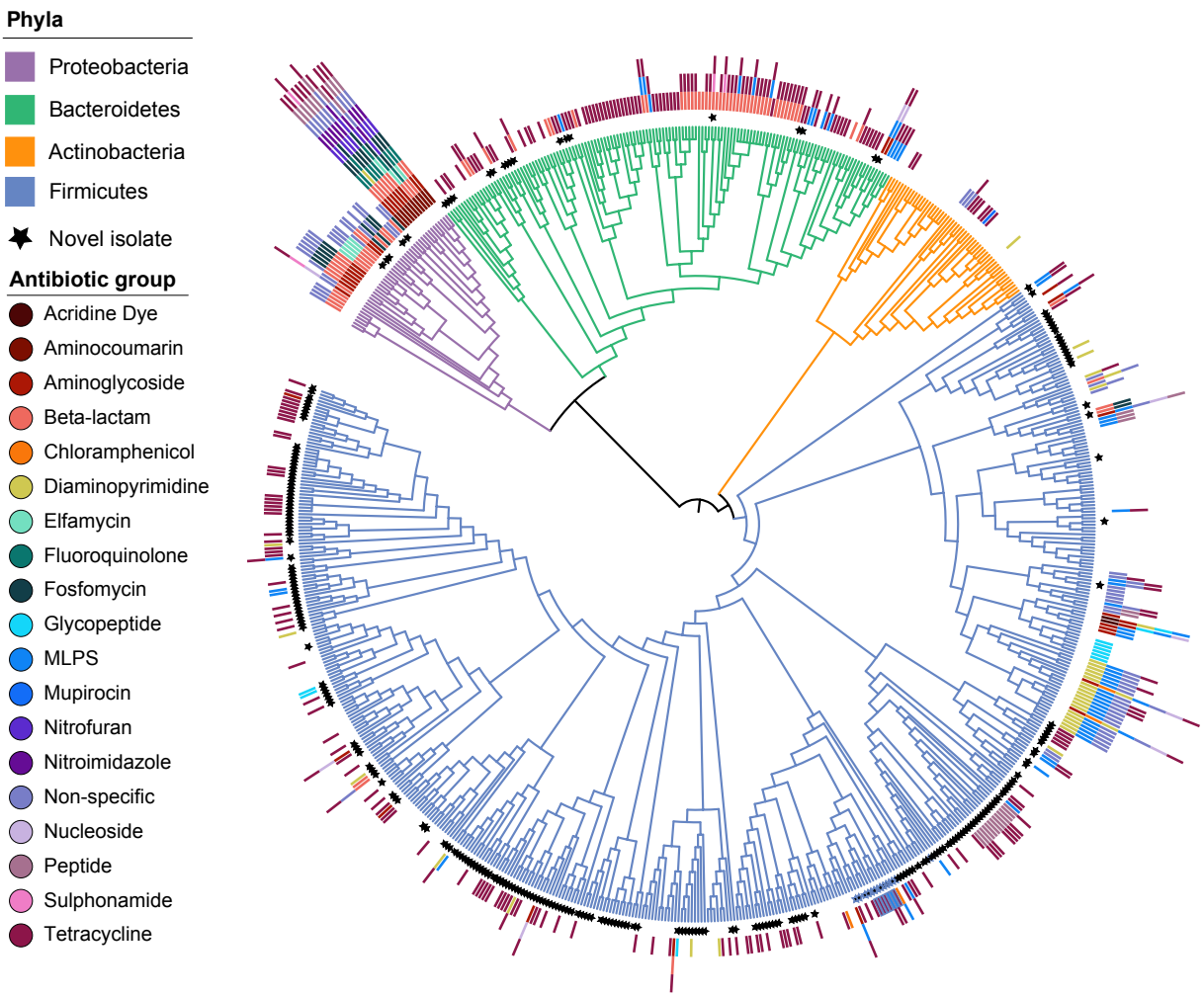


Figure 3.5. Observations of predicted antibiotic resistance in the HBC isolates against the core genome phylogeny. The phylogeny represents the core genomes of 737 whole genome sequences of gut microbiota isolated from healthy human faecal samples. Antibiotic resistance genes and mutations described in CARD were identified in these genomes, which were grouped by the corresponding class of antibiotic. If more than one, then these were classified as 'non-specific resistances' or MLPS if Macrolide, Lincosamide, Pleuromutillin or Streptogramin A. The outer rings of coloured bars show the presence of at least one resistance determinant to a particular antibiotic class. Proteobacteria isolates appear to have the highest number of resistances to different antibiotic classes.

Having determined the overall occurrence of known antibiotic resistance determinants in the HBC and gained a broad idea of their distribution, I next sought to understand the prevalence of antibiotic resistance predicted in each of the four main gut microbiota. From the phylogeny in Figure 3.5, Proteobacteria have the highest number of resistances to different antibiotic classes. Looking more specifically at the proportion of isolates within phyla (Fig. 3.6) further demonstrates this: 95.6 % of Proteobacteria are predicted to harbour resistances compared

to 82.5 % of Bacteroidetes, 43.5 % of Firmicutes and just 24.5 % of Actinobacteria. Proteobacteria and Bacteroidetes also have a significantly higher proportion of predicted resistant isolates than expected based on the HBC proportion overall (q values of < 0.0001, determined by Fisher exact tests and corrected for multiple testing; significant < 0.05).

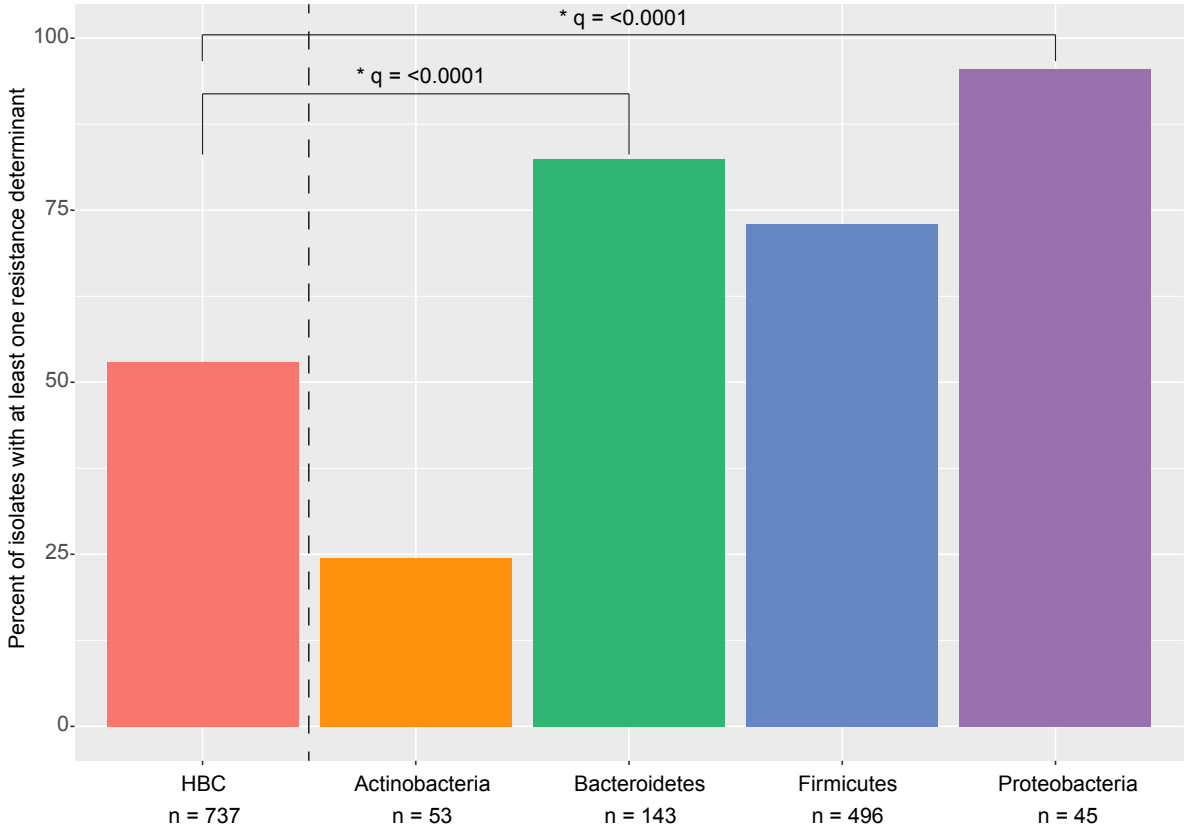


Figure 3.6. Proportions of isolates with at least one genetic antibiotic resistance determinant in each phyla, compared to the overall HBC. The numbers of genomes with at least one genetic antibiotic resistance determinant were counted for the complete HBC and for each of the four HBC phyla. HBC 52.9 %; Actinobacteria 24.5 %; Bacteroidetes 82.5 %; Firmicutes 43.5 %; Proteobacteria 95.6 %. Proteobacteria and Bacteroidetes had significantly more isolates with predicted resistance than expected based on the overall HBC collection. Statistical significance determined by Fisher exact tests, and corrected using the Benjamini, Hochberg, and Yekutieli method; q = significant < 0.05. n = total number of isolates in that group.

Moreover, Proteobacteria showed a bigger range and higher maximum number of antibiotic resistance determinants per isolate than the other three phyla (Fig. 3.7) and again, this was found to be statistically significant (q < 0.0001). Overall, the data from Figures 3.5 to 3.7 so far show that antibiotic resistance determinants are not distributed evenly between or within

phyla: more Bacteroidetes isolates are found to contain resistance determinants than expected, but the number of determinants per isolate was not significantly different to the HBC overall. The Proteobacteria phylum is significantly enriched for the number of isolates with resistance and the number of resistance determinants per isolate.

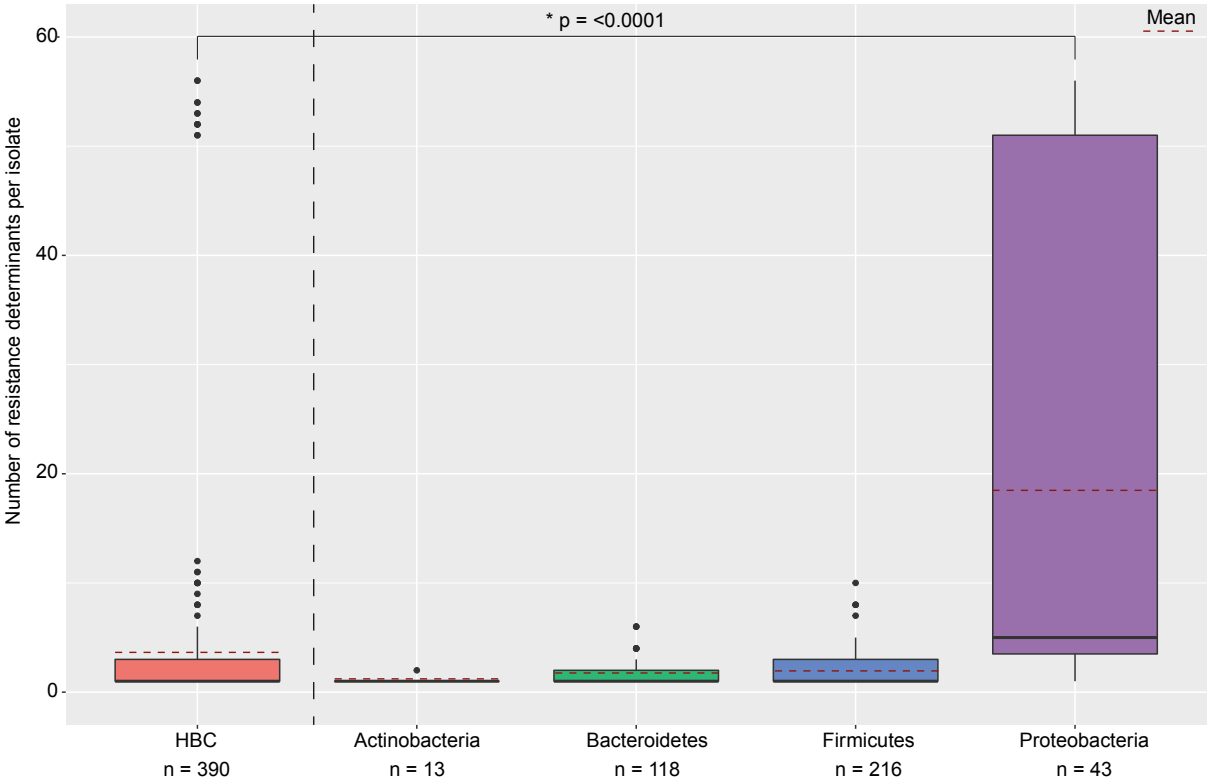


Figure 3.7. Interquartile range of number of antibiotic resistance determinants per isolate in each phyla, compared to the overall HBC. The total number of antibiotic resistance determinants was calculated for each isolate and the interquartile range in this number plotted for all 737 HBC genomes and for each of the four HBC phyla. Actinobacteria had the smallest range (0-2) and lowest total number of antibiotic resistance determinants per isolate of all the phyla. In contrast, Proteobacteria had the biggest range (0-56) and highest total number of antibiotic resistance determinants per isolate. Interquartile range was determined using the Tukey method; the black circles indicate outlier observations and the thick black line represents the median. The red dashed line represents the mean. The mean number of determinants per isolate was statistically higher in Proteobacteria than the HBC. Statistical significance determined by Mann-Whitney U test; p = significant < 0.05. n = number of isolates in that group with predicted resistance (i.e., excluding isolates without resistance).

Similarly, Proteobacteria display a bigger range in the number of specific antibiotic categories resistance is predicted to (i.e., excluding non-specific antibiotic resistances; Fig. 3.8): the Proteobacteria isolates are on average predicted to be resistant to four different antibiotics but could be resistant to up to 10 different antibiotics. Actinobacteria, Bacteroidetes and

Firmicutes are only predicted to be resistant to 1.0 or 1.5 antibiotics on average respectively. Again, these results were found to be statistically significant ($p < 0.0001$). This analysis excludes the non-specific antibiotic resistances. That the range in number of predicted resistances appears similar to the range in number of antibiotic resistance determinants per isolate suggests that non-specific antibiotic resistance is not a major factor for Actinobacteria, Bacteroidetes and Firmicutes. There is a bigger difference in the Proteobacteria, indicating that these isolates contain more non-specific resistances.

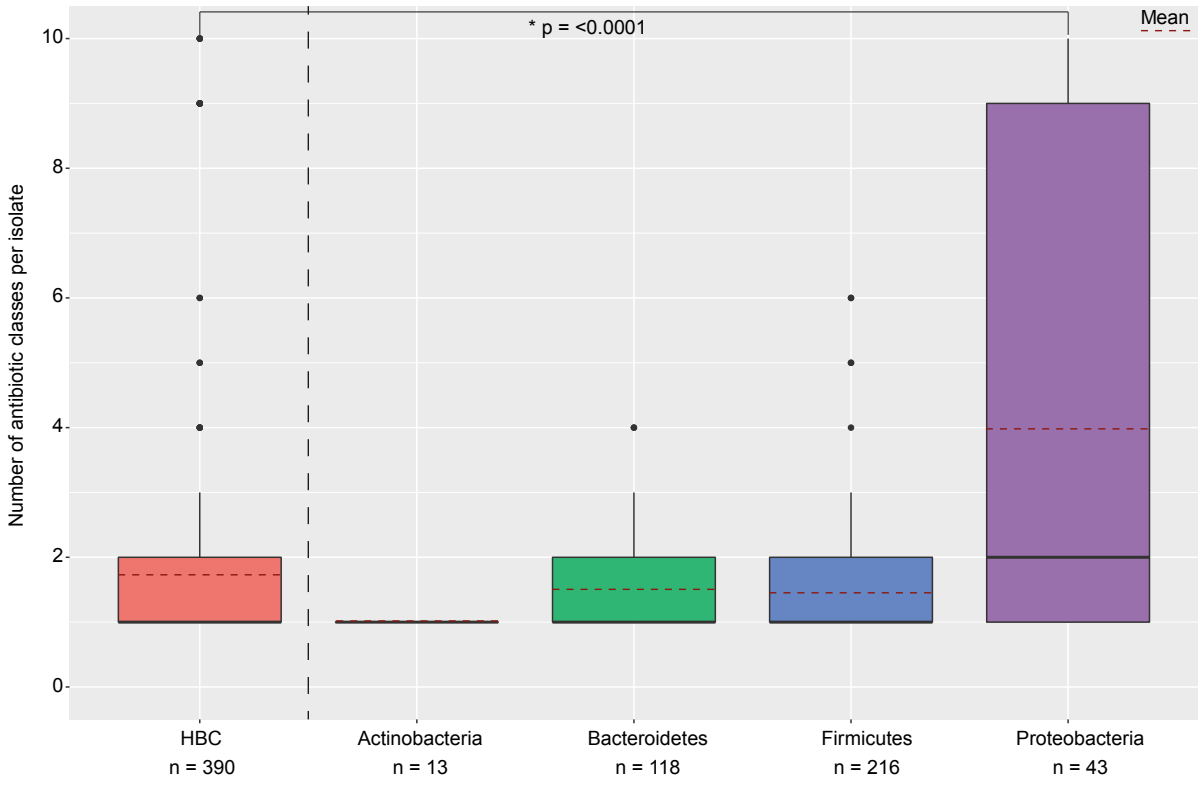


Figure 3.8. Interquartile range of number of antibiotic classes resistance is predicted to per isolate in each phylum, compared to the overall HBC. After grouping the determinants by the antibiotic class they are described as conferring resistance against by CARD, the total number of specific antibiotic classes a single isolate was predicted to harbor genetic resistances against was counted (i.e., excluding non-specific resistances), Again, Proteobacteria had the biggest range and highest maximum number of antibiotic classes per isolate. Interquartile range was determined using the Tukey method; the black circles indicate outlier observations and the thick black line represents the median. The red dashed line represents the mean. The mean number of different antibiotic classes per isolate was statistically higher in Proteobacteria than the HBC. Statistical significance determined by Mann-Whitney U test; p = significant < 0.05. n = number of isolates in that group with predicted resistance (i.e., excluding isolates without resistance).

Next, I considered whether certain phyla were enriched for particular antibiotic resistances. Fisher exact tests were performed to compare the proportion of isolates with resistance determinants for a particular antibiotic category within phyla compared to the proportion of all HBC isolates; p-values were adjusted for multiple testing using the Benjamini, Hochberg, and Yekutieli method (Table 3.2). Overall, every phylum except Actinobacteria is enriched for resistance to specific antibiotic classes. Bacteroidetes were enriched for beta-lactam resistance and tetracycline resistance (q values both < 0.001, significant < 0.05) and Firmicutes were enriched for diaminopyrimidine antibiotic resistances (q value 0.001). Proteobacteria were statistically enriched for determinants belonging to nine different classes of antibiotics, as well as non-specific antibiotic resistance determinants (q values all < 0.001). In several cases, resistance to antibiotic classes was observed significantly less often than expected – for example, beta-lactam resistance in Actinobacteria. This further indicates that genetic determinants of resistance are unevenly distributed between phyla. Overall, Proteobacteria are enriched for the most antibiotic classes and known resistance determinants, thus appearing to harbour more clinically relevant antibiotic resistance determinants than the other three key gut microbiota phyla.

Table 3.2. The proportion of isolates with resistance to particular antibiotic categories is compared between phyla and the overall HBC. Percentage represents proportion of predicted resistances for an antibiotic in the HBC or individual phyla. The individual phylum proportions were compared to the HBC proportion to determine the direction of statistical significance. The arrows denote the direction of significance; ↑ indicates that antibiotic resistance was observed more than expected and ↓ that antibiotic resistance was observed less than expected. No arrow means that any change in the proportion was not statistically significant. Q values significant < 0.05 (Fisher exact tests and adjusted using the Benjamini, Hochberg, and Yekutieli method).

		Phylum (number of genomes)											
		Actinobacteria (53)			Bacteroidetes (143)			Firmicutes (496)			Proteobacteria (45)		
Antibiotic	**HBC %**	**Phylum %**	**q-value***	**Direction**	**Phylum %**	**q-value***	**Direction**	**Phylum %**	**q-value***	**Direction**	**Phylum %**	**q-value***	**Direction**
Acridine Dye	0.3	0.0	1.000		0.0	1.000		0.4	1.000		0.0	1.000	
Aminocoumarin	1.6	0.0	1.000		0.0	1.000		0.0	< 0.001*	↓	26.7	< 0.001*	↑
Aminoglycoside	5.2	0.0	1.000		1.4	0.227		3.4	0.096		33.3	< 0.001*	↑
Beta-lactam	12.1	0.0	0.023*	↓	31.5	< 0.001*	↑	1.4	< 0.001*	↓	82.2	< 0.001*	↑
Chloramphenicol	0.5	0.0	1.000		0.0	1.000		0.8	1.000		0.0	1.000	
Diaminopyrimidine	5.2	1.9	1.000		0.0	0.007*	↓	7.3	0.001*	↑	2.2	1.000	
Elfamycin	1.1	0.0	1.000		0.0	1.000		0.0	0.002*	↓	17.8	< 0.001*	↑
Fluoroquinolone	1.6	0.0	1.000		0.0	1.000		0.0	< 0.001*	↓	26.7	< 0.001*	↑
Fosfomycin	3.5	0.0	1.000		0.0	0.066		0.4	< 0.001*	↓	53.3	< 0.001*	↑
Glycopeptide	1.5	0.0	1.000		0.0	1.000		2.2	0.227		0.0	1.000	
MLPS	9.4	0.0	0.087		11.9	1.000		14.1	1.000		6.7	1.000	
Mupirocin	0.1	1.9	0.767		0.0	1.000		0.0	1.000		0.0	1.000	
Nitrofuran	1.5	0.0	1.000		0.0	1.000		0.0	< 0.001*	↓	24.4	< 0.001*	↑
Nitroimidazole	1.6	0.0	1.000		0.0	1.000		0.0	< 0.001*	↓	26.7	< 0.001*	↑
Non-specific	11.8	5.7	1.000		0.0	< 0.001*	↓	8.7	0.004*	↓	91.1	< 0.001*	↑
Nucleoside	1.2	0.0	1.000		1.4	1.000		1.2	1.000		2.2	1.000	
Peptide	4.2	0.0	1.000		0.0	0.026*	↓	3.8	1.000		26.7	< 0.001*	↑
Sulphonamide	0.5	0.0	1.000		1.4	1.000		0.0	0.140		4.4	0.227	
Tetracycline	37.2	17.0	0.227		76.9	< 0.001*	↑	28.6	< 0.001*	↓	28.9	1.000	

3.2.4 Variation of predicted genomic resistance across different human commensal bacterial Families

Having identified significantly more antibiotic resistance overall in Proteobacteria so far, I wanted to see whether this was also true at the family level. Families with more than 5 isolates were selected to be analysed to allow a certain degree of robustness. Those 14 families were ranked by the proportion of isolates with predicted resistance (Fig. 3.9).

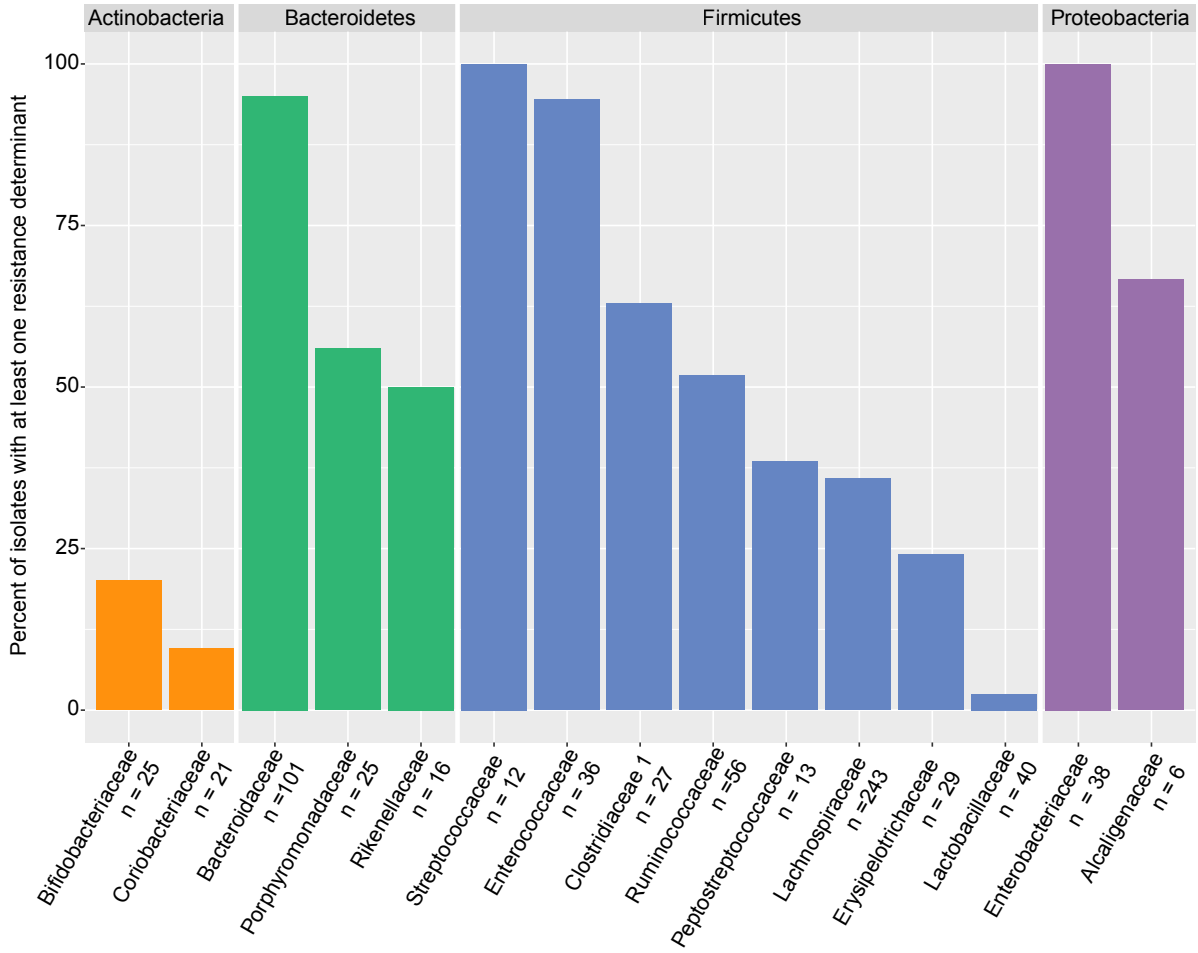


Figure 3.9. Proportions of isolates with at least one genetic antibiotic resistance determinant in taxonomic families. The numbers of genomes with at least one genetic antibiotic resistance determinant were counted for families in the HBC. The two families with the highest average number of resistance determinants per isolate or with more than five isolates are shown for each phylum.

This shows that as for phyla, families vary in their proportion of isolates with predicted antibiotic resistance. In particular, the highest rates of isolates with predicted antibiotic

resistance occur in Enterobacteriaceae, Streptococcaceae, Enterococcaceae, and Bacteroidaceae. These families are known to contain opportunistic pathogenic species. However, so is the family Peptostreptococacceae (namely, the diarrhoea-causing *Clostridioides difficile*), yet this family has a much lower proportion of isolates with predicted resistance. If we consider the number of genetic determinants per isolate between families (Fig. 3.10), even though Streptococcaceae and Bacteroidaceae have high proportions of isolates with predicted resistance, they only harbour relatively few resistance determinants per isolate. In comparison, Enterococcaceae and Enterobacteriaceae both have significantly more resistance determinants per isolate on average compared to Bacteroidaceae and Streptococcaceae. However, Enterobacteriaceae did not have significantly more resistance determinants on average than Enterococcaceae. Together so far, this data suggests that gut microbiota taxa known to contain isolates of species that can act as opportunistic pathogens are enriched for genetic determinants of antibiotic resistance.

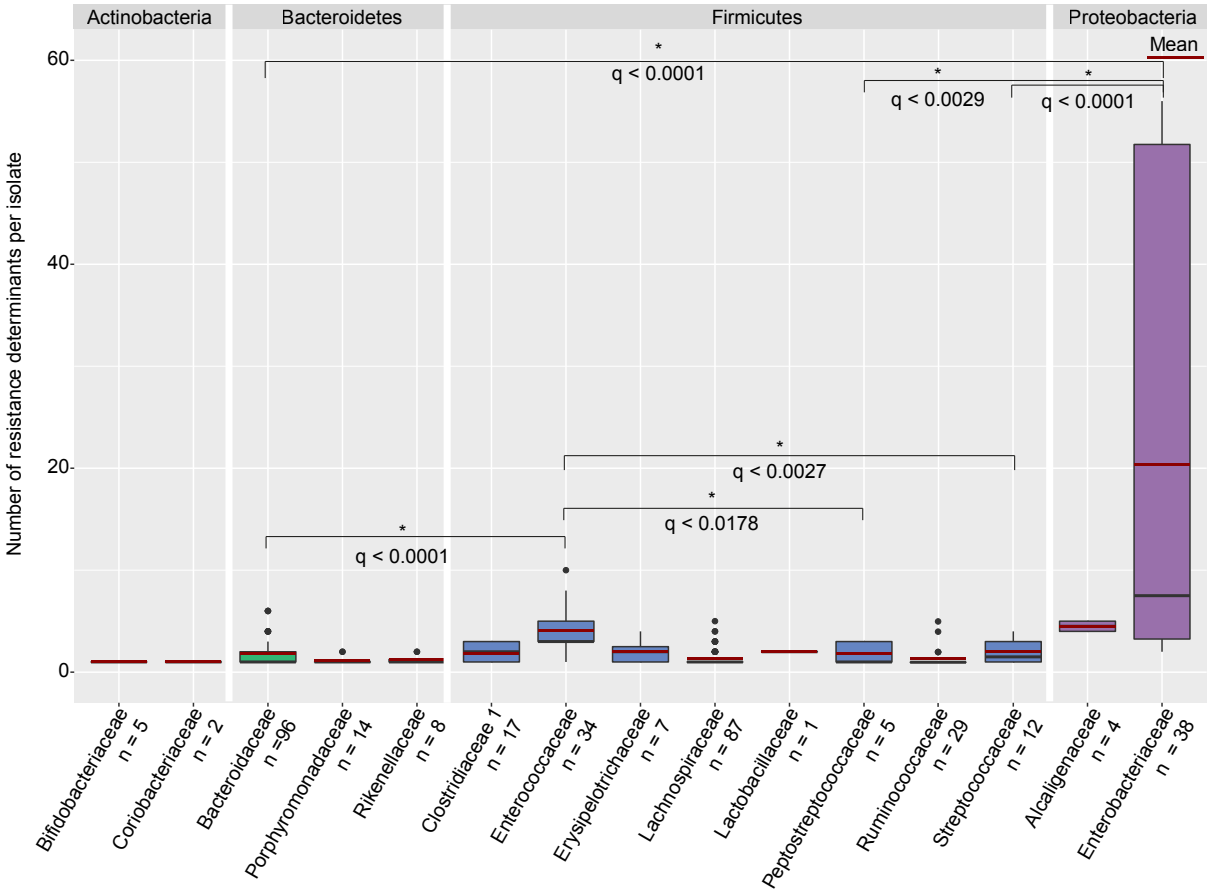


Figure 3.10. Interquartile range of number of antibiotic resistance determinants per isolate in commensal gut bacterial families. The total number of antibiotic resistance determinants was calculated for each isolate and the distribution in this number plotted at the family level. This included non-specific antibiotic resistance determinants. Interquartile range was determined using the Tukey method; the points show outlier observations and the thick black line represents the median. The red line represents the mean. n = number of isolates in that group with predicted resistance. Only families with more than five isolates were included. Statistical significance was determined between families known to contain pathogenic bacteria (Bacteroidaceae, Enterococcaceae, Peptostreptococacceae, Streptococcaceae and Enterobacteriaceae) by Kruskal-Wallis tests and corrected using the Benjamini, Hochberg, and Yekutieli method; q = significant < 0.05. Only significant results are shown.

3.2.5 Distribution of predicted genomic resistance between known and novel isolates

So far, I have identified that families containing species that can act as opportunistic pathogens are enriched for genetic determinants of antibiotic resistance. Having such a diverse and novel collection of gut bacteria has also allowed me to investigate whether clinically relevant antibiotic resistance determinants are harboured not just by bacteria belonging to previously published or described (“known”) taxa, but those considered to be novel and thus uncharacterised. As we do not have novel isolates for many of the families in

the HBC, I performed this analysis at the phylum level. Fewer novel genomes overall harbour clinically relevant antibiotic resistances (Fig. 3.11); this is also true for the Bacteroidetes phylum. There were no novel Actinobacteria in this dataset so the rates in this phylum could not be compared. In Proteobacteria and Firmicutes (which has the most novel genomes at 253), it is the novel genomes that harbour more clinically relevant antibiotic resistances. However, the only statistically significant results were that the HBC overall and Bacteroides phyla had more known isolates with resistance than novel isolates; the observations in the Firmicutes were not significantly different. Nonetheless, these results indicate that uncharacterised bacteria contribute to the intestinal reservoir of known antibiotic resistance.

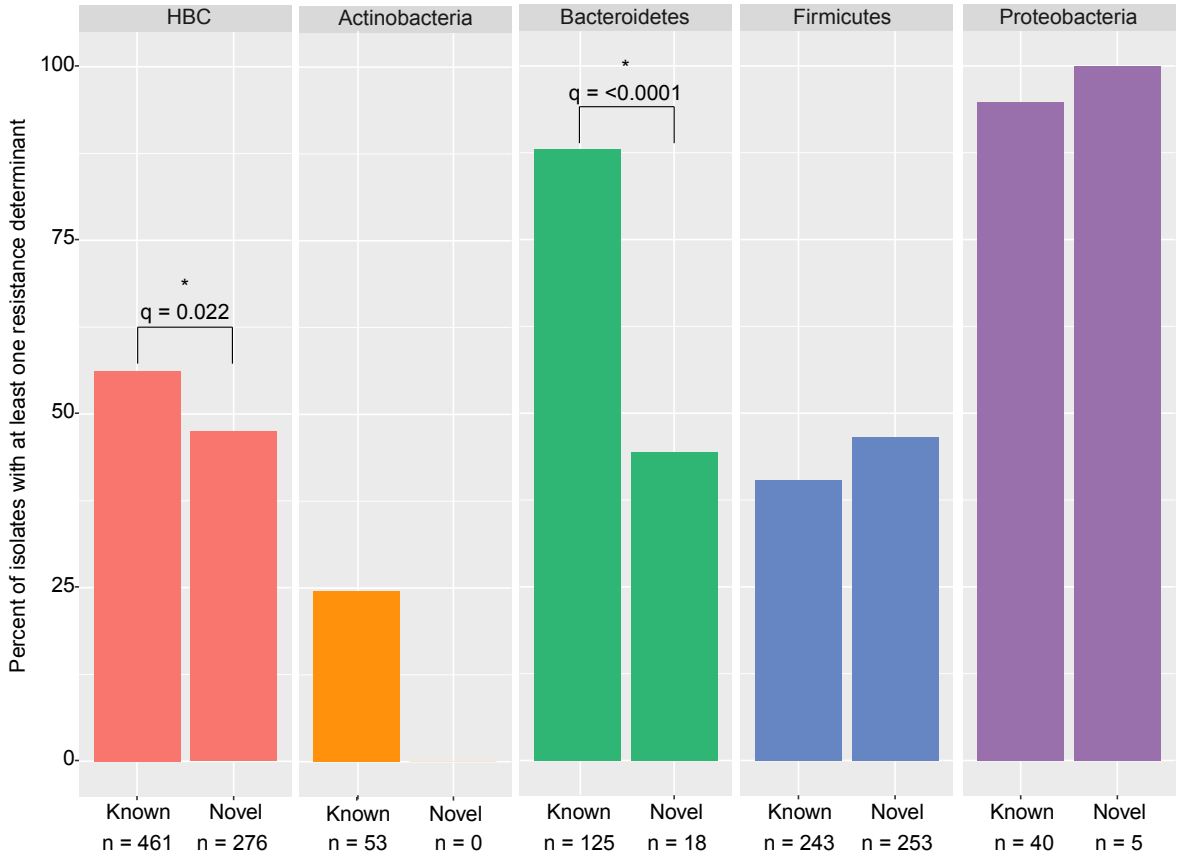


Figure 3.11. Proportions of resistant isolates in known versus novel isolates. The numbers of isolates that were novel were counted for all 737 HBC genomes and for each of the four HBC phyla. For each sub-group (known or novel), isolates with at least one genetic antibiotic resistance determinant were counted as before. 47.4 % of the novel genomes in the HBC were predicted to have at least one genetic antibiotic resistance determinant overall. There were no novel Actinobacteria genomes; the majority (91.7 %) of novel genomes belong to Firmicutes, with 46.6 % of novel Firmicutes predicted to have at least one genetic antibiotic resistance determinant overall. Overall, the HBC had significantly more known isolates with predicted resistance than novel isolates; within phyla only Bacteroides showed a significant difference. Statistical significance determined by Fisher exact tests, and corrected using the Benjamini, Hochberg, and Yekutieli method; q = significant < 0.05. n = total number of isolates in that group.

Overall, the known genomes of the HBC harboured resistances to all 19 categories of antibiotics, whereas the novel genomes only harboured resistances to 12 categories of antibiotics (Figure 3.12, showing the proportion of predicted resistances). Known and novel Bacteroidetes genomes harboured resistances to the same six antibiotics, although there was proportionately more diversity in the resistances of novel Bacteroidetes. Similarly, novel Firmicutes harboured resistances to the same 11 categories of antibiotic as known Firmicutes, though there were more tetracycline resistances in the novel Firmicutes compared to known Firmicutes. The proportion of resistances in known and novel Bacteroidetes and Firmicutes appears broadly similar, largely dominated by tetracycline resistance. In contrast, the Actinobacteria and Proteobacteria have different patterns in predicted resistances: the known Actinobacteria have resistances to all 19 classes of antibiotics and known Proteobacteria to 14 classes of antibiotics, whereas the novel Proteobacteria only have eight classes of antibiotic resistance predicted. In addition, the novel Proteobacteria have different classes of antibiotic resistance predicted (i.e., acridine dye, diaminopyrimidine, glycopeptide) than the known Proteobacteria. Together, these observations make it clear that although Proteobacteria is enriched for antibiotic resistance, non-Proteobacteria and uncharacterised gut microbiota harbour diverse antibiotic resistances, underlining the importance of understanding this extensive reservoir and its clinical relevance.

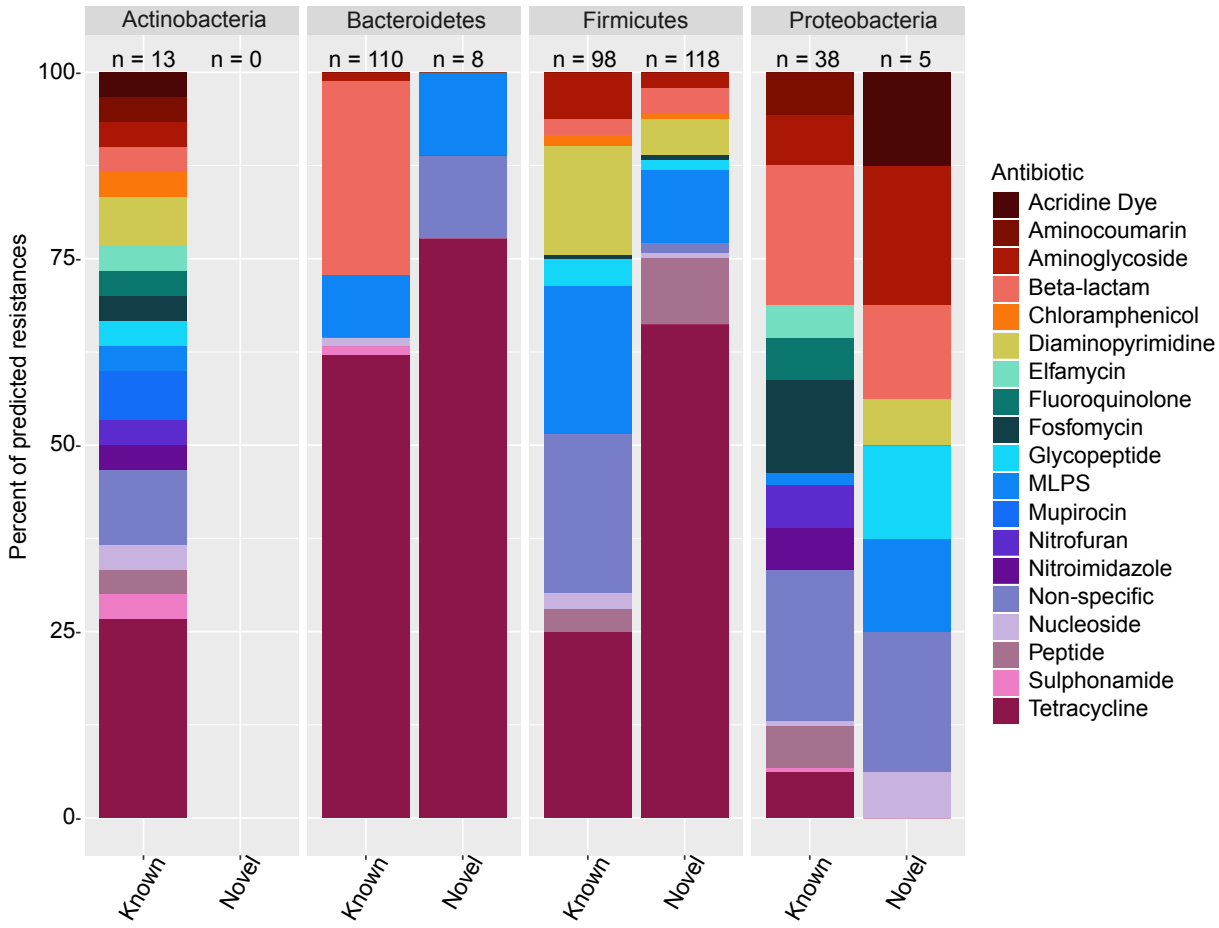


Figure 3.12. Proportions of antibiotic categories that resistances are predicted against in HBC genomes. The number of predicted resistance phenotypes in known and novel genomes per antibiotic class was calculated for the HBC and each phylum. The majority of predicted resistances in those novel Firmicutes were against tetracycline antibiotics. r = total number of predicted resistance phenotypes for that group; g = genomes with predicted resistance in that group.

3.2.6 Comparison of predicted genomic resistance in commensal versus pathogenic isolates

I next compared the presence of antibiotic resistance in commensals of the HBC to pathogens. To do this, I searched PATRIC[194] for genomes for the ESKAPE pathogens (*E. faecium*, *S. aureus*, *K. pneumoniae*, *A. baumannii*, *P. aeruginosa* and *Enterobacter species*). *E. coli* and *C. difficile* were also included as both are important causes of bacterial infections, and there are commensal isolates of these species in the HBC for comparisons. I filtered for bacteria isolated from humans in clinical settings to ensure they were definitely pathogenic isolates. In addition, I chose bacteria isolated from 2010 or after, and from Canada, the US or the UK, to be

contemporaneous with the HBC isolates. From this list, I selected isolates that were considered to have 'good' quality genomes and with a high level of completeness (CheckM[195] completeness score equal to or greater than 98). PATRIC considers genomes good or poor quality based on summary annotation statistics and from comparisons with other PATRIC genomes after they have been through the PATRIC comprehensive genome analysis service[194]. There were no *Enterococcus faecium* isolates with official isolation dates post-2010, but looking at other meta-data identified 37 isolates from 2012 or later, two from 2001, one from 2000 and one from 1997; these were all included in the subsequent analyses. For the 1725 genomes, I predicted the presence of genetic antibiotic resistance determinants in CARD using ARIBA as for the HBC.

Importantly, 97 % of these pathogenic genomes had at least one predicted resistance. However, there is more variation in the number of resistance determinants per isolate between species (Fig. 3.13) than was seen for the commensal isolates in section 3.2.4. This shows that although resistance is predicted in the majority of pathogenic isolates, different pathogenic species of bacteria have different antibiotic resistance potential. Pathogenic *E. coli* has the highest maximum, median and mean number of determinants per genome; this is similar to the observations in commensal HBC Proteobacteria.

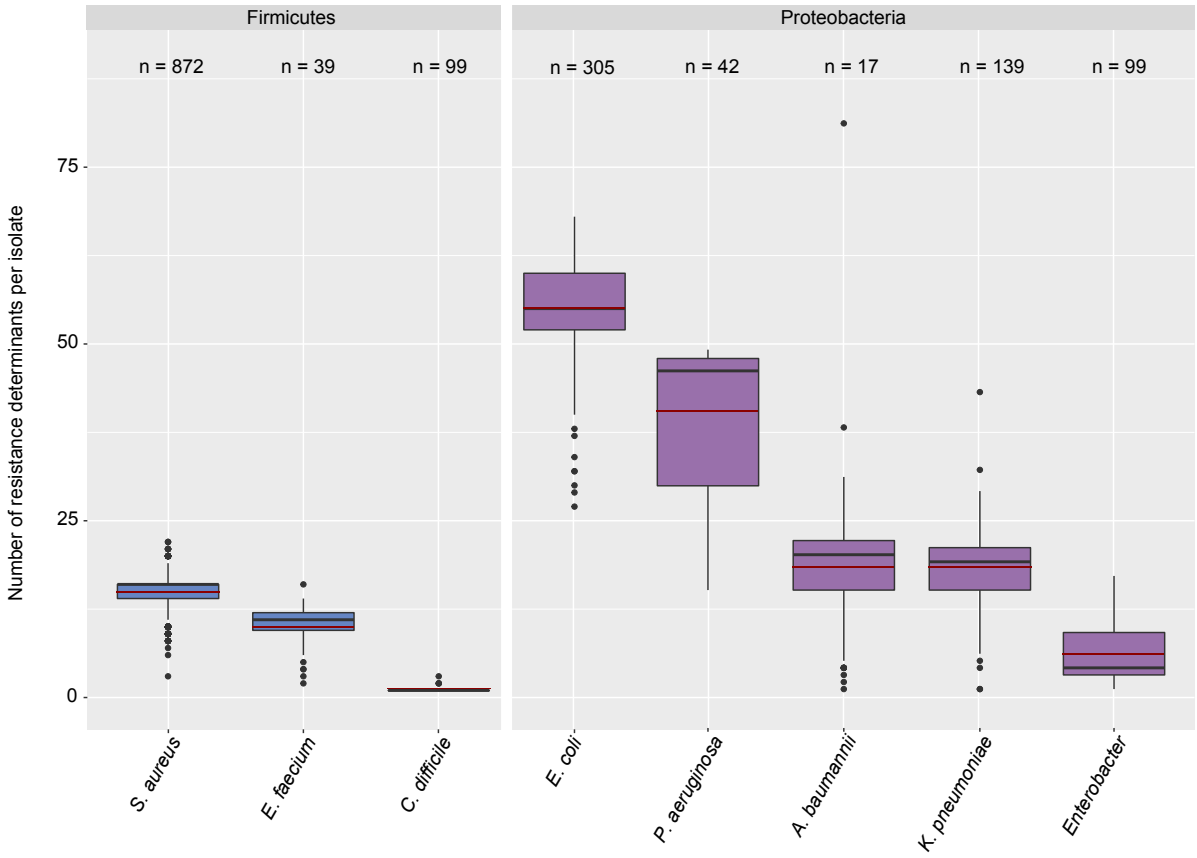


Figure 3.13. Interquartile range of number of antibiotic resistance determinants per isolate in pathogenic bacterial species. The total number of antibiotic resistance determinants was calculated for each pathogenic genome studied and the distribution in this number plotted at the species level. This included non-specific antibiotic resistance determinants. n = number of isolates in that group with predicted resistance. Interquartile range was determined using the Tukey method; the points show outlier observations and the thick black line represents the median. The red line represents the mean.

In addition, different pathogenic species have different profiles of predicted resistances (Fig. 3.14). Thus, not all of these pathogen groups are equal in their predicted propensity to harbour antibiotic resistances. The Proteobacteria species and *S. aureus* appear broadly similar, with eight or more different resistances predicted. In contrast, *C. difficile* and *E. faecium* appear less similar, with only three and seven categories of antibiotic resistances predicted respectively.

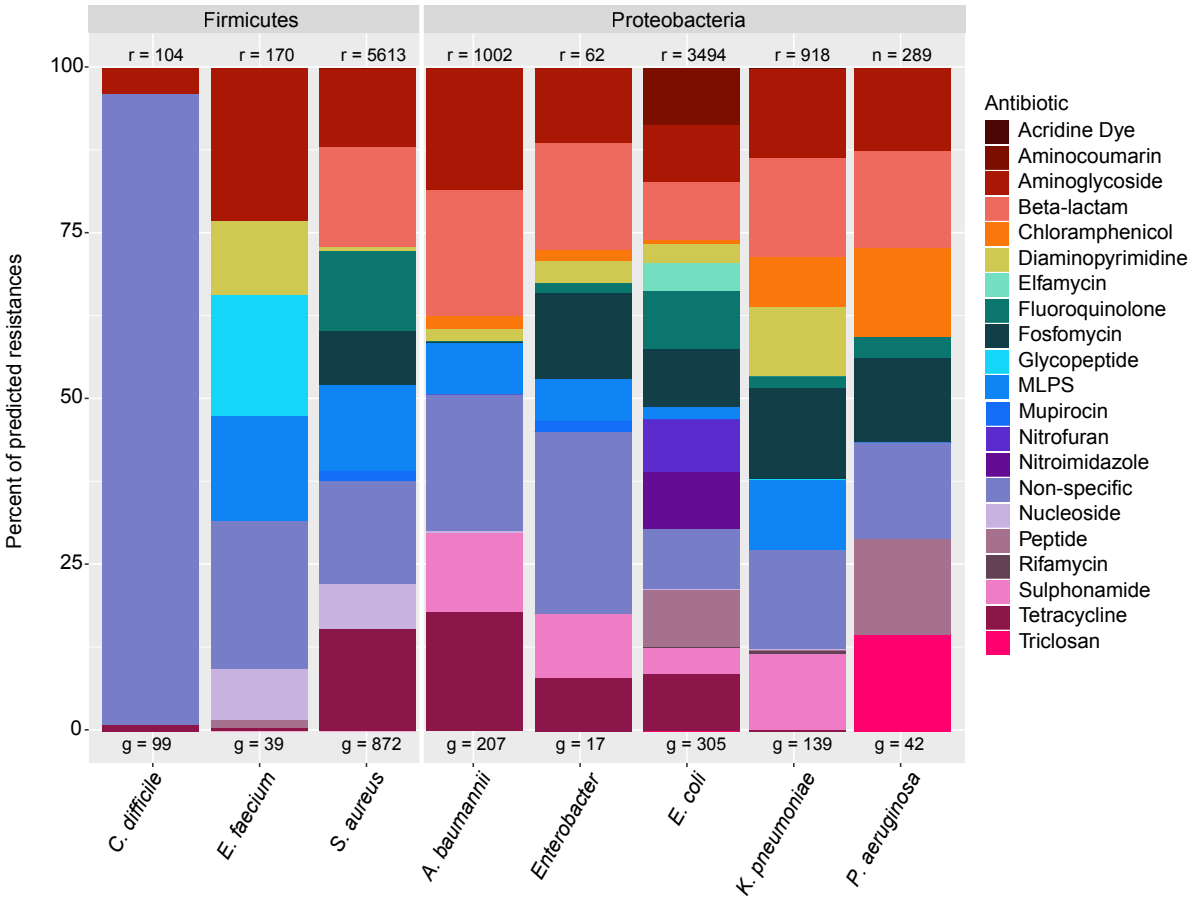


Figure 3.14. Proportions of antibiotic categories that resistances are predicted against in pathogenic bacterial genomes. The percentage of predicted resistance phenotypes in pathogenic genomes per antibiotic class was calculated for each species. r = total number of predicted resistance phenotypes for that group; g = genomes with predicted resistance in that group.

Having assessed the frequency and proportions of genomic resistance in certain pathogenic species, I wanted to directly compare between pathogenic and commensal (from the HBC) genomes of related species. I have performed this analysis at the family level for Enterococcaceae (includes pathogenic *E. faecium*), Peptostreptococacceae (includes pathogenic *C. difficile*), Staphylococcaceae (includes pathogenic *S. aureus*), and Enterobacteriaceae (includes pathogenic *Enterobacter*, *E. coli* and *K. pneumoniae*) due to limitations in numbers of HBC genomes for certain species. Comparing the number of determinants per genome directly between commensal HBC and pathogenic isolates of the

same families, we can see that the pathogenic isolates have significantly more determinants per genome on average, except for Peptostreptococcaceae (Fig. 3.15).

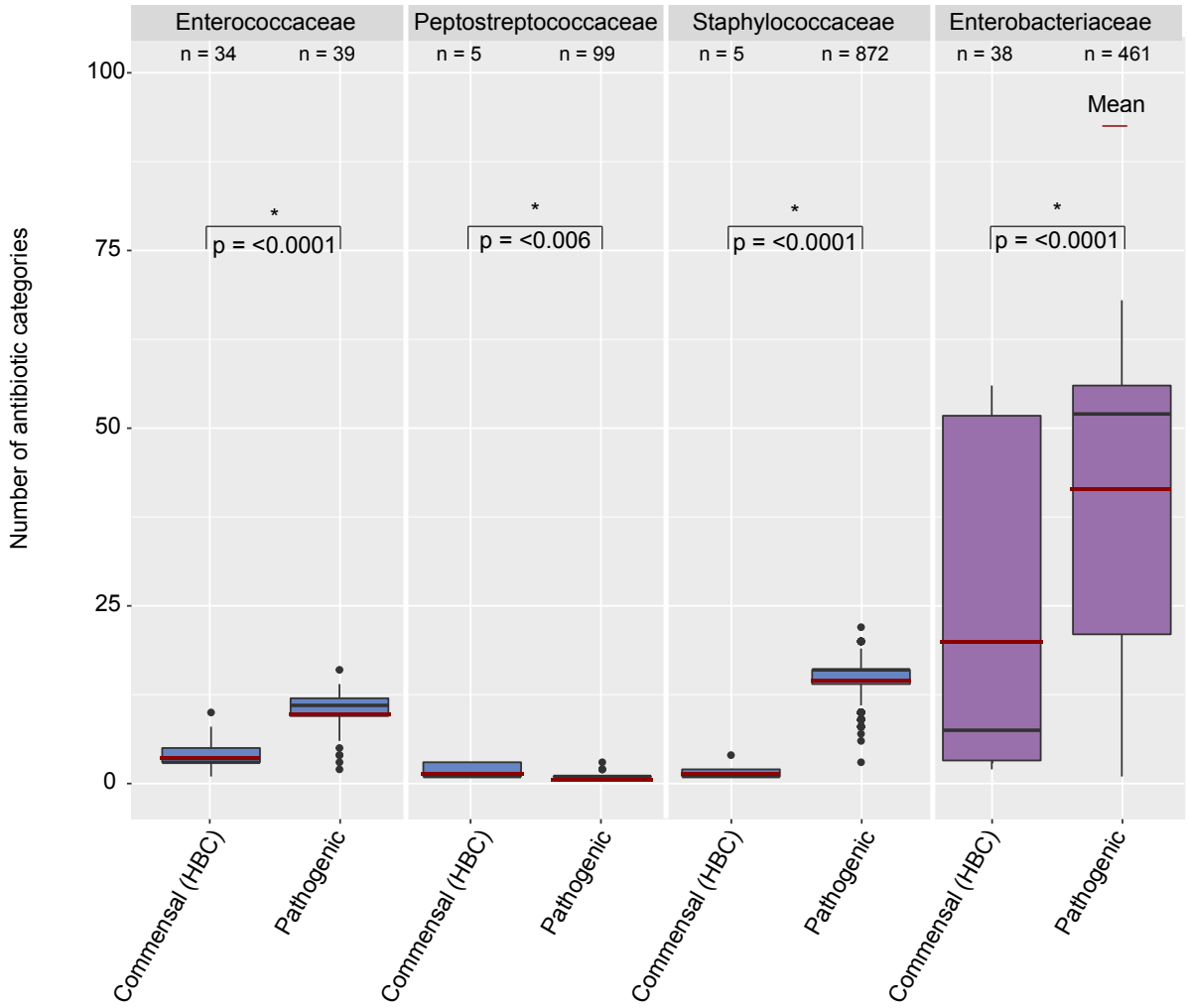


Figure 3.15. Interquartile range of number of antibiotic resistance determinants per isolate in commensal versus pathogenic isolates. The total number of antibiotic resistance determinants was calculated for each genome studied and the distribution in this number plotted at the family level, where families were represented by more than one isolate in both datasets. n = number of isolates in that group with predicted resistance. Interquartile range was determined using the Tukey method; the black circles show outlier observations and the thick black line represents the median. The red line represents the mean. Statistical significance determined by Mann-Whitney U test; p = significant < 0.05.

However, we can also see that Enterococcaceae and Enterobacteriaceae each have broadly similar resistance profiles in both commensal HBC and pathogenic isolates of these families (Fig. 3.16). In contrast, pathogenic Staphylococacceae have more categories of antibiotic resistances predicted than commensal HBC isolates. Peptostreptococacceae demonstrate a

different pattern again: the pathogenic isolates are dominated by non-specific resistances, whereas the commensal isolates have more tetracycline and MLPS resistances predicted. This data shows that although pathogenic isolates of Enterococcaceae and Enterobacteriaceae carry more resistance determinants, overall commensal and pathogenic isolates in these families are predicted to be resistant to similar antibiotics. In contrast, commensal isolates of Peptostreptococacceae and Staphylococcaceae are predicted to be resistant to different antibiotics than pathogenic isolates.

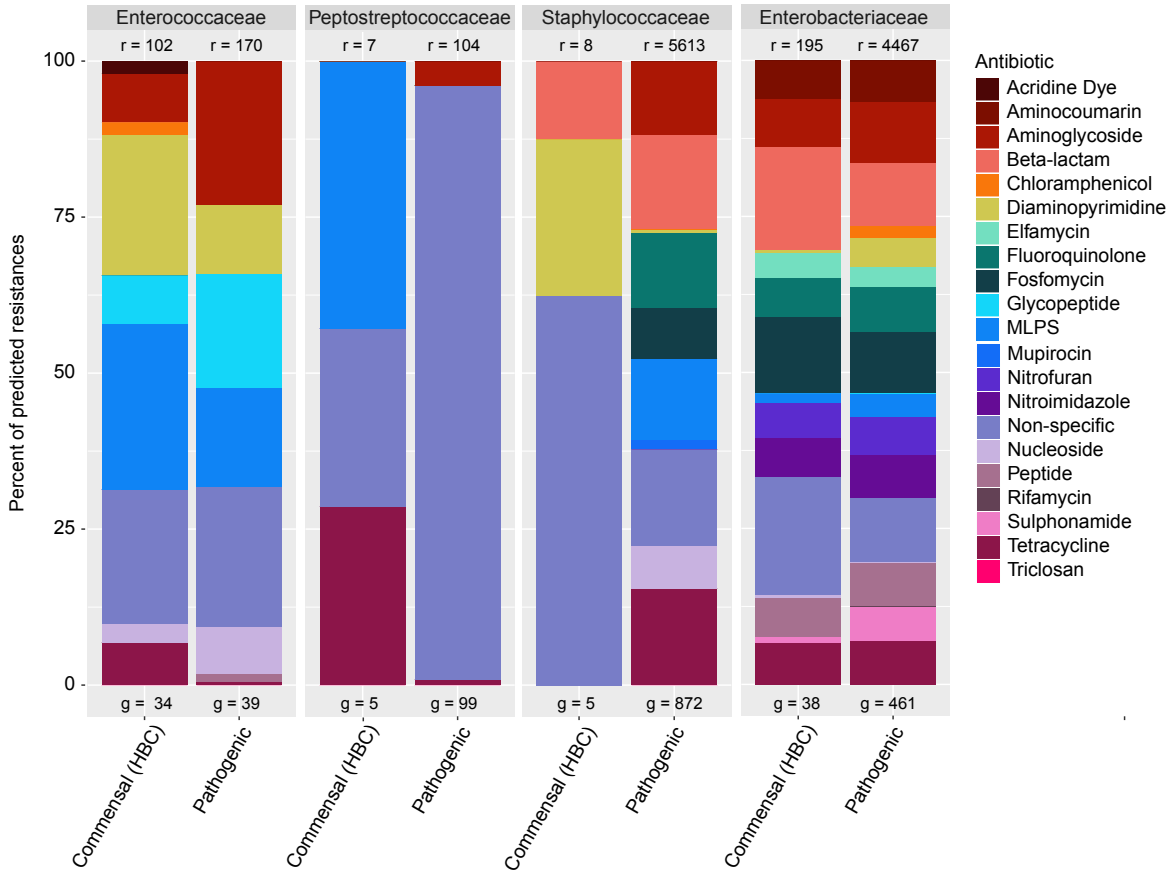


Figure 3.16. Proportions of antibiotic categories that resistances are predicted against in commensal HBC versus pathogenic bacterial genomes. The number of predicted resistance phenotypes in genomes per antibiotic class was calculated for each family. r = total number of predicted resistance phenotypes for that group; g = genomes with predicted resistance in that group.

3.3 Discussion

In this chapter, I aimed to determine the presence of known antibiotic resistance determinants in pure isolates of commensal gut bacteria and see if the whole gut microbiome may contribute to the problem of antibiotic resistance or only certain members. I therefore screened a collection of 737 phylogenetically diverse human gut bacterial genomes for known antibiotic resistance determinants described by CARD. This identified 178 unique antibiotic resistance determinants across the HBC, predicted to confer resistance against 19 antibiotic categories (including a category for non-specific antibiotic resistance). This included all but one of the antibiotics on the WHO Essential Medicines List, oxazolidinone. As in previous studies[140,142], tetracycline resistance genes were some of the most prevalent across the diversity of the HBC.

There were nine other antibiotics described in CARD to which there were not any predicted specific resistances: rifamycin, polyamine, triclosan, isoniazid, para-aminosalicylic acid, antibacterial free fatty acids, bicyclomycin and pyrazinamide (listed in order of the number of resistance determinants against these antibiotics in CARD). Rifamycin has broad-spectrum activity against Gram-negatives and Gram-positives, but is mainly used to treat TB infections[223]. There are no *Mycobacterium* isolates in the HBC, which we would expect given the criteria for the human donors for the HBC study to be healthy and without any bacterial infections. Rifamycin resistance develops by mutations in the RNA polymerase gene, and if the donors had not received rifampicin treatment recently then this may explain the absence of rifamycin-conferring resistance mutations in genomes of gut bacteria isolated from those people. Isoniazid[224], para-aminosalicylic acid and pyrazinamide are also used to treat TB and so a similar explanation may apply. However, there were non-specific determinants present in these genomes that included rifamycin as one of the antibiotics they are described as

conferring resistance to (Appendix II). Similarly, triclosan and oxazolidinone were included in some non-specific determinant descriptions. Bicyclomycin is an old antibiotic with weak activity that is being revived for combinatorial therapy of drug resistant infections[225]. Polyamine is a compound that can increase the susceptibility of bacteria to other antibiotics[226], such as beta-lactams, that is also being developed as an option combinatorial therapy[227]. Antibacterial free fatty acids are another experimental alternative to antibiotics[228]. Since these antibiotics are not currently commonly used as treatment of bacterial infections, resistance is less likely to have had the chance to develop or spread as often as resistance to other, more commonly used antibiotics. This could explain the absence of resistance determinants to these three antibiotics in the HBC.

Overall, resistances were enriched among the Proteobacteria, and particularly the Enterobacteriaceae, in terms of diversity of resistance determinants, abundance of resistance determinants per isolate, and the number of isolates with predicted resistance. Isolates of Enterobacteriaceae harboured between 1 and 31 non-specific resistance determinants; this means the Proteobacteria could be resistant to more antibiotics than assumed when just considering the specific antibiotic resistance determinants. Since the HBC isolates were isolated from the guts of healthy adults who had not taken antibiotics in at least six months, and the microbiome is a diverse population of microorganisms, it is unlikely that the Enterobacteriaceae isolates have been more exposed to antibiotics than non-Proteobacteria isolates. As many of the most important bacterial pathogens are Gram-negatives – often Enterobacteriaceae or other Proteobacteria – some of the most commonly used antibiotics are those that target Gram-negatives. Therefore, it is possible that when taking antibiotics the commensal Enterobacteriaceae are theoretically more likely to be impacted and thus under more selective pressure to develop antibiotic resistance. If this were maintained long-term

(beyond the period of antibiotic treatment), this may explain the observations in this chapter. This is plausible, however there is often a fitness cost associated with antibiotic resistance[229], implying that any acquired antibiotic resistance genes or mutations are not likely to be maintained in the absence of antibiotic selective pressure.

Alternatively, it is possible that the HBC Enterobacteriaceae acquired resistance genes from bacteria in the environment or diet, in line with the "One Health" concept discussed in the Introduction to this thesis, and that this has happened more commonly in this taxon. However, horizontal gene transfer is not necessarily more common in Proteobacteria than other phyla: a recent study found that phylogeny did not have a significant impact on genome fluidity[230]. Conversely, and more likely, these results may reflect how the databases are predominantly created from information generated by studying bacterial pathogens; throughout this chapter it is the families of bacteria known to contain pathogenic species that have the most antibiotic resistance predicted. The databases would therefore be biased towards predicting antibiotic resistance genes or mutations in bacterial genomes more similar to those studied for the creation of said databases. This bias could explain the enrichment of resistance in Enterobacteriaceae over more distantly related bacteria, such as the Firmicutes.

Despite the apparent enrichment in Proteobacteria, there is still a diverse range of antibiotic resistance predicted in large proportions of the other phyla in the HBC. For example, Firmicutes were enriched for diaminopyrimidine (trimethoprim) resistance; *dfr* trimethoprim resistance genes have been described in both Gram-negative and Gram-positive bacteria, but are thought to be intrinsic to enterococci[231]. Moreover, Bacteroidetes were enriched for beta-lactam resistance; beta-lactamases are often endogenously produced by these bacteria[232]. Finally, tetracycline resistance was enriched in Bacteroidetes, and was the most common type of resistance in the HBC. Tetracycline resistance has been reported as increasingly common in

Bacteroidetes[233] and the gut microbiota generally[123], and serves to highlight how antibiotic resistance can become widespread in commensal bacteria. Whilst tetracyclines are now infrequently used in human medicine for this reason, there is a possibility of cross resistance with the last resort glycylcycline antibiotic tigecycline, and so the presence of tetracycline resistance determinants remains relevant. There were also instances of resistance to antibiotic classes being observed significantly less often than expected, such as to beta-lactams in the Actinobacteria isolates; this would suggest either the isolates of those phyla are all susceptible to that antibiotic, or that there are alternative resistance mechanisms not identified using the methods applied here.

Many of the predicted resistances in novel HBC genomes, most of which belong to the phylum Firmicutes, were to tetracycline. Other resistances were also predicted such as MLPS, beta-lactams and diaminopyrimidine (trimethoprim). If known resistance genes were not identified in novel isolates, then the fact there are still gut bacteria to be discovered[216] does not necessarily have serious consequences for the problem of antibiotic resistance. However, almost half of the novel HBC isolates did contain antibiotic resistance genes or mutations. Since there are uncharacterised gut bacteria, this emphasises that we do not yet fully understand the gut microbiota and the extent of their contribution to antibiotic resistance. Overall, significantly more known HBC genomes contained resistance determinants than novel HBC genomes; the same statement is true for the Bacteroidetes phylum specifically. There was no significant difference between the numbers of known or novel Firmicutes or Proteobacteria with resistance determinants. For the Proteobacteria, it is possible that this is due to sample bias, as there were only five novel Proteobacteria isolates, whereas in Firmicutes it is almost a 50:50 division of novel and known genomes. This suggests that known bacteria, which are more likely to be more closely related to known pathogenic bacteria, are

more likely to contain antibiotic resistance, and that more distantly related (novel and uncharacterised) bacteria are less likely to contain antibiotic resistance. However, it is possible that again database bias is impacting these results, and that known bacteria are more likely to be predicted to contain antibiotic resistance than more distantly related bacteria. Overall though, this data emphasises that commensal gut microbiota do harbour clinically relevant antibiotic resistances and that this reservoir, much of which remains uncharacterised, is more extensive than currently realised.

Antibiotic resistances were commonly predicted throughout the HBC and there was variation observed even between closely related isolates; these statements are known to be true for pathogens as well as commensals. The commensal families most enriched for genetic determinants of antibiotic resistance are those known to contain opportunistic pathogenic species or isolates, such as Enterobacteriaceae and Enterococcaceae. In the last part of this chapter, I studied the differences in antibiotic resistance genotypes between commensal and pathogenic bacteria. I found that resistance determinants were observed more frequently in pathogenic isolates, and again especially in Proteobacteria.

As mentioned, antibiotics are often targeted towards these bacteria since many pathogenic species are Gram-negative Proteobacteria (Fig. 1.3); it is possible then that these bacteria are more exposed to antibiotics and thus more likely to develop antibiotic resistance. This may be partially true, but antibiotics can have “off-target” effects on other bacteria; other antibiotics are broad-spectrum and thus designed to target several types of bacteria, including Gram-positives. This means that other members of the gut microbiome will also be impacted by antibiotic use and thus also under selective pressure to develop resistance. Potentially, Proteobacteria might be more genetically capable of developing antibiotic resistance: Gammaproteobacteria (which includes Enterobacteriaceae) are known to have super-

integrons as an integral component in their genomes[234], allowing for multiple-drug resistance to develop rapidly. However, this concept of increased HGT in particular taxa was discussed earlier in this section and is unlikely to be the case. Alternatively, the enrichment of antibiotic resistance determinants in Proteobacteria may be due to efflux pumps being more common in Gram-negative bacteria than in Gram-positives[235]. Here, the majority of resistance determinants in the Proteobacteria were non-specific, of which efflux mechanisms dominated. This confirms that these non-specific mechanisms are a major contributor to the antibiotic resistance potential in Proteobacteria and may explain why pathogenic Proteobacteria are so prone to cause multi-drug resistant infections.

However, the average number of determinants per genome was lower in *Enterobacter* spp. than in *E. faecium* and *S. aureus*. Since *Enterobacter* spp. are Proteobacteria, this is perhaps unexpected based on my earlier findings, but could be explained by the fact that these bacteria have only relatively recently been described as an emerging multi-drug resistant threat[236]. In addition, pathogenic *E. faecium* and *C. difficile* genomes had resistances predicted to fewer antibiotic categories than pathogenic *S. aureus* or the Proteobacteria species. Pathogenic Gram-negative Proteobacteria are known for having mobile genetic elements that can contain several resistance genes at once that can confer resistance to different antibiotics[237]. In addition, pathogenic *S. aureus* is well-known for being resistant to multiple antibiotics[238], despite being Gram-positive. This may explain the difference in predicted resistance profiles observed.

When I compared the pathogenic versus commensal genomes, pathogenic isolates were generally found to have more genetic resistance determinants per genome on average than commensal isolates, particularly Enterobacteriaceae. The exception was for Peptostreptococcaceae, although it is possible that this is due to the number of genomes

studied (five commensals versus 99 pathogens) and studying additional genomes of commensal Peptostreptococcaceae will determine whether this pattern continues to be seen. In general for the comparisons between commensals and pathogens, the sample bias means these results should be treated with caution.

Finally, I found that for Peptostreptococacceae and Staphylococacceae, commensal genomes had different profiles of predicted antibiotic resistances to pathogenic isolates, whereas Enterobacteriaceae and Enterococcaceae shared similar predicted resistance profiles between commensal and pathogenic isolates. It is possible that this could be explained by more frequent horizontal sharing of resistance determinants within Enterobacteriaceae or Enterococcaceae than other bacterial families, though as discussed above phylogeny has not been observed to impact this process[230]. However, that study was performed at the species level, and so in the future it would be interesting to more specifically estimate and compare the frequency of horizontal gene transfer at other taxonomic levels, such as within and between bacterial families. This will also help to understand the spread of antibiotic resistance better. Yet again, it is important to consider the potential bias of the database of antibiotic resistance genes and mutations. This bias could also explain the observed higher proportions of pathogenic species with predicted resistance (and the high numbers of resistance determinants per genome) compared to commensal HBC isolates: the pathogenic isolates will be more similar to the isolates studied and used to create the database.

It is important to acknowledge that this study has only used one collection of commensal gut bacteria and one database of antibiotic resistance determinants; repeating these analyses with additional genomes and alternative databases will help confirm these findings. However, since the CARD database is a regularly updated and extensive collection of antibiotic resistance determinants, and the HBC is a recent and diverse collection of gut bacteria, these

new findings add to current knowledge regarding the gut microbiome as a reservoir of antibiotic resistance. The most important message from this chapter is that antibiotic resistance genotypes are common in commensals and often share similar predicted resistance profiles to related pathogens. In addition, many of the resistance genotypes are to antibiotics on the WHO list of essential medicines[21], emphasising the potential clinical relevance of these observations.

As the findings in this chapter are based purely on predicted genotypes using known antibiotic resistance determinants, this does not necessarily preclude that the isolates studied here are not resistant to other antibiotics – or even that they are resistant to the ones predicted. As discussed, there may be a database bias that makes it more likely to predict antibiotic resistance in bacteria more closely related to pathogenic species. Only phenotypic testing of antibiotic sensitivity of HBC isolates will confirm whether or not the observations in this chapter are accurate predictions of phenotypic antibiotic resistance. I will investigate this in the next chapter.

Chapter 4: Determination of phenotypic antibiotic resistance in commensal gut bacteria and the accuracy of genomic predictions

4.1 Introduction

4.1.1. Overview

In the previous chapter, I identified the presence of a range of clinically relevant genetic antibiotic resistance genes and mutations in the 737 genomes of the HBC. This analysis identified a significant enrichment of ARGs in the Proteobacteria members of the HBC. In this chapter I investigate whether Proteobacteria are phenotypically enriched for antibiotic resistance using the HBC culture collection, or whether CARD is biased towards identifying resistance genotypes in this phylum. In addition, I assess how accurate these predictions of antibiotic resistance in commensal gut bacteria are by determining phenotypic susceptibility and resistance to a range of clinically relevant antibiotics in a subset of the HBC.

The gold-standard method for determining isolate-specific and phenotypic antibiotic susceptibility is culture-based antibiotic susceptibility testing (AST). Developed mainly for clinical isolates of pathogenic bacteria, AST involves culturing the bacterium of interest in the presence of an antibiotic and observing its ability to grow. One common method is measuring the size of a zone of inhibition: this features a paper disk containing a single concentration of the antibiotic that is placed on an agar plate that has been inoculated all over with the isolate of interest – after 24 hours incubation this produces a bacterial lawn. The sensitivity of the isolate to that antibiotic determines how close to the disk it can grow: the more sensitive the isolate, the less close to the disk it will be able to grow. The diameter of the zone where no growth occurs (the zone of inhibition) is measured. The antibiotic disks are 0.5cm in diameter,

allowing several disks to be placed on a single inoculated plate and many antibiotic and isolate combinations to be tested using few resources[239]. Another method includes measuring the minimum inhibitory concentration (MIC). This involves exposing a bacterial isolate to a stepwise increasing range of antibiotic concentrations. This can be performed using broth microdilution (a series of prepared liquid culture mediums each with a different antibiotic concentration is inoculated with the isolate of interest) or using antibiotic gradient strips. These are small rectangular paper strips that contain a gradient of an antibiotic and are placed on top of an agar plate that has been inoculated to produce a bacterial lawn. The antibiotic diffuses into the agar – the more concentrated end will diffuse further – and then after 24 hours incubation the antibiotic concentration at which the isolate is no longer able to grow alongside the strip (the MIC), is measured.

4.1.2. Defining isolates as antibiotic-susceptible or -resistant

The zone of inhibition or MIC is compared to guidelines provided by the European Committee on Antimicrobial Susceptibility Testing (EUCAST) or the Clinical and Laboratory Standards Institute (CLSI); international agencies who study and determine zone of inhibition breakpoints that categorise a certain isolate as resistant, susceptible or having intermediate sensitivity. This depends on having culturable, purified isolates of the bacteria of interest – which, until recently has been difficult to achieve for the majority of gut bacteria. Therefore, these breakpoint guidelines are only available for a small number of pathogenic bacteria species.

As the costs of DNA sequencing decrease, whole genome sequencing antibiotic susceptibility testing (WGS-AST) is being investigated as an alternative method to culture-based AST to

determine antibiotic sensitivity. For example, the genome of a disease-causing bacterium isolated from a patient is sequenced and then searched for known antibiotic resistance determinants; any found are assumed to confer phenotypic resistance to the corresponding antibiotic. This approach is considered a rule-based WGS-AST method: the presence of antibiotic resistance determinant A in the genome confers resistance to antibiotic B. Therefore, rule-based methods rely on databases that contain information regarding the nucleotide or amino acid sequence of resistance determinants and which antibiotics they confer resistance to[240]. These methods are attractive as they offer the potential for results to be available in a matter of hours rather than days and to screen many more bacteria than culture-based methods. Moreover, rule-based approaches offer easily interpretable results and are currently the most popular methods for translation of WGS-AST into a clinical setting. However, the reliance on databases of known resistance determinants can introduce problems; as discussed in the previous chapter, these databases may introduce bias towards certain bacteria. This can lead to a "false negative" result: no genetic resistance determinant is present in an isolate, but the isolate is phenotypically resistant to a particular antibiotic. These rule-based methods therefore vary in their accuracy for antibiotic/taxon combinations[241]; to my knowledge, how accurate they are for gut bacteria has not yet been assessed.

In this chapter, I generate and study the *in vitro* phenotypes of a diverse set of human commensal gut bacteria in the HBC against nine clinically relevant antibiotics (all of which belong to classes on the WHO List of Essential Medicines). As previously discussed, phenotypic data can be used to determine breakpoints of antibiotic concentrations where isolates are considered susceptible or resistant. Since published breakpoints are limited for gut bacteria, I defined a system for determining whether the isolates I studied should be considered

susceptible or resistant. I then determined the enrichment of phenotypic resistance in isolates representing the four main gut phyla of bacteria: Actinobacteria, Bacteroidetes, Firmicutes and Proteobacteria[242]. I also combined this phenotypic data with the genotypic data from the previous chapter to determine whether a rule-based prediction method of antibiotic resistance inferred from the presence of known resistance determinants can be accurately applied to human commensal gut bacteria. I then applied a comparative genomics and phenotype approach to identify and study candidate novel antibiotic resistance genes.

4.2 Results

4.2.1 Phenotypic screening of antibiotic resistance in a subset of 73 HBC isolates

Having identified the presence of genetic antibiotic resistance determinants using CARD, I leveraged our ability to culture anaerobic gut bacteria[146,151] to assess the phenotypic response to antibiotics. 73 phylogenetically diverse isolates from the HBC (Fig. 4.1) were selected, representing all four phyla, approximately 10 % of the culture collection and 14 % of the number of different species (39/273) contained in the HBC: 16 isolates of Proteobacteria (seven species in two families); 21 isolates of Bacteroidetes (11 species in two families); 11 isolates of Actinobacteria (five species in two families); 25 isolates of Firmicutes (16 species in three families). These isolates have a total of 115 predicted resistance phenotypes against 16 of the 17 antibiotic resistance classes discussed in Chapter 4 (acridine dye was not included since it is not a clinically relevant antibiotic[243]). Each isolate was tested for *in vitro* sensitivity against nine antibiotics that are clinically relevant and on the WHO list of essential medicines[21]: amoxicillin and ceftriaxone (two different sub-types of beta-lactams), ciprofloxacin (a fluoroquinolone), erythromycin (a macrolide), gentamicin (an

aminoglycoside), metronidazole (a nitroimidazole), tetracycline, trimethoprim (a diaminopyrimidine), vancomycin (a glycopeptide).

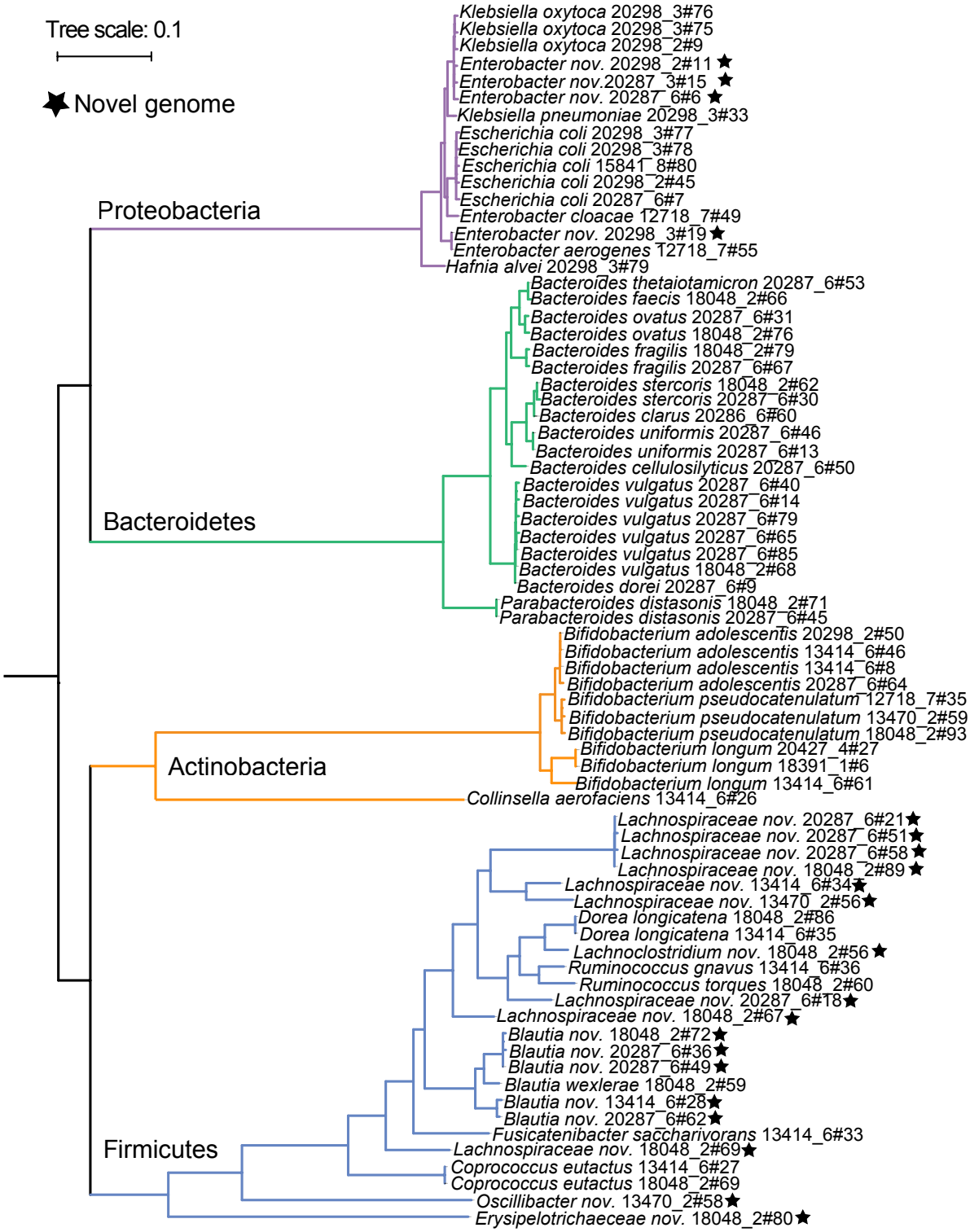


Figure 4.1. A phylogeny of 73 isolates from the HBC selected for selective phenotypic screening of antibiotic sensitivity. The phylogeny was inferred from 40 core genes of 73 whole genome sequences of gut microbiota isolated from healthy human faecal samples. These samples are a subset of the 737 isolates in the Human Gastrointestinal Bacterial Culture Collection (HBC) chosen to represent all four phyla.

To measure antibiotic sensitivity, I performed zone of inhibition analysis, due to the benefits described and because it is easy to perform in the confines of an anaerobic cabinet, where culturing anaerobic gut bacteria must take place. This method places single-concentration antibiotic disks (Table 4.1) on agar plates inoculated with the isolate of interest and measures the diameter of the zone of inhibition (the area surrounding the antibiotic disk where the isolate does not grow). An average zone of inhibition diameter was calculated from three biological replicates of each isolate/antibiotic combination (73 isolates x 9 antibiotics = 657 combinations). These phenotypic tests were set up by myself and Mr Mark Stares.

Table 4.1. Single-concentration antibiotic disks used for phenotypic sensitivity testing. Disks were manufactured by Oxoid and the concentration chosen was based on CLSI or EUCAST or BSAC guidelines as indicated in the table. Disk concentrations are usually selected based on guidance from EUCAST or CLSI breakpoint information; however, information for anaerobes is lacking. Therefore, concentrations were mainly selected based on advice for *Enterococcus*, as a Gram-positive, facultative anaerobic Firmicute. Though Bacteroides are Gram-negative, the same concentration disks were used for consistency and to allow comparisons.

Antibiotic	Antibiotic class	Disk Concentration (μg)	Based on guidelines
Amoxicillin	Beta-lactam: penicillin	10	*Enterococcus*; CLSI 2015
Ceftriaxone	Beta-lactam: cephalosporin	10	*Enterococcus*; EUCAST 2018
Ciprofloxacin	Fluoroquinolone	10	*Enterococcus*; CLSI 2015
Erythromycin	Macrolide	15	*Enterococcus*; CLSI 2015
Gentamicin	Aminoglycoside	30	*Enterococcus*; EUCAST 2018
Metronidazole	Nitroimidazole	5	Miscellaneous; BSAC 2015
Tetracycline	Tetracycline	30	*Enterococcus*; CLSI 2015
Trimethoprim	Diaminopyrimidine	5	*Enterococcus*; EUCAST 2018
Vancomycin	Glycopeptide	30	*Enterococcus*; CLSI 2015

Density curves were plotted for average zone of inhibition diameter for each antibiotic (Fig. 4.2), where each dot represents an average zone size for a particular isolate/antibiotic combination. These graphs therefore represent the proportions of isolates in each phylum with a certain zone size. Visualising the data in this way identifies trends in the range and frequency of zone of inhibition sizes for each antibiotic and each phylum. A larger zone of inhibition size (e.g. 60 mm) suggests that that isolate was very sensitive to that antibiotic; in contrast, a smaller zone of inhibition size (e.g. 5 mm) suggests that the isolate was less sensitive.

A very narrow range of zone of inhibition sizes occurs several times: Actinobacteria with amoxicillin, gentamicin, tetracycline and vancomycin; Bacteroidetes with gentamicin; Proteobacteria with ciprofloxacin, gentamicin, metronidazole, tetracycline and vancomycin (Fig. 4.2) These narrow distributions generally appear around the mid-range of average zone of inhibition sizes (20-40 mm), suggesting that all the isolates tested in those categories were sensitive (i.e., the antibiotic had an effect on these isolates). The exception is for Bacteroidetes and gentamicin, where the majority of isolates were not sensitive at all (zone of inhibition = 0 mm). Only two Bacteroidetes isolates were slightly sensitive with a zone of inhibition that was always smaller than 10 mm. This suggests that this phylum is particularly unaffected by gentamicin. Moreover, in some cases all isolates of a phylum were not sensitive to an antibiotic, such as Proteobacteria for metronidazole and vancomycin. Much bigger ranges in zone of inhibition size and thus sensitivity are also present: for example Actinobacteria with metronidazole; Bacteroidetes with amoxicillin, ceftriaxone, erythromycin and tetracycline; Firmicutes with ceftriaxone, erythromycin, tetracycline and trimethoprim. This indicates that some isolates in those phyla are more resistant to these antibiotics than other isolates. Since only eight bacterial families are represented by these 73 isolates, and the majority of isolates

in each phylum belong to a single family, the observed variation in antibiotic sensitivity also occurs within bacterial families as well as phyla.

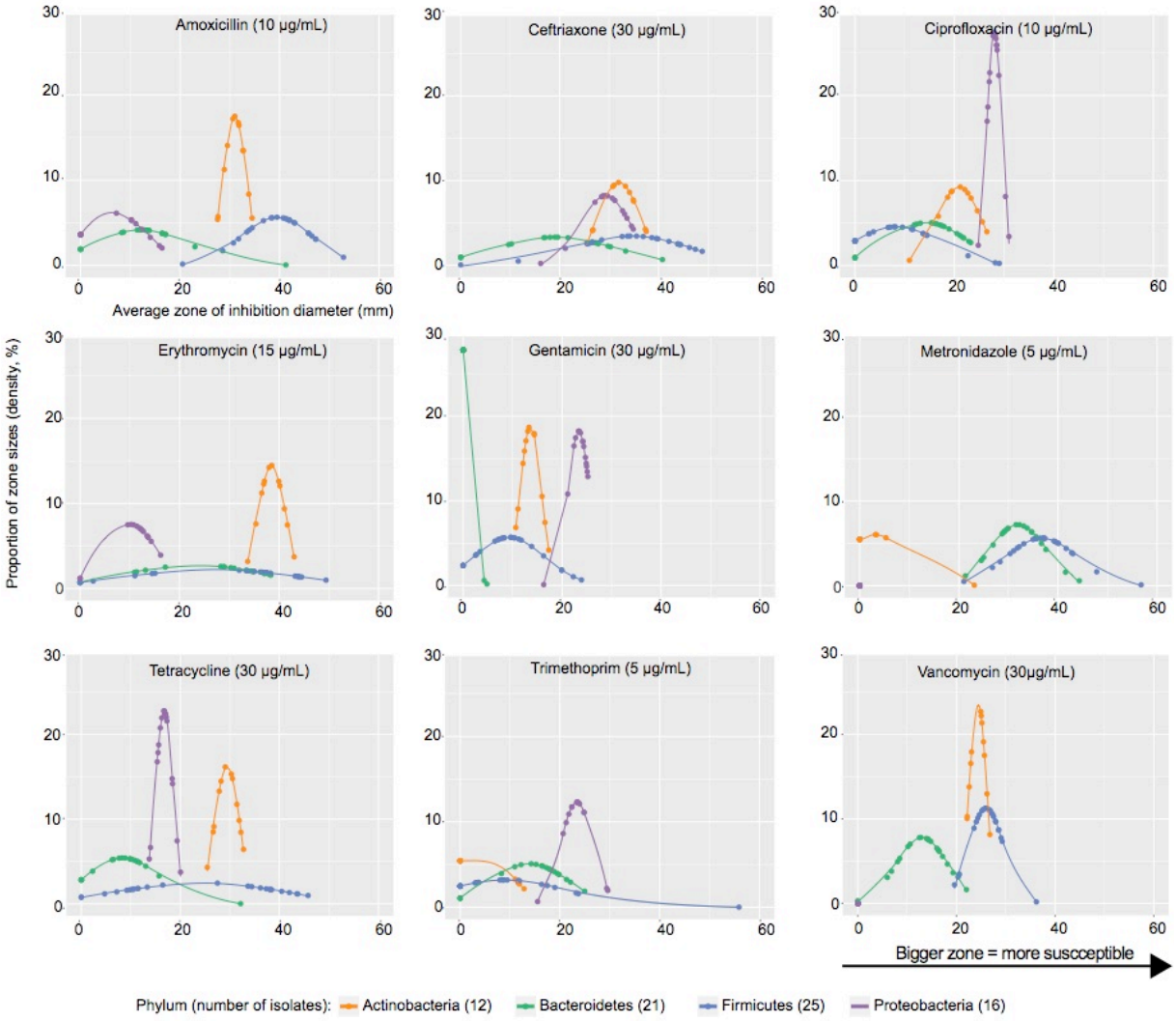


Figure 4.2. Density curves of zone of inhibition size among 73 isolates of the four main phyla of human gut bacteria. The isolates were screened for phenotypic antibiotic resistance against the following antibiotics: amoxicillin and ceftriaxone (beta-lactams), ciprofloxacin (a fluoroquinolone), erythromycin (a macrolide), gentamicin (an aminoglycoside), metronidazole (a nitroimidazole), tetracycline, trimethoprim (a diaminopyrimidine), and vancomycin (a glycopeptide) using disks containing a single concentration of an antibiotic. The screen was repeated in three biological replicates and zone of inhibition diameter averaged per isolate. Density curves for average zone of inhibition size was plotted for each phyla and antibiotic: each dot represents an average zone of inhibition size for a particular isolate/antibiotic combination. The x axis is the zone of inhibition size in millimetres: a larger average zone of inhibition means that isolate was more sensitive to an antibiotic than a smaller zone of inhibition. The y axis is the density or proportion of isolates with a particular zone size. Some phyla have a very narrow range of zone of inhibition sizes, e.g. Proteobacteria and ciprofloxacin or Actinobacteria and vancomycin. Other phyla have a much bigger range in zone of inhibition or susceptibility, e.g. Firmicutes and tetracycline. Mr Mark Stares helped generate zone of inhibition data.

4.2.2 Comparison of zone of inhibition sizes between isolates with and without genetic antibiotic resistance determinants

Combining the phenotypic data for these 73 isolates with the genomic data generated in the previous chapter it is possible to assess whether the antibiotic resistance genotype (or predicted phenotype) corresponds with the actual antibiotic resistance phenotype. Thus, the accuracy of the proposed rule-based method of predicting antibiotic resistance in gut bacteria is determined. To my knowledge, this is the first time this comparison of genotypic and phenotypic antibiotic resistance data has been performed for a diverse collection of gut bacteria isolated from healthy humans.

Initially, I used the Tukey method to study the range of zone sizes in isolates with and without predicted resistance and identify outliers with abnormally small zone diameters for each antibiotic (Fig. 4.3). This showed that there were isolates completely resistant (zone diameter = 0 mm) to an antibiotic despite not harbouring any resistance determinants from CARD to that drug. This occurs for all antibiotics and indicates that the CARD based predictions of genetic resistance determinants are not completely accurate for these gut bacteria. To investigate this further, the bacteria need to be categorised as resistant or susceptible to each antibiotic.

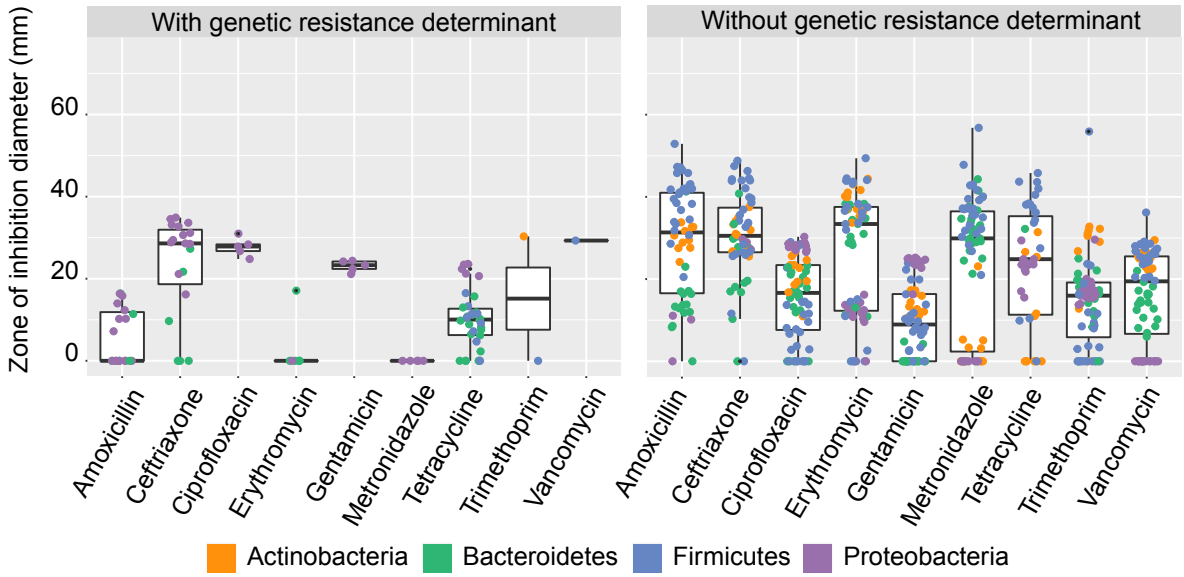


Figure 4.3. Range of average zone of inhibition between isolates with and without the presence of genetic resistance determinants. The average zone of inhibition diameters in isolates with predicted genetic resistance determinants against a certain antibiotic were compared against those in isolates with no predicted resistance. The boxplots show the interquartile range, median and limits at 1.5IQR above and below. Overall, there was a large range in zone of inhibition size in isolates without predicted genetic resistance determinants, with some isolates completely resistant to each antibiotic studied even in the absence of genomic resistance.

4.2.3 Defining a system for categorising gut bacteria as resistant or susceptible to antibiotics and considering the spectrum of antibiotics

Guidelines for defining resistance and susceptibility exist for a range of pathogens, though are less well-characterised in anaerobes and Gram-positive bacteria. Moreover, they are used for categorising closely related taxa, usually at the species level. In this thesis I wanted to compare at higher taxonomic levels, therefore I defined a scale using the phenotypic data I generated. Where very narrow ranges in sensitivity occur, cut-off points (breakpoints) in zone of inhibition size for categorising isolates as susceptible or resistant will be closer together (e.g. zone > 7.5 mm = susceptible but zone < 2.5 mm = resistant). Where broad ranges in sensitivity occur, breakpoints will be further apart (e.g. zone > 45 mm = susceptible and zone < 15 mm = resistant). This makes it clear that breakpoints for susceptibility/resistance must be defined for each antibiotic. Ideally, this would also be defined for individual taxa – in pathogens this is

done at the species level using thousands of clinical and sometimes environmental isolates. In this study, there are not enough isolates to be able to do that and a scale was defined using all 73 HBC isolates for each individual antibiotic. I determined the interquartile range of the zone of inhibition sizes and categorised isolates 'resistant' if the zone of inhibition size was in the lower quartile or smallest 25 % of all zone sizes for a particular antibiotic (Fig. 4.4, Table 4.2). Similarly, isolates were categorised 'susceptible' if the zone size was in the upper quartile or largest 25 % of all zone sizes for a particular antibiotic. Isolates with zone sizes in the middle 50 % require further testing to determine whether or not they should be considered susceptible or resistant and are referred to as intermediate sensitivity; these will be excluded from downstream analyses.

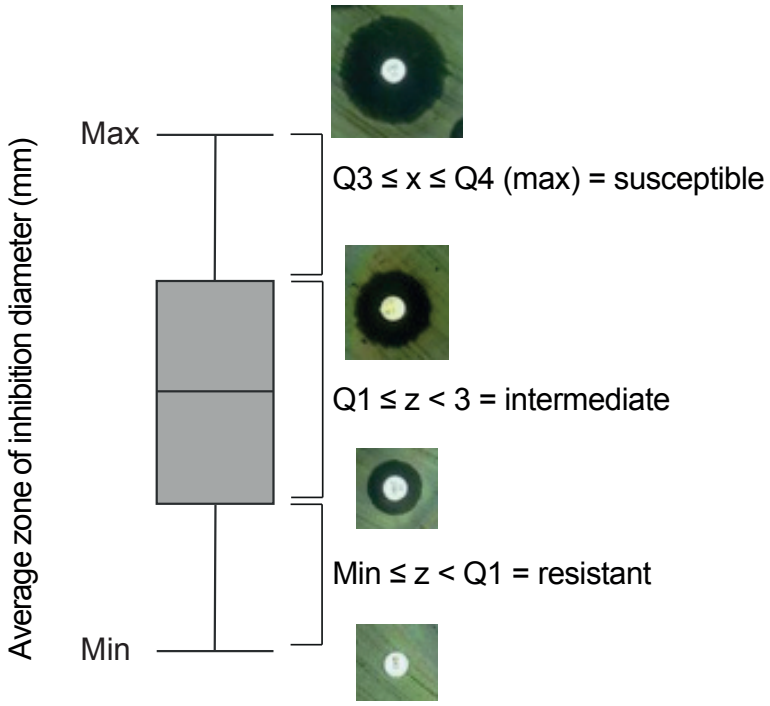


Figure 4.4. Proposed scale to define antibiotic resistance and susceptibility in human gut microbiota. I measured antibiotic susceptibility using single-concentration antibiotic disks and zone of inhibition size for nine antibiotics in 73 isolates (657 phenotypes in total). I ordered the zone of inhibition sizes for an individual antibiotic from largest to smallest and determined the interquartile range. I considered isolates 'resistant' if the zone size was in the lower quartile or smallest 25 % of all zone sizes for a particular antibiotic. Similarly, isolates were considered 'susceptible' if the zone size was in the upper quartile or largest 25 % of all zone sizes for a particular antibiotic. Isolates with zone sizes in the middle 50 % require further testing to determine whether or not they should be considered susceptible or resistant. Z = zone of inhibition diameter.

Table 4.2. Zone of inhibition limits used to categorise isolates of human gut bacteria as antibiotic-resistant or -susceptible. Zone of inhibition sizes for nine antibiotics in 73 isolates of human gut bacteria were determined. I ordered the zone of inhibition sizes for an individual antibiotic from largest to smallest and determined the interquartile range. I considered isolates 'resistant' if the zone size was in the lower quartile (min ≤ z < q1, where z = average of zone of inhibition in millimetres) or smallest 25 % of all zone sizes for a particular antibiotic. Similarly, isolates were considered 'susceptible' if the zone size was in the upper quartile (q3 ≤ z ≤ q4 (max)) or largest 25 % of all zone sizes for a particular antibiotic. Isolates with zone sizes in the middle 50 % (q1 ≤ z < q3) require further testing to determine whether or not they should be considered susceptible or resistant. For gentamicin, metronidazole and trimethoprim, all isolates with no zone of inhibition (0.00 mm) were considered resistant; any zone larger than 0.00 mm but smaller than the q3 value was considered intermediate.

Antibiotic	Resistant (min ≤ z < q1)	Intermediate (q1 ≤ z < q3)	Susceptible (q3 ≤ z ≤ q4 (max))
Amoxicillin	0.00 ≤ z < 11.09	11.09 ≤ z < 36.76	36.76≤ z ≤ 52.85
Ceftriaxone	0.00 ≤ z < 25.88	25.88 ≤ z < 34.90	34.90 ≤ z ≤ 48.85
Ciprofloxacin	0.00 ≤ z < 8.12	8.12 ≤ z < 25.61	25.61 ≤ z ≤ 30.98
Erythromycin	0.00 ≤ z < 11.35	11.35 ≤ z < 37.19	37.19 ≤ z ≤ 49.45
Gentamicin	0.00 ≤ z ≤ 0.00	0.00 < z < 19.92	19.92 ≤ z ≤ 25.21
Metronidazole	0.00 ≤ z ≤ 0.00	0.00 < z < 35.72	35.72 ≤ z ≤ 56.80
Tetracycline	0.00 ≤ z < 9.87	9.87 ≤ z < 30.31	30.31 ≤ z ≤ 45.80
Trimethoprim	0.00 ≤ z ≤ 0.00	0.00 < z < 21.32	21.32 ≤ z ≤ 55.91
Vancomycin	0.00 ≤ z < 6.85	6.85 ≤ z < 25.47	25.47 ≤ z ≤ 36.205

With this data it is possible to comment on the spectrum of antibiotic efficacy against diverse human commensal gut microbiota. For example, ceftriaxone appears a good example of a broad-spectrum antibiotic (i.e., an antibiotic that impacts several different bacterial taxa): all four phyla contain isolates that are very sensitive to this antibiotic (zone of inhibition > 25 mm, the midpoint in the range of zone of inhibitions measured). However, the proportions of isolates vary: for Proteobacteria, only one in 16 isolates (6.3 %) was susceptible to ceftriaxone, with one isolate (6.3 %) being resistant and the other 14 of intermediate sensitivity (87.5 %). In Bacteroidetes, one in 21 isolates (4.8 %) was susceptible to ceftriaxone, 12 isolates were resistant (57.1 %), and eight isolates were of intermediate sensitivity (38.1 %). In Actinobacteria, three in 12 isolates (25 %) were susceptible, one isolate was resistant (8.3 %) and eight isolates were of intermediate sensitivity (66.7 %). In Firmicutes, 14 in 25 isolates (56 %) were susceptible to ceftriaxone, four isolates were resistant (16 %) and seven were of intermediate sensitivity (28 %). On the other hand, gentamicin would be considered a more

narrow spectrum antibiotic (i.e., an antibiotic that impacts only specific taxa): no Bacteroidetes isolates were susceptible, but 19 were resistant (90.5 %; the other two isolates were of intermediate sensitivity (9.5 %)). In addition, no Actinobacteria were susceptible to gentamicin; all Actinobacteria isolates were of intermediate sensitivity. Four Firmicutes isolates were resistant to gentamicin (16 %) and another four were susceptible, with 17 isolates (68 %) of intermediate sensitivity. Overall, these observations suggest that sensitivity varies within bacterial families and phyla and that different antibiotics affect certain families and phyla to different extents.

4.2.4 Comparison of genomic predictions of antibiotic resistance with bacterial phenotypes and identification of unpredicted resistances

Four key genotype/phenotype combinations can be defined by comparing these two datsets (Table 4.3): Confirmed Resistance (genetic resistance and phenotypic resistance both observed); Confirmed Susceptibility (no genetic or phenotypic resistance); Unpredicted Susceptibility (genetic resistance predicted but phenotypically susceptible) and Unpredicted Resistance (no genetic resistance predicted but phenotypically resistant).

Table 4.3. Genotype/phenotype combinations of antibiotic sensitivity. Genotypes were determined by predicting the presence of antibiotic resistance genes and mutations described in CARD in the genomes of 73 isolates of human gut bacteria. Phenotypes were determined by zone of inhibition antibiotic susceptibility testing and categorized as resistant or susceptible. This produces four possible genotype/phenotype combinations.

		Presence of genetic resistance determinant (genotype)	
		Yes (Predicted Resistance)	No (Predicted Susceptibility)
Phenotype	Resistant	Confirmed Resistance	Unpredicted Resistance
	Susceptible	Unpredicted Susceptibility	Confirmed Susceptibility

The distribution of these genotype/phenotype combinations among the 73 isolates is visualized in Figure 4.5 against their core genome phylogeny. Each cell represents a genotype/phenotype combination for a particular isolate and antibiotic. Visualising the data in this way shows patterns that can be generalized to each phylum. For example, Confirmed Resistances appear most common in the Proteobacteria and Bacteroidetes, whereas Actinobacteria have the fewest Confirmed Resistances. In contrast, Firmicutes appear to have the most Confirmed Susceptible isolates. In addition, the Proteobacteria isolates have the most Unpredicted Susceptibility genotype/phenotype combinations. All phyla feature Unpredicted Resistances; in some cases, these occur in every isolate of a particular phylum. For example, Proteobacteria and vancomycin; Bacteroidetes and Gentamicin; Firmicutes and ciprofloxacin. These observations will now be explored in more detail.

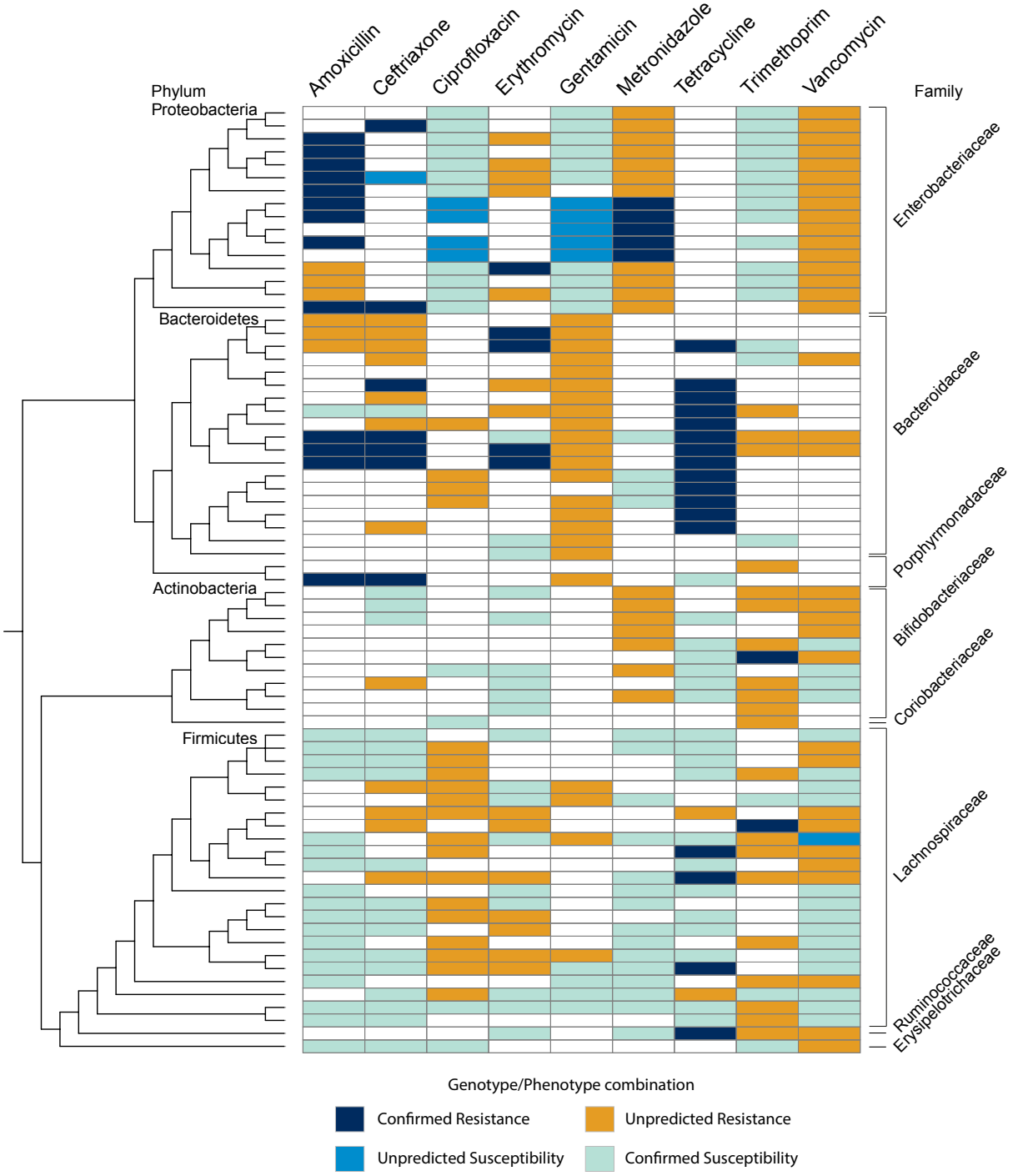


Figure 4.5. Distribution of antibiotic sensitivity genotype/phenotype combinations in 73 phylogenetically diverse isolates of human gut bacteria reveals many "unpredicted" resistances. The phylogeny was inferred from 40 core genes of 73 whole genome sequences of gut microbiota isolated from healthy human faecal samples. Each isolate was screened for antibiotic resistance to 9 antibiotics; in total 657 phenotypes were determined. Each cell in the figure represents an isolate's genotype/phenotype combination for a particular antibiotic. Dark blue = Confirmed Resistance (genetic resistance and phenotypic resistance both observed); Light blue = Confirmed Susceptibility (no genetic or phenotypic resistance); Mid-blue = Unpredicted Susceptibility (genetic resistance predicted, but phenotypically susceptible) and Unpredicted Resistance (no genetic resistance predicted, but phenotypically resistant). White cells represent combinations involving intermediate antibiotic sensitivity.

Examining the proportion of these genotype/phenotype combinations per phylum more closely (Fig. 4.6) showed that Confirmed Resistance is significantly higher in Bacteroidetes compared to the overall set of 73 isolates (q value < 0.0001; p-values determined by Fisher exact tests, adjusted using the Benjamini, Hochberg, and Yekutieli method for q-values, significant when q < 0.05) and Confirmed Susceptibility is significantly higher in Firmicutes (q value < 0.0001). Unpredicted Susceptibility mainly occurs in Proteobacteria, where it was significantly enriched (q value < 0.001), plus a very small amount in Firmicutes. All phyla demonstrate Unpredicted Resistance, but this occurs significantly more so than expected in Bacteroidetes (q value < 0.001). The proportion of Unpredicted Resistance genotype/phenotype combinations can be considered as the rate of False Negatives (i.e., the absence of CARD resistance determinants but phenotypic resistance indicates susceptibility was falsely predicted). In addition, the proportion of Unpredicted Susceptibility can be considered as the rate of False Positives (i.e., the presence of CARD resistance determinants but phenotypic susceptibility indicates resistance was falsely predicted). Thus, the overall False Negative rate for all genotype/phenotype combinations in all isolates was 38.6 %: the highest False Negative rate (Unpredicted Resistance, 51.9 %) occurs in Bacteroidetes, followed by 40.5 % in Actinobacteria, 39.6 % in Proteobacteria and 29.7 % in Firmicutes. The overall False Positive (Unpredicted Susceptibility) rate is 3.2 %: the highest False Positive rate occurs in Proteobacteria (11.0 %) and the lowest is 0.72 % in Firmicutes. There were no False Positives in Actinobacteria or Firmicutes. Therefore, antibiotic resistance seems to be more accurately predicted in some phyla of human gut microbiota than others.

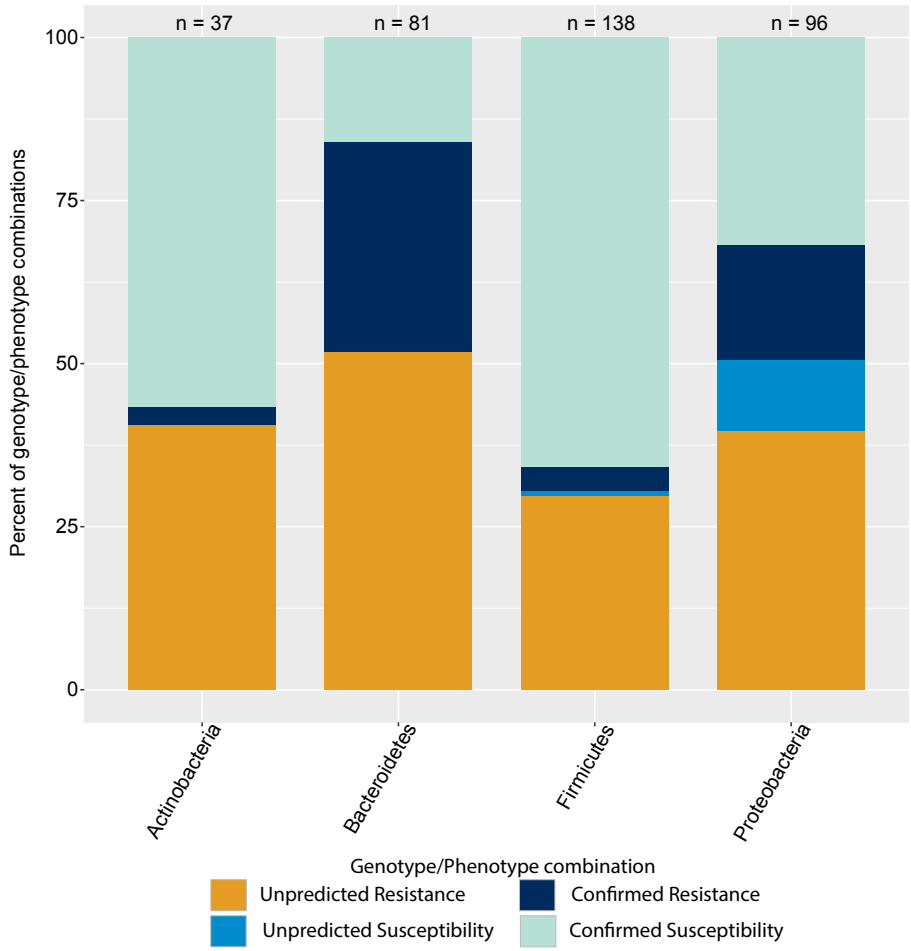


Figure 4.6. The proportion of genotype/phenotype combinations for each phylum. The number of genotype/phenotype combinations overall for all 73 isolates and each phylum was counted: All – 352; Actinobacteria – 37; Bacteroidetes – 81; Firmicutes – 138; Proteobacteria – 96. The proportion of specific combinations e.g. Confirmed Resistance, Confirmed Susceptibility, Unpredicted Resistance and Unpredicted Susceptibility was also determined across all 73 isolates tested and in each phylum. Confirmed Resistance was significantly enriched in Bacteroidetes (q value < 0.0001) compared to the overall Confirmed Resistance rate and was significantly lower in Firmicutes (q < 0.0001). Confirmed Susceptibility was enriched in Firmicutes (q < 0.0001) and occurred significantly less in Bacteroidetes than expected (q < 0.0001). Unpredicted Resistance was found in all phyla, but significantly more in Bacteroidetes (q = 0.0354). Unpredicted Susceptibility significantly occurs in Proteobacteria (q < 0.0001), with a very small amount in Firmicutes. P-values determined by Fisher exact tests, adjusted for multiple-testing using the Benjamini, Hochberg, and Yekutieli method; significant when q < 0.05. Unpredicted Resistance can also be considered a False Negative result and Unpredicted Susceptibility can be considered a False Positive result.

Assessing the proportion of genotype/phenotype for each antibiotic (Fig. 4.7) reveals that resistance is also more accurately predicted for some antibiotics than others. For example, both amoxicillin and tetracycline have significantly more Confirmed Resistances (34.2 % and 44.7 % respectively) than expected compared to the overall rate of Confirmed Resistances

(13.9 %) across all genotype/phenotype combinations (q values 0.035 and < 0.0001 respectively; p-values determined by Fisher exact tests, adjusted using the Benjamini, Hochberg, and Yekutieli method for q-values; significant when q < 0.05). Confirmed Susceptibility was not significantly different for any antibiotic compared to the overall rate (45.2 % of genotype/phenotype combinations across all antibiotics, ranging from 33.3 % to 50.0 % for individual antibiotics). Unpredicted Susceptibility is only observed for ceftriaxone, ciprofloxacin, gentamicin, and vancomycin, although was not significantly enriched in any of these antibiotics. The overall False Positive (Unpredicted Susceptibility) rate was 3.13 % of genotype/phenotype combinations, with the highest False Positive rate occurring for gentamicin (11.9 %), closely followed by ciprofloxacin (10.5 %), then dropping to 2.63 % for vancomycin; False Positives were not observed for amoxicillin erythromycin, metronidazole, tetracycline and trimethoprim. The overall False Negative (Unpredicted Resistance) rate was much higher, at 37.8 % of genotype/phenotype combinations. Whilst no antibiotic was enriched for more Unpredicted Resistances than expected, the False Negative rate was also highest for gentamicin (54.8 %), and similarly high for ciprofloxacin (50.0 %), vancomycin (50.0 %), and trimethoprim (48.8 %). The lowest False Negative rates occurred for amoxicillin (15.8 %) and tetracycline (5.26 %); indeed, False Negatives (Unpredicted Resistances) were observed significantly less often for tetracycline (q value < 0.0001) than expected, based on the overall rate. Overall, that all nine antibiotics have False Negative results (Unpredicted Resistances) further indicates that the isolates examined contain more antibiotic resistance than was predicted using the CARD database of known, clinically relevant antibiotic resistance determinants.

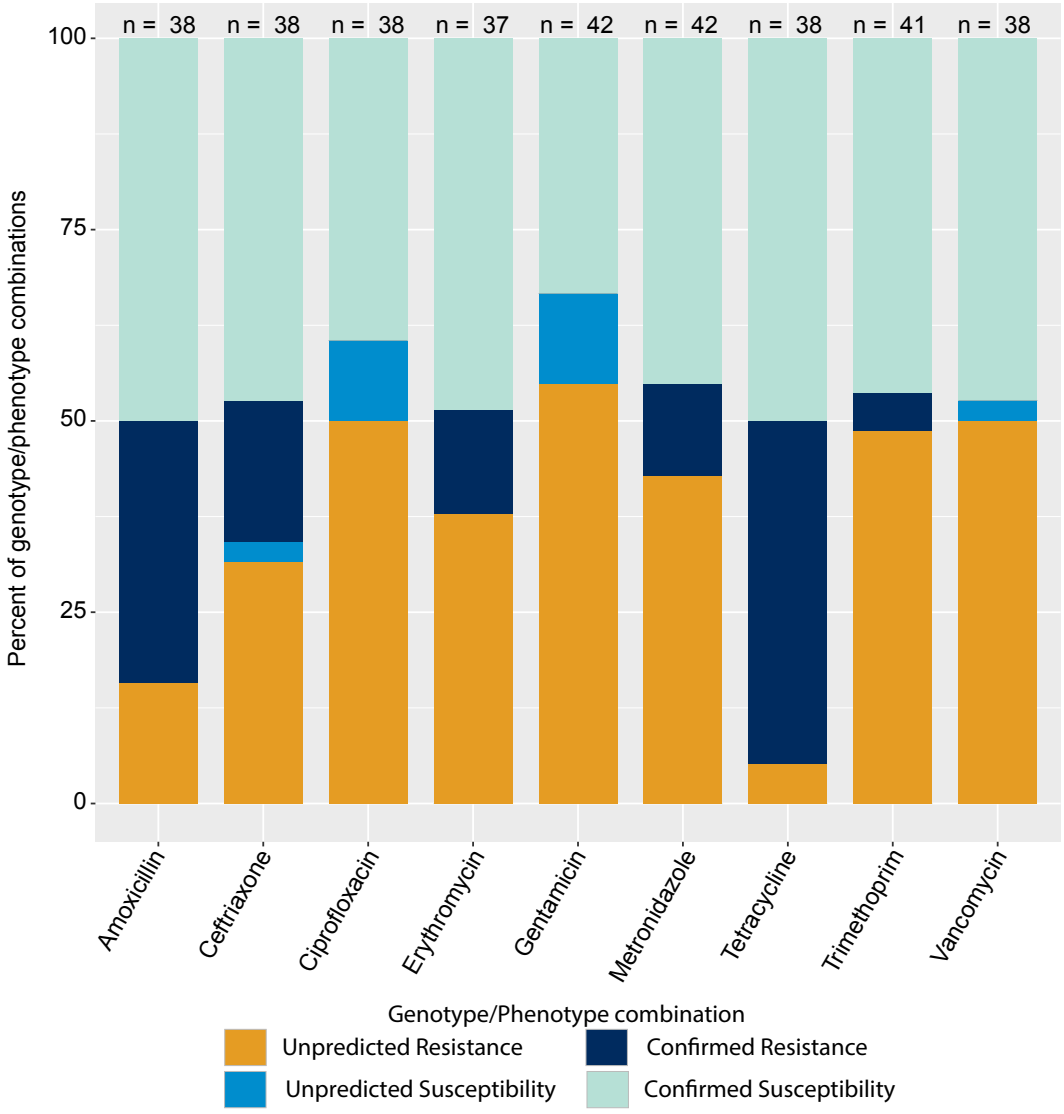


Figure 4.7. The proportion of genotype/phenotype combinations for each antibiotic. The number of genotype/phenotype combinations overall for all antibiotics ("All") and each antibiotic was counted: All – 352; Amoxicillin – 38; Ceftriaxone – 38; Ciprofloxacin– 38; Erythromycin – 37; Gentamicin – 42; Metronidazole – 42; Tetracycline – 38; Trimethoprim – 41; Vancomycin – 38. The proportion of specific combinations e.g. Confirmed Resistance, Confirmed Susceptibility, Unpredicted Resistance and Unpredicted Susceptibility was also determined across all and for each antibiotic. Confirmed Resistance is enriched in amoxicillin and tetracycline (q values = 0.035 and < 0.0001 respectively). Unpredicted resistance was observed significantly fewer times than expected for tetracycline compared to all antibiotics ($q < 0.0001$). P-values determined by Fisher exact tests, adjusted for multiple-testing using the Benjamini, Hochberg, and Yekutieli method; significant when $q < 0.05$. Unpredicted Resistance can also be considered a False Negative result and Unpredicted Susceptibility can be considered a False Positive result.

4.2.5 Comparison of antibiotic resistance databases and prediction methods

Having identified that the ARIBA with CARD rule-based method is not completely accurate, other databases and methods were applied to the 73 isolates to provide alternative predictions of resistance genotypes (Fig. 4.8): the CARD's own Resistance Gene Identifier tool (CARD-RGI) and ARIBA with the MegaRes, ResFinder, and SRST2-ARGANNOT databases were used. I then compared the newly generated resistance genotypes to the phenotypic data generated for the nine antibiotics to determine proportions of each genotype/phenotype combination (Confirmed Resistance, Confirmed Susceptibility, Unpredicted Resistance and Unpredicted Susceptibility) for each method. The proportion of each combination in the four new methods was compared to the proportion from the initial ARIBA with CARD analysis performed in this thesis. This shows that the ResFinder and SRST2-ARGANNOT databases had significantly higher proportions of Confirmed Susceptibility, but no significant differences in the other three genotype/phenotype combinations. All databases perform similarly in terms of the rate of False Negatives (Unpredicted Resistances), though it was highest when using ARIBA with the MegaRes database (45.5 %, compared to 45.1 % for ARIBA with CARD, 39.8 % for CARD-RGI, 39.2 % for ARIBA with ResFinder, and 38.9 % for ARIBA with SRST2-ARGANNOT). Despite a significant increase in Confirmed Resistance genotype/phenotype combinations with CARD-RGI, this tool also had a higher rate of False Positives (Unpredicted Susceptibility; 8.5 % compared to 3.1 % for ARIBA with CARD, 2.8 % for ARIBA with MegaRes, and 0.9 % for ARIBA with SRST2-ARGANNOT). Therefore, CARD-RGI in particular seems to overpredict resistance using genomic data (predicting resistance when the isolate is susceptible). For these reasons, the original ARIBA with CARD results are used for subsequent analyses.

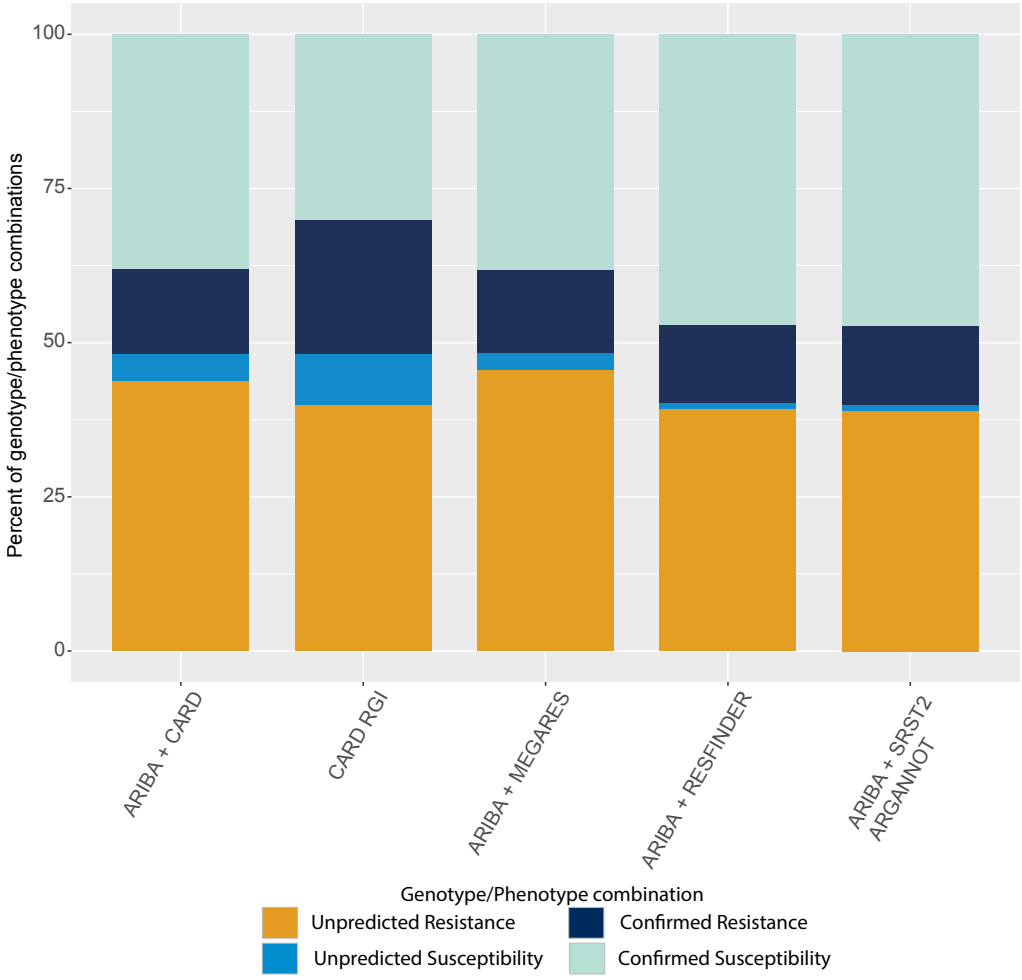


Figure 4.8. The proportion of genotype/phenotype combinations for each resistance database or method tested. The proportion of specific combinations e.g. Confirmed Resistance, Confirmed Susceptibility, Unpredicted Resistance and Unpredicted Susceptibility was determined for each database and method (total combinations = 352 for each method). These proportions were compared to the original ARIBA+CARD proportions using a two-proportion z test with Yates correction (q value significant when $q < 0.05$). The CARD Resistance Gene Identifier (CARD-RGI) method had significantly lower proportion of Confirmed Susceptibility ($q = 0.0385$) and significantly higher proportions of Confirmed Resistance and Unpredicted Susceptibility (q values 0.0103 and 0.0377 respectively). ARIBA with the RESFINDER and SRST2-ARGANNOT databases both had significantly higher proportions of Confirmed Susceptibility than ARIBA+CARD ($q = 0.0147$ and 0.0119 respectively). Unpredicted Resistance can also be considered a False Negative result and Unpredicted Susceptibility can be considered a False Positive result.

4.2.6 Identifying enrichment of unpredicted resistance to certain antibiotics in particular phyla

Determining which phyla have Unpredicted Resistances to certain antibiotics (Fig. 4.9) shows that Actinobacteria and Proteobacteria only have Unpredicted Resistance against three and four antibiotics respectively, whereas Firmicutes and Bacteroidetes demonstrate Unpredicted Resistance for six and eight antibiotics respectively. In particular, Actinobacteria were enriched in unpredicted metronidazole and trimethoprim resistance (q values 0.027 and 0.042 respectively; p-values determined by Fisher exact tests, adjusted using the Benjamini, Hochberg, and Yekutieli method for q-values, significant when $q < 0.05$). Bacteroidetes were especially enriched in unpredicted gentamicin resistance ($q < 0.0001$). Firmicutes were enriched in unpredicted ciprofloxacin resistance (q value 0.014), but fewer unpredicted metronidazole and vancomycin resistances were observed than expected (q values both 0.005). Proteobacteria were not enriched for any unpredicted resistances. Therefore, each phylum has different profiles of Unpredicted Resistances. In addition, this data supports the previous findings that Proteobacteria appears to have resistance more accurately predicted, and that unpredicted resistance overall was more common in non-Proteobacteria. However, as we previously saw that the Proteobacteria have large proportions of Unpredicted Susceptibility (Fig. 4.6), it appears that resistance may also be overpredicted in these isolates.

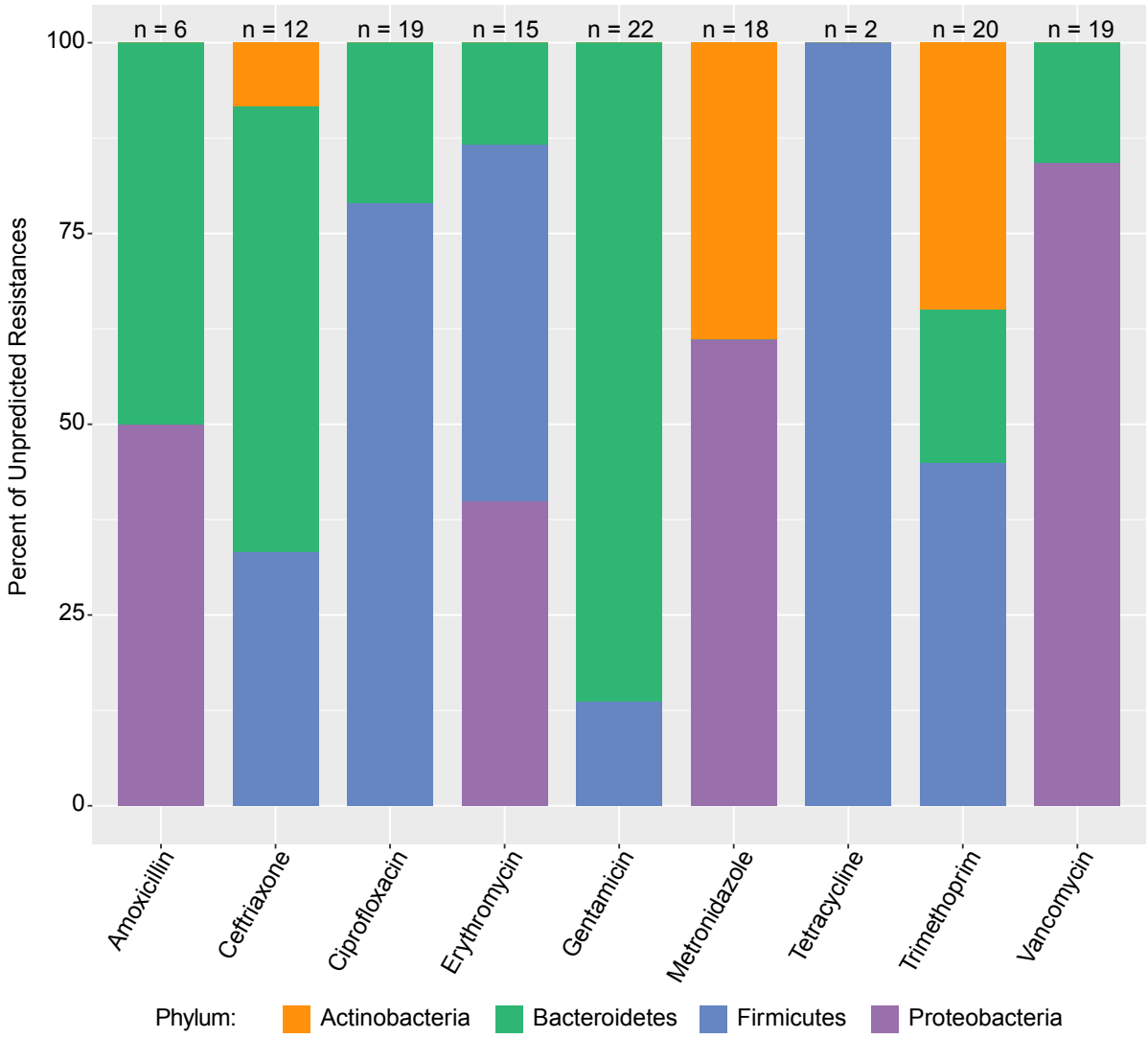


Figure 4.9. The percentage of Unpredicted Resistance antibiotic genotype/phenotype combinations by which phyla those combinations were observed in. Proteobacteria and Actinobacteria only demonstrate unpredicted resistance for three antibiotics, whereas Bacteroidetes and Firmicutes demonstrate unpredicted resistance for eight and six antibiotics respectively. Actinobacteria were enriched for Unpredicted Resistance to Metronidazole and Trimethoprim ($q = 0.027$ and 0.042 respectively). Bacteroidetes were enriched overall for Unpredicted Resistance ($q = 0.042$) but especially gentamicin Unpredicted Resistance ($q < 0.0001$). Firmicutes were enriched for ciprofloxacin Unpredicted Resistance ($q = 0.014$) but significantly fewer metronidazole and vancomycin Unpredicted Resistances were observed compared to the complete dataset (q values both 0.005). Proteobacteria were not enriched for any Unpredicted Resistances; significantly fewer Unpredicted Resistances were observed for ciprofloxacin ($q = 0.042$), gentamicin ($q = 0.027$), metronidazole ($q = 0.042$), trimethoprim ($q = 0.042$) and vancomycin ($q < 0.0001$). P-values determined by Fisher exact tests, adjusted for multiple-testing using the Benjamini, Hochberg, and Yekutieli method; significant when $q < 0.05$.

Considering the data so far, it is clear that the ARIBA + CARD method used here to initially predict antibiotic resistance in these human commensal gut bacteria is not accurate. In addition, the isolates studied contain more antibiotic resistance than is predicted using a rule-based method with the CARD database of known, clinically relevant antibiotic resistance determinants. In particular, the Bacteroidetes isolates were enriched for False Negatives (Unpredicted Resistance) and Proteobacteria isolates were enriched for False Positives (Unpredicted Susceptibility).

4.2.7 Further investigations of unpredicted resistance

Unpredicted Resistance genotype/phenotype combinations represent instances of a mismatch between two important methods for determining antibiotic sensitivity (culture-based- and WGS-AST). Some of these unpredicted resistance observations are likely explained by intrinsic resistance; for example, where they occur in all isolates of a particular phyla, including gentamicin in the Bacteroidetes isolates studied or vancomycin in the Proteobacteria (Fig. 4.5). However, where these unpredicted resistances occur in isolates that are closely related to isolates susceptible to that same antibiotic, these may be explained by genetic resistance determinants that are novel, or not described in CARD. This offers the opportunity to look for candidate novel antibiotic resistance genes or mutations in isolates with unpredicted resistance. In particular, there are instances of unpredicted ceftriaxone resistances observed in Bacteroidetes and Firmicutes where closely related isolates demonstrate Confirmed Susceptibility or Confirmed Resistance. As ceftriaxone resistance is often mediated by beta-lactamase enzymes these examples may indicate the presence of novel beta-lactamases in these human commensal gut microbiota.

Two particular isolates appeared good targets for novel beta-lactamases: *Bacteroides faecis* 18048_2#66 and *Lachnospiraceae nov.* 20287_6#18. These isolates both had ceftriaxone zone of inhibition sizes of 0 mm, indicating complete resistance to this beta-lactam antibiotic. They were also in the top five most ceftriaxone-resistant isolates (Fig. 4.10), but were the only two of those five without genetic determinants of beta-lactam resistance in their genomes.

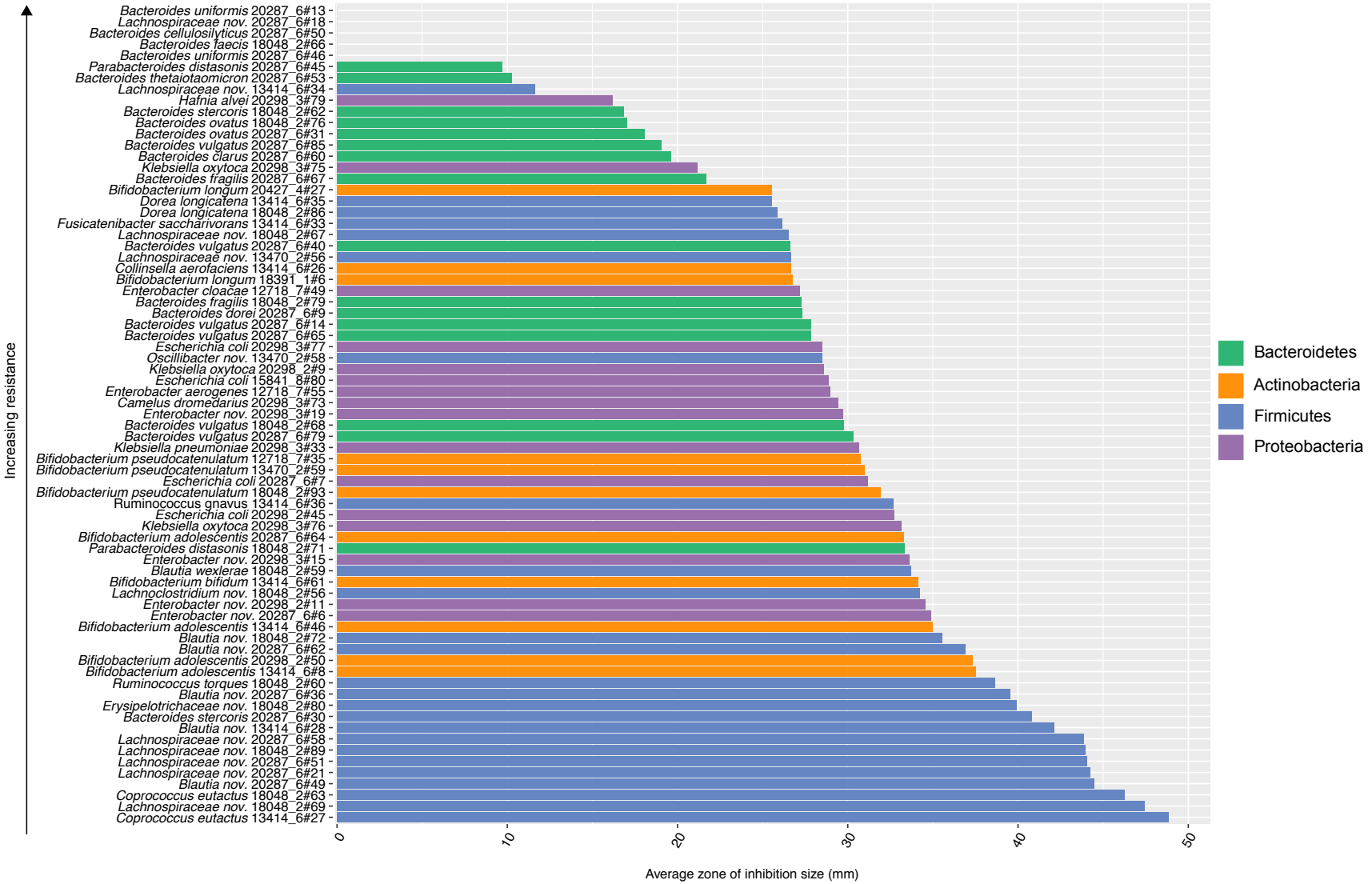


Figure 4.10. Ranking of isolates by ceftriaxone sensitivity. The average zone of inhibition sizes was used to rank the 73 HBC isolates by ceftriaxone sensitivity (bottom = most sensitive, largest zone of inhibition; top = least sensitive, smallest zone of inhibition). Five isolates were completely resistant to ceftriaxone (no zone at all).

4.2.8 Searching for novel antibiotic resistance determinants in human gut commensal microbiota

A comparative phenotyping and genomics approach was used to identify candidate novel resistance genes from *Bacteroides faecis* 18048_2#66 and *Lachnospiraceae nov.* 20287_6#18. I determined the average nucleotide identity (ANI), a measure of genomic similarity between the coding regions of two genomes, for each isolate under investigation and its closest two relatives from the HBC (Fig. 4.11). The two related *B. faecis* isolates, *B. faecis* 13470_2#65 and *B. faecis* 12718_7#26 were both more than 99 % similar by ANI to *B. faecis* 18048_2#66. *Lachnospiraceae nov.* 20287_6#18 had a very close relative of 99.24 % ANI, *Lachnospiraceae nov.* 8080_1#94, but the second next most closely related isolate from the HBC (*Coprococcus nov*. 20298_3#65) was only 81 % similar by ANI.

I determined the average Minimum Inhibitory Concentration (MIC) for ceftriaxone for each of the six isolates using Biomerieux Etests (antibiotic gradient strips) across three biological replicates (Figure 4.11). The average MIC for the two isolates with Unpredicted Resistance to ceftriaxone, *B. faecis* 18048_2#66 and *Lachnospiraceae nov.* 20287_6#18, was at least 256 µg/ml. The maximum concentration of ceftriaxone in the Etest strip was 256 µg/ml; these results mean that those two isolates were completely resistant to ceftriaxone at the maximum concentration tested and so the MIC is greater than or equal to 256 µg/ml. For *B. faecis* 18048_2#66, both related isolates *B. faecis* 13470_2#65 and *B. faecis* 12718_7#26 were more sensitive to ceftriaxone, with MICs under 60 µg/ml. The *Lachnospiraceae nov.* 8080_1#94 isolate shared the same phenotype as *Lachnospiraceae nov.* 20287_6#18 with an MIC of at least 256 µg/ml, and the *Coprococcus nov*. 20298_3#65 was much more sensitive with an MIC of just 0.5 µg/ml.

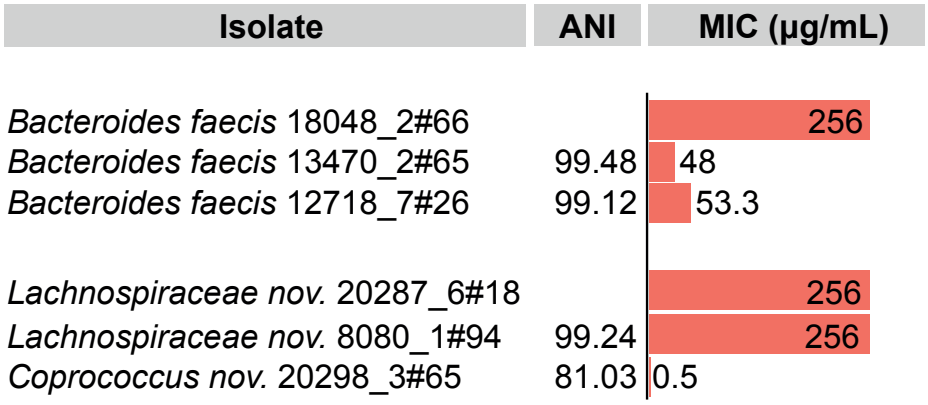


Figure 4.11. Comparison of ceftriaxone sensitivity in two sets of isolates from the HBC. *Bacteroides faecis* 18048_2#66 and *Lachnospiraceae nov.* 20287_6#18 were identified as having unpredicted ceftriaxone resistance. Their two closest relatives each were identified from the HBC using a 40 core gene phylogeny and Average Nucleotide Identity (ANI) was determined for the two closest relatives compared to the isolate in which unpredicted resistance was observed. The ceftriaxone Minimum Inhibitory Concentration (MIC, µg/ml) was determined for each isolate.

This data offers the opportunity to identify genomic differences between those isolates that may explain phenotypic differences. The *Lachnospiraceae nov.* 20287_6#18 isolate will be excluded since its most closely related isolate shared a ceftriaxone resistant phenotype and the ceftriaxone-susceptible *Coprococcus nov*. 20298_3#65 is a different species making genomic identification impractical. The *B. faecis* isolates, however, represent an ideal situation of very closely related isolates with differing phenotypic ceftriaxone sensitivity. In the rest of this section, I will investigate genomic differences between these *B. faecis* isolates to identify potential candidate beta-lactamase genes or mutations that may confer the unpredicted ceftriaxone resistance observed in *B. faecis* 18048_2#66.

Core genome analysis was performed using Roary[201] on the three *B. faecis* isolates and identified 3652 genes shared by all three isolates ("shared core genes") and 614 genes unique to the ceftriaxone resistant *B. faecis* 18048_2#66 ("resistant-unique genes", absent from the two more sensitive *B. faecis* isolates). ShortBRED[209] was used to reduce the amino acid sequences of 235,009 proteins containing the phrase "beta-lactamase" in their name from the

NCBI Protein database as of July 2018) into a database of reference amino acid markers. This includes beta-lactamase regulatory proteins so can possibly account for regulatory mutations too. These markers were used to search the translated amino acid sequences of the 614 resistant-unique genes and the 3652 shared core genes with 90 % identity. Other cut offs were tested (Table 4.4); however for subsequent analyses the 90 % cut off was used as a high level of similarity is typically required to infer functionality. Seven of the shared core genes were found to have amino acid sequences 90 % similar to reference amino acid markers from beta-lactamases (Table 4.5). A single gene labelled "Group 2384" was annotated as a candidate beta-lactamase.

Table 4.4. Numbers of resistant-unique and shared core genes with similarity to beta-lactamase markers in the human gut bacteria isolate *Bacteroides faecis* 18048_2#66. ShortBRED[194] was used to reduce the amino acid sequences of 235,009 proteins containing the phrase "beta-lactamase" in their name from NCBI Protein database (as of July 2018) into a database of reference amino acid markers. These markers were used to search the translated amino acid sequences of the 614 resistant-unique genes and the 3652 shared core genes with a variety of similarity cut offs.

Identity cut off (%)	No. of resistant-unique genes matched	No. of shared core genes matched
90	1	7
80	1	15
70	1	20
60	2	30
50	3	52
25	23	244

Table 4.5. A summary of candidate beta-lactamases that may explain an unpredicted ceftriaxone resistance phenotype observed in the human gut bacteria isolate *Bacteroides faecis* 18048_2#66. This isolate was phenotypically resistant to ceftriaxone in the absence of genetic determinants of beta-lactam resistance described in CARD. The closest two relatives from the HBC were identified and also phenotyped; they were both more sensitive to ceftriaxone. Roary core genome analysis was performed to identify genes unique to the resistant isolate ("resistant unique") and genes shared by all three isolates ("shared core"). These genes were searched for sequences with 90 % similarity to amino acid markers derived from 230,009 beta-lactamase related proteins in the NCBI Protein database. The table describes the genes that were found to contain markers of these proteins, what the genes were annotated as by Roary and the protein that the observed marker is derived from.

Category	Gene	Annotation	NCBI Beta-lactamase marker hit
Resistant unique	*Group 2384*	Beta-lactamase domain-containing protein	WP004329300 MULTISPECIES: MBL fold metallo-hydrolase [Bacteroidales]
Shared core	*ampG1*	Major Facilitator Superfamily	NP812531 AmpG protein, beta-lactamase induction signal transducer [*Bacteroides thetaiotaomicron* VPI-5482]
Shared core	*ampG2*	Signal transducer	NP809947 signal transducer [*Bacteroides thetaiotaomicron* VPI-5482]
Shared core	*blaR1*	Transcriptional regulator	WP010538315 MULTISPECIES: M56 family metallopeptidase [*Bacteroides*]
Shared core	*Group 106*	TonB	WP062695069 M56 family peptidase [*Bacteroides thetaiotaomicron*]
Shared core	*Group 3492*	Protein of unknown function (DUF2874)	WP062695288 hypothetical protein [*Bacteroides thetaiotaomicron*]; Putative beta-lactamase-inhibitor-like, PepSY-like; pfam11396
Shared core	*Group 4547*	Protein of unknown function (DUF2874)	WP008766859 hypothetical protein [*Bacteroides thetaiotaomicron*]; Putative beta-lactamase-inhibitor-like, PepSY-like; pfam11396
Shared core	*Group 6146*	Putative exported beta-lactamase protein	WP008769828 DUF302 domain-containing protein [*Bacteroides fragilis*]; Beta-lactamase; pfam00144

Subsequently, I investigated whether any of these candidate beta-lactamases might explain the unpredicted ceftriaxone resistance in *B. faecis* 18048_2#66. Firstly, I determined whether the candidate beta-lactamase gene Group 2384 unique to the resistant isolates corresponded with increased ceftriaxone MIC. To do this, I looked for the presence of Group 2384 in the complete set of HBC genomes with 100 % sequence length and nucleotide identity. This gene was identified in 16 other HBC isolates, all in the Bacteroidetes phylum. Five of these Bacteroidetes isolates were excluded from further analysis due to the presence of other genetic beta-lactam resistance determinants (identified in the analyses discussed in Chapter 3). For the 11 remaining Bacteroidetes isolates, I measured the ceftriaxone MICs for their closest relatives in the HBC (Fig. 4.12). In theory, if Group 2384 was responsible for the ceftriaxone-resistance phenotype in *B. faecis* 18048_2#66 and potentially other Bacteroidetes isolates, I would expect the presence of Group 2384 in an isolate's genome to correspond with a higher ceftriaxone MIC. This correlation was not observed, suggesting that this candidate beta-lactamase may not be responsible for this phenotype, or that it is not functional in the other eleven Bacteroidetes isolates.

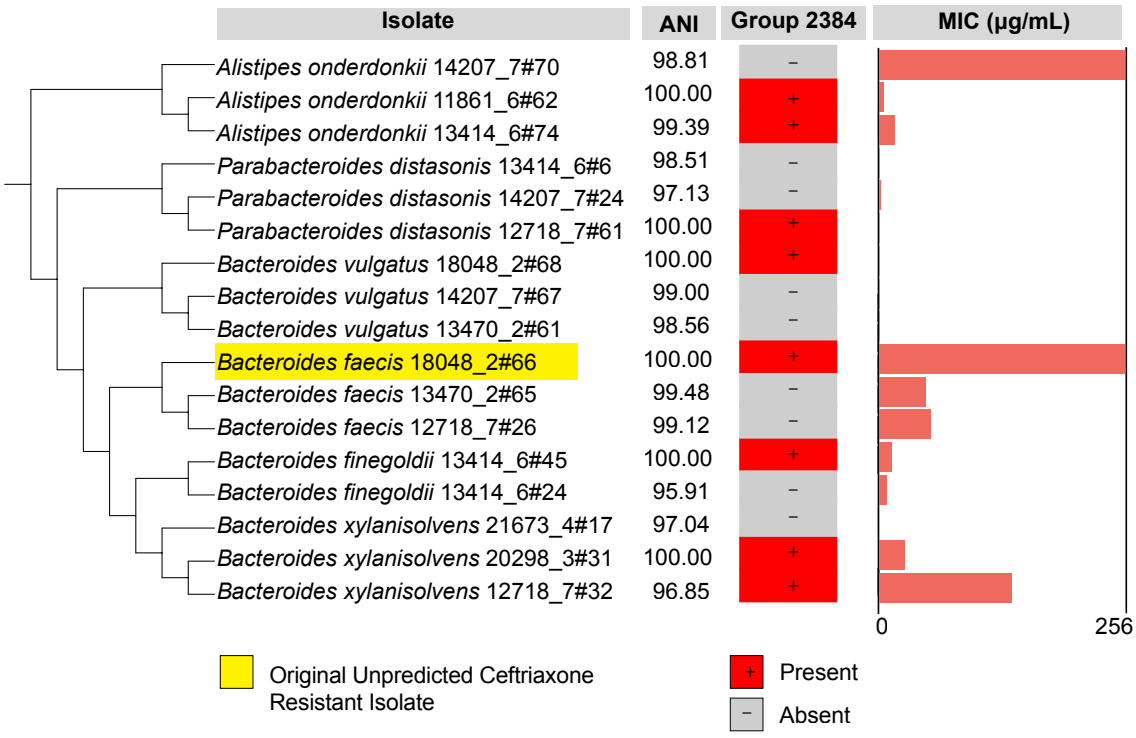


Figure 4.12. Ceftriaxone MIC in isolates with and without the Group 2384 candidate beta-lactamase gene. Group 2384 is a candidate beta-lactamase first identified in the Bacteroides faecis 18048_2#66, an isolate with unpredicted ceftriaxone resistance (highlighted in yellow). The presence of the Group 2384 gene was searched for in the HBC (100 % sequence length and ID) and was identified in 16 isolates. 11 Group 2384-positive isolates without any other predicted beta-lactam resistance (determined using CARD) plus each of their two closest relatives from the HBC were selected. ANI analysis was used to determine how similar the Group2384-negative isolates were to their closest Group 2384-positive relative. Ceftriaxone MICs were measured using Etests in three biological replicates of ach isolate. The presence of Group2384 was not correlated with a lower MIC for amoxicillin and ceftriaxone.

Although the presence of Group 2384 did not correlate with increased ceftriaxone MIC and thus resistance in other HBC isolates, this does not necessarily rule out its function as a ceftriaxone-resistance gene in *B. faecis* 18048_2#66. It is possible that it is not expressed in the other isolates or contains mutations that leave it non-functional. To rule out the latter hypothesis, I extracted and aligned the Group 2384 sequences from the twelve Group 2384-positive isolates I tested for phenotypic ceftriaxone sensitivity to infer a phylogenetic tree (Fig. 4.13). Whilst the Group 2384 genes were not identical, very few mutations were identified: *Bacteroides vulgatus* 18048 2#68 has base T at position 545 where the other eleven isolates have an A, seven of the Group 2384-positive isolates (including *B. faecis* 18048_2#66) have

base C at gene position 15, whereas the other five Group 2384-positive isolates have base T (Fig. 4.13). These mutations were all synonymous and did not alter the amino acid sequence. Therefore, it is unlikely that these mutations had any direct impact on the hypothetical function of Group 2384 as a ceftriaxone beta-lactamase, especially since six of the eleven Group 2384 sequences were identical to the one from *B. faecis* 18048_2#66.

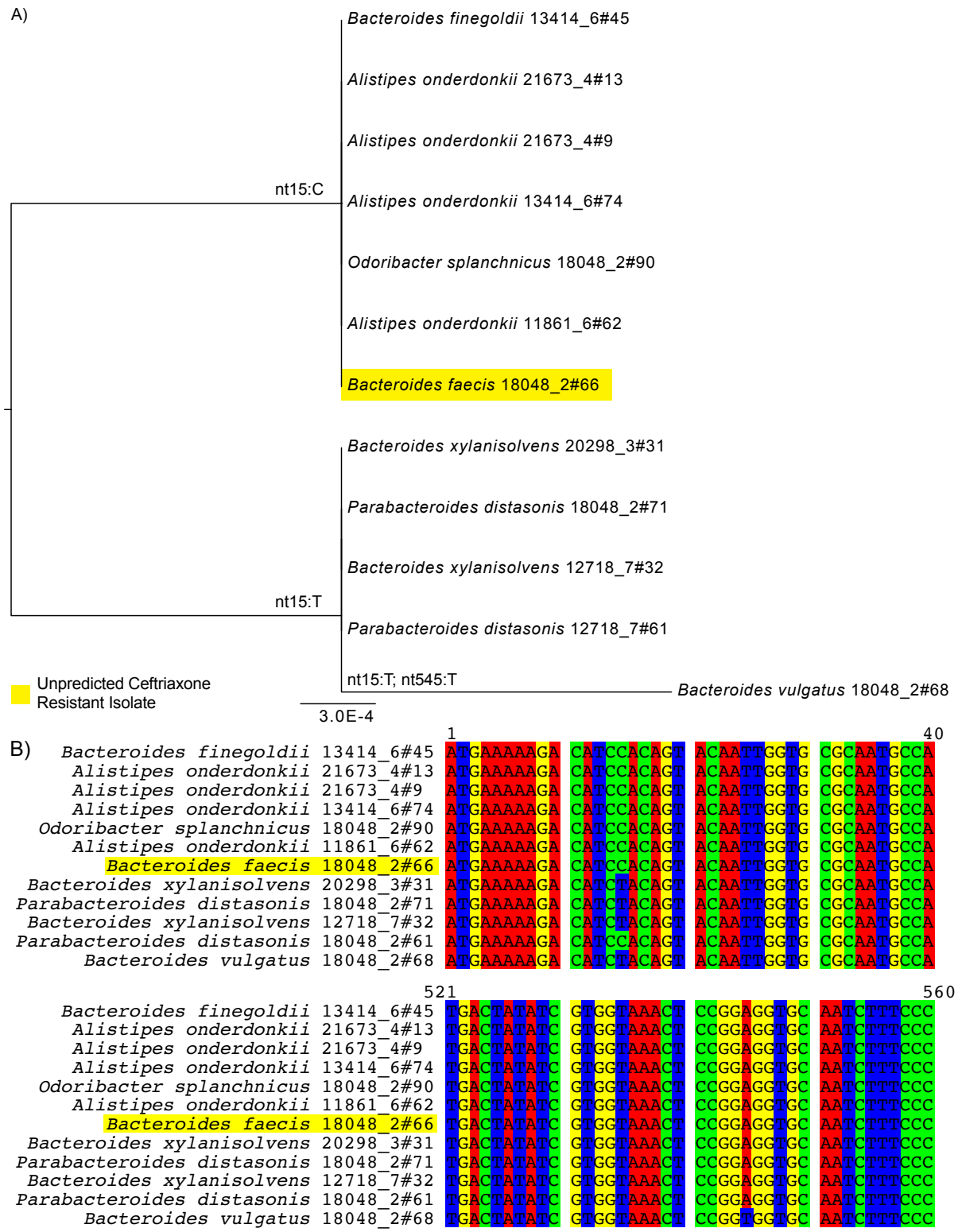


Figure 4.13. Group 2384 gene sequences from twelve HBC isolates. Group 2384 is a candidate beta-lactamase first identified in the *Bacteroides faecis* 18048_2#66, an isolate with unpredicted ceftriaxone resistance (highlighted in yellow). A: the nucleotide sequence for the Group 2384 gene was extracted from 11 other HBC isolates in which it was identified, aligned and used to infer a phylogeny and identify mutations (labelled at branch points). B: the alignment of nucleotides 1-40 and 521-560 are shown to illustrate the single nucleotide polymorphisms in the Group 2384 gene sequences.

To further investigate the potential function of Group 2384 as a ceftriaxone resistance-conferring beta-lactamase I performed a gain-of-function cloning experiment. A plasmid carrying a chloramphenicol resistance marker gene and the Group 2384 gene was designed with GeneArt (ThermoFisherScientific; Fig. 4.14). The Group 2384 gene was inserted within a tetracycline resistance gene, under the control of the tetracycline resistance gene promoter. The synthesised construct was transformed into electrocompetent *E. coli* cells via electroporation with the assistance of Mr Matthew Dorman. The transformed cells were grown on LB agar plates containing chloramphenicol (12.5 µg/mL) to check that the vector had been taken up by the *E. coli* cells. Transformed cells were also plated on LB agar containing chloramphenicol (12.5 µg/mL) and ceftriaxone at a concentration of 256 µg/mL (representing the observed *B. faecis* 18048_2#66 phenotype) and 4 µg/mL (slightly above the MIC of the untransformed *E. coli*). Whilst colonies were observed on the chloramphenicol control plates, no colonies were observed in the presence of either concentration of ceftriaxone. This suggests the Group 2384 gene may not infer ceftriaxone resistance and explain the unpredicted phenotype in *B. faecis* 18048_2#66.

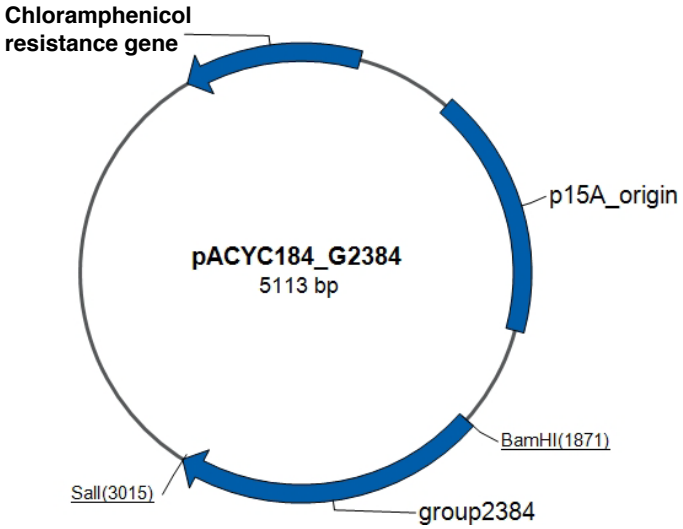


Figure 4.14. GeneArt construct containing Group 2384 candidate beta-lactamase gene. The plasmid pACYC184 was used as the cloning vector, designed and synthesised using GeneArt (ThermoFisherScientific) to contain the gene sequence of the Group 2384 candidate beta-lactamase identified from *Bacteroides faecis* 18048_2#66. The Group 2384 gene was inserted within a tetracycline resistance gene, under the control of the tetracycline resistance gene promoter. Mr Matthew Dorman assisted with the design of the plasmid construct.

It is possible that a mutation in a shared candidate beta-lactamase in the *B. faecis* isolates could cause the unpredicted ceftriaxone resistance phenotype. The gene sequence for each of the seven shared candidate beta-lactamases were extracted from the three *B. faecis* isolates and used to infer phylogenetic trees for each gene from the nucleotide sequences (Fig. 4.15). I looked for non-synonymous mutations that would alter the amino acid sequence of the protein product and could result in altered activity between the resistant *B. faecis* 18048_2#66 and more sensitive *B. faecis* 12718_7#26 and *B. faecis* 13470_2#65. Amino acid substitutions were identified in five of seven shared candidate beta-lactamases; in only one of these, "*ampG2*", was there a substitution unique to resistant *B. faecis* 18048_2#66. AmpG2 is a putative transporter related to AmpG, which possibly transports signal molecules into *P. aeruginosa* cells for the induction of the *ampC* beta-lactamase[244]. AmpG2 may function in a similar way; a non-synonymous mutation in this gene could potentially explain the phenotypic differences in these three isolates; however, it does not tell us exactly which beta-lactamase is degrading the ceftriaxone and amoxicillin antibiotics in *B. faecis* 18048_2#66. It could be controlling Group 2384, but further experiments would be required to investigate this.

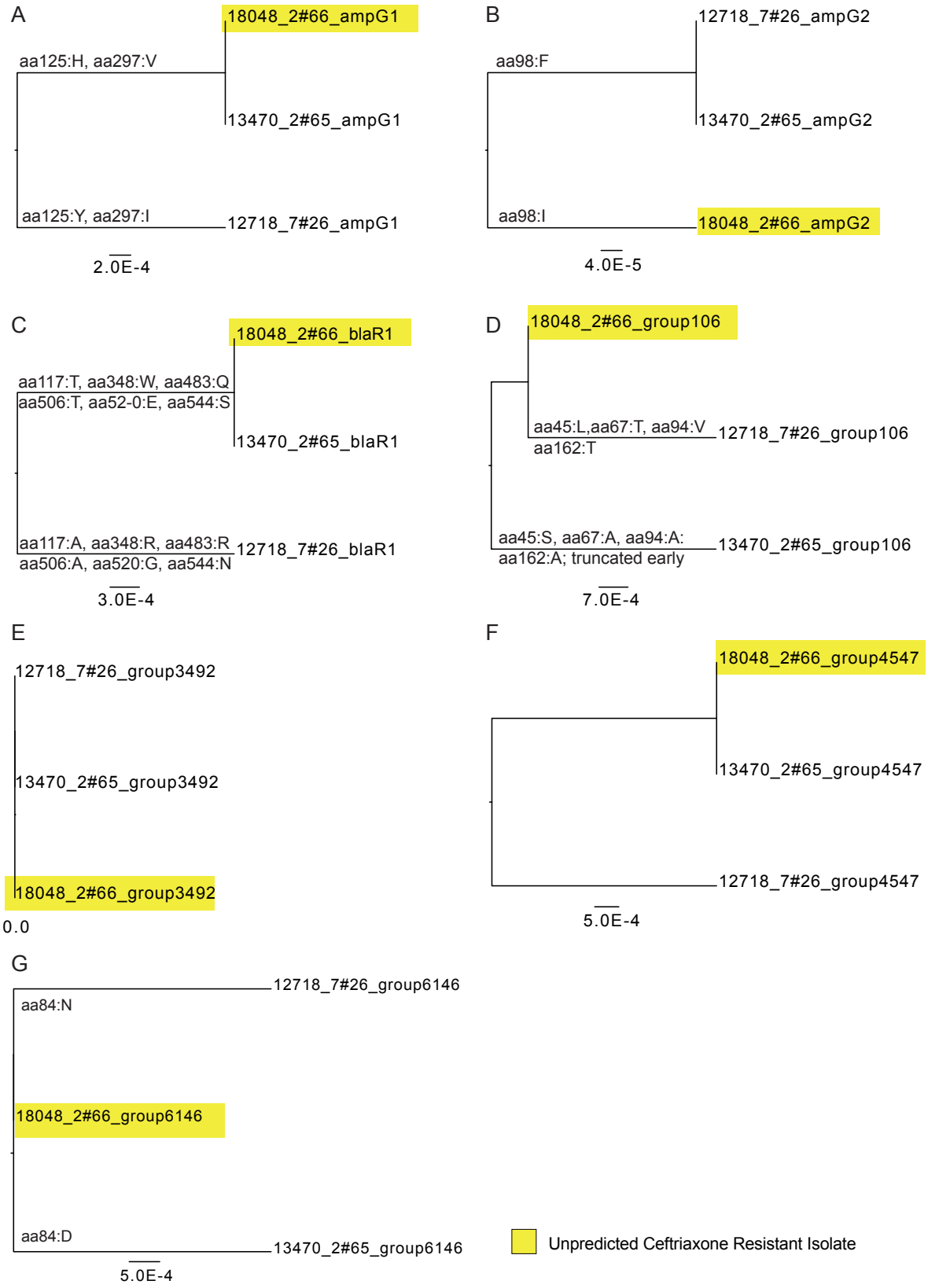


Figure 4.15. Gene phylogenies of candidate shared beta-lactamases in three HBC *Bacteroides faecis* isolates. Seven genes (A-G) shared between three isolates of *B. faecis* from the HBC were found to contain markers of beta-lactamase related proteins from the NCBI Protein database. One of the isolates was observed to be resistant to ceftriaxone in the absence of any known clinically relevant beta-lactam resistance determinant. Amino acid substitutions were observed in five of the seven genes; only in one gene was the mutation unique to the resistant isolate (B, ampG2).

Following these experiments, I applied a shotgun cloning approach to look for novel beta-lactam resistance genes from the whole genomes of *B. faecis* 18048_2#66 and *Lachnospiraceae nov.* 20287_6#18. Both isolates were regrown from HBC glycerol stocks kept at -80 °C, single colonies purified and grown in culture overnight. The full length 16S rRNA sequences were amplified from these cultures using 7f and 1510r PCR primers and sequenced by Sanger sequencing at Eurofins Scientific (Germany). The 16S rRNA sequences were then checked against the whole genome sequence data for these isolates to confirm their identity and check for contamination. Following this quality control, genomic DNA was extracted from the culture pellets of each isolate using phenol:chloroform extraction method by Mr Mark Stares. These genomic samples were used as starting material for the Copy Control Fosmid Cloning Kit (Lucigen) to clone 25-40 kb fragments into *E. coli* using fosmid vectors and Lambda phages. I determined the ceftriaxone MIC of the recipient *E. coli* strain using Etests as before.

The Copy Control Fosmid Cloning method was applied to each isolate individually and therefore represents shotgun cloning from purified isolates rather than mixed samples. This enables identification of the host of any novel antibiotic resistance genes discovered in these experiments. The kit includes control DNA of 40kb fragments, which was included as quality control alongside shotgun cloning of the two isolates with unpredicted ceftriaxone resistance. The fosmid vector carries a chloramphenicol resistance marker gene; the transformed cells were grown on LB agar plates containing chloramphenicol to check that the vector had been taken up by the *E. coli* cells. The transformed *E. coli* cells were also grown on LB agar plates containing ceftriaxone at a concentration of 256 µg/mL (representing the observed *B. faecis* 18048_2#66 phenotype) and 4 µg/mL (slightly above the MIC of the untransformed *E. coli*).

This procedure was repeated four times. Each time, hundreds of colonies were observed on the LB plates with chloramphenicol from the *E. coli* transformed with the control insert DNA.

However, only one or two colonies were observed from *E. coli* transformed with DNA from *Bacteroides faecis* 18048_2#66 or *Lachnospiraceae nov.* 20287_6#18 on the LB plates with chloramphenicol and no colonies on plates containing chloramphenicol and ceftriaxone.

4.3 Discussion

In this chapter I have determined the susceptibility or resistance of 73 phylogenetically diverse human commensal gut bacteria isolates against nine commonly used, clinically relevant antibiotics from the WHO list of essential medicines[21]. This phenotypic data helps to determine a comprehensive view of the impact of commonly used antibiotics across the diversity of gut microbiota. This offers insights into the spectrum of antibiotics which may be used to inform healthcare practices – such as which antibiotics to prescribe for infections caused by opportunistic pathogens from the gut. The main antibiotic from this study that would be useful to specifically target opportunistic anaerobic or gut bacteria is metronidazole, as most Bacteroidetes and Firmicutes tested were generally sensitive to this antibiotic, although this increases the likelihood of impacting more members of the gut microbiota.

In addition to knowing which antibiotics to use in the case of infections by opportunistic gut bacteria, this phenotypic data can advise on which antibiotics to avoid if trying to minimise the impact on commensal gut microbiota. For example, ceftriaxone is a broad-spectrum antibiotic[245] and therefore is useful for treating infections of unknown cause or with resistance to narrower spectrum antibiotics. Although typically administered via injection, not orally, ceftriaxone is known to have an impact on the commensal gut microbiota[246] and here I have showed that members of all four key gut microbiota could be affected. On the other hand, amoxicillin is also considered relatively broad spectrum, but here the only isolates that were

sensitive were Firmicutes, and one Bacteroidetes. Amoxicillin is one of the most commonly prescribed drugs in the world and typically administered orally for e.g. ear infections or throat infections. Amoxicillin also showed high levels of Confirmed Resistance; therefore, this antibiotic should also be used with caution as it may target Firmicutes, some of our most important gut microbiota. Moreover, there are already relatively high levels of amoxicillin resistance caused by known genetic determinants in commensal Proteobacteria that could potentially be transferred to pathogenic bacteria, especially under the selective pressure of amoxicillin therapy.

Another example of note is that of gentamicin: gentamicin is considered a broad-spectrum antibiotic, but one that does not work on anaerobes and streptococci. This is because gentamicin relies on oxygen-dependent transport into bacteria cells[24]. The results in Figure 4.5 demonstrate that this appears true for anaerobic Bacteroidetes, which were largely considered resistant to gentamicin. However, several Firmicute isolates were considered intermediate sensitivity or even susceptible, despite Lachnospiraceae (making up the majority of the Firmicutes isolates screened here) reported to be obligate anaerobes[247]. Therefore, gentamicin may have a more extensive impact on commensal gut microbiota than previously realized. Interestingly, there were no antibiotics that both Bacteroidetes and Firmicutes were both generally resistant to: all the antibiotics tested here have the potential to cause harm to common commensal gut microbiota.

The antibiotic resistance phenotypes were compared to the predicted resistance profiles based on the presence of genetic resistance determinants described in CARD, as determined in the previous chapter. The observations in this chapter follow a rule-based method of predicting antibiotic sensitivity phenotypes from antibiotic resistance genotypes. Rule-based methods have been found to be accurate for predicting antibiotic resistance in several species

of bacterial pathogens[240,242,248]. However, to my knowledge they have not been tested for accuracy in human gut commensal bacteria. I created a system to define the combined data as Confirmed Susceptibility, Confirmed Resistance, intermediate sensitivity (with or without genetic resistance determinants present), Unpredicted Susceptibility and Unpredicted Resistance. This revealed Unpredicted Susceptibility (False Positive results) to be most common in Proteobacteria and Unpredicted Resistance (False Negative results) most common in Bacteroidetes, but observed for all four phyla and all nine antibiotics.

By identifying unpredicted resistances, this suggests that the results in the previous chapter where antibiotic resistance determinants are enriched in Proteobacteria is not a true reflection of an enrichment of phenotypic antibiotic resistance. As in the previous chapter, it is important to note that database bias may explain this: a database designed from pathogenic bacteria (which are predominantly Proteobacteria, see Fig 1.3) may be more likely to identify similar antibiotic resistance genes or mutations in bacterial isolates more closely related to pathogens. It is likely that this partially explains the difference in observed enrichment of antibiotic resistance genotypes and antibiotic resistance phenotypes. Indeed, using alternative databases with generally similar False Positive and False Negative rates supports this. Moreover, the system I defined to classify isolates as phenotypically susceptible or resistant used all the isolates studied. In the future, the system should be redefined using isolates within more closely related taxa. Once studies approach the magnitude of those performed in pathogenic isolates (e.g. hundreds or thousands of isolates per species), the species level would be the most appropriate taxon to use. However, that resistance is better predicted to some antibiotics than others may also reflect the main mechanism of resistance for these enzymes: antibiotic resistance caused by the presence or absence of a particular gene (e.g. beta-lactamases or tetracycline resistance proteins are more likely to be called

accurately than a single nucleotide mutation). The largest proportion of the Unpredicted Resistances in Bacteroidetes were for gentamicin (46 %), which as discussed is thought not to work on anaerobic organisms such as Bacteroidetes. However, there are several other types of Unpredicted Resistance observed in the Bacteroidetes isolates, as well as relatively high levels of Confirmed Resistance. This indicates that Bacteroidetes make important contributions to the antibiotic resistance potential of the gut microbiota, which may have implications for the treatment of opportunistic pathogens caused by members of this phylum.

It was difficult to account for potential intrinsic resistances in this study to allow for the fact that current rules of antibiotic spectrum might not apply across the diversity of gut bacteria. I have already discussed one instance where Unpredicted Resistances might have been due to intrinsic resistance (gentamicin), but all the Proteobacteria isolates were also resistant to vancomycin as well. This antibiotic targets Gram-positive bacteria specifically, so this result is not surprising. In other cases, where there is variation on genotype/phenotype combinations between very closely related isolates, Unpredicted Resistances might indicate instances of novel antibiotic resistance genes or mutations.

In this study I identified two isolates with the highest possible ceftriaxone resistance measured in the absence of any genetic determinants of beta-lactam resistance. I investigated these isolates, *Bacteroides faecis* 18048_2#66 and *Lachnospiraceae nov.* 20287_6#18, for novel antibiotic resistance genes. This included detailed analysis and experiments regarding one particular candidate beta-lactamase (“Group 2384”), although I was not able to confirm its function. The Group 2384 gene was integrated into the pACYAC184 plasmid within a tetracycline resistance gene and thus was under the control of the tetracycline resistance gene promoter. Accordingly, the Group 2384 gene should have been expressed; however, the mRNA may not have been translated into the protein product with potential beta-lactamase

activity. It is also possible that the observed phenotypic variation in the three *B. faecis* isolates could be caused by differences in gene expression that cannot be detected with WGS alone. Q-PCR or RNA-sequencing of isolates growing in the presence of antibiotics, such as ceftriaxone, are two methods that may help determine if this is the case. In particular, qPCR experiments would help identify if the expression level of the candidate Group 2384 beta-lactamase is associated with the presence of mutations (as discussed towards the end of section 4.2.8) and/or differences in ceftriaxone sensitivity.

Furthermore, I was unable to identify any other candidate novel antibiotic resistance genes with the comparative genomics method. This could be due to the presence of novel beta-lactamases of less than 90 % similarity to beta-lactamases in the NCBI Protein database, though above 60 % similarity, no additional candidate beta-lactamases were identified. Below 50 % similarity, additional candidate beta-lactamase genes were identified and so these may be of interest; however, high sequence similarity is usually required to infer functional similarity[249]. This can be extended to other observations of Unpredicted Resistance; less stringent similarity cut offs when searching for the presence of antibiotic resistance determinants may produce more hits that could explain these observations, however, the Unpredicted Susceptibility rate is likely to rise in response.

The shotgun cloning method did not prove successful in identifying candidate novel antibiotic resistance genes during this study either. As hundreds of colonies were yielded from the control input DNA, this suggests the problem lies with the input DNA from *Bacteroides faecis* 18048_2#66 and *Lachnospiraceae nov.* 20287_6#18. The input DNA is therefore likely a highly critical factor for this protocol. In the future, this method will continue to be optimised within our laboratory as it would be a valuable tool to have available. Identifying genes conferring phenotypes such as antibiotic resistance and improve genome annotation in these organisms,

many of which are novel and/or uncharacterized, would be especially useful. In particular, any novel antibiotic resistance genes from gut bacteria should be curated into databases of known antibiotic resistance genes so that they can be included in antibiotic resistance surveillance programmes, such as the European Antimicrobial Resistance Genes Surveillance Network (EURGen-Net)[250].

Although susceptibility was generally better predicted than resistance, there were relatively high rates of Unpredicted Susceptibility in the Proteobacteria. This means that antibiotic resistance can be overpredicted. This also has relevance for healthcare: if a bacterium was predicted to be resistant to a particular antibiotic, a patient might be prescribed a different antibiotic that could pose more harm by being broader spectrum or more toxic than was required. Unpredicted susceptibility could be due to mis-calling a mutation as present when it is in fact absent, or by identifying the presence of housekeeping genes involved in the regulation of antibiotic resistance genes (such as the *vanR* regulator of vancomycin resistance[251]) . It could also be due to lack of expression or compensatory mutations to offset any fitness cost associated with the predicted antibiotic resistance mechanism. Alternatively, the observed genetic determinants of resistance could be taxon-specific. If a resistance determinant is observed in a taxon other than what it has been described in in the literature, something may be missing from the original host that is required for the determinant to be expressed or functional. It is again important to acknowledge the caveat of using CARD, whereby the majority of its antibiotic resistance genes and mutations have been described in just a few species of pathogenic bacteria. Therefore, in the context of commensal gut bacteria, they may not function as described in pathogenic isolates. It would be interesting to investigate these Unpredicted Susceptibilities in more depth. For example, the impact of individual genes on false predictions could be studied: if the presence of a gene or mutation

always results in Unpredicted Susceptibility (i.e., despite its presence the isolate is susceptible to that antibiotic), then determinants with this pattern could be filtered out as a poor predictor of antibiotic resistance. Taking this work forward must also involve looking for antibiotic resistance genes from other sources (e.g. curating those from functional metagenomic studies of bacterial communities) and determining specific relationships between the presence of antibiotic resistance determinants in a genome and antibiotic sensitivity. To really understand these relationships, more isolates of commensal gut bacteria should be studied (on a scale similar to that of pathogenic bacteria) and MICs should be determined for a more specific antibiotic sensitivity measurement.

This chapter highlights that currently a rule-based approach to estimating antibiotic sensitivity in human gut microbiota is not without flaws. Databases of antibiotic resistance determinants established through research on a relatively small number of pathogens, such as CARD, should therefore be used with caution when applied to more diverse, less well-characterised organisms – such as the human gut microbiota. As demonstrated, these methods can under-predict the antibiotic resistance of such isolates and show bias towards Proteobacteria. These databases can be useful and accurate for well-studied pathogenic bacteria[168,252], but if we are to accurately predict antibiotic resistance in the human gut microbiota and in metagenomic samples, more comprehensive databases of resistance genes are required. It may be necessary to have separate databases for common pathogens and for other bacteria, such as opportunistic pathogens, commensal gut microbiota, or other types of environmental bacteria. Since pathogens can acquire antibiotic resistance genes from environmental bacteria, these types of databases would be useful to help monitor the emergence of clinically relevant antibiotic resistance in clinical isolates of disease-causing bacteria. However, this will require additional similar studies comparing antibiotic resistance genotypes and phenotypes

but at a much larger scale, perhaps using high throughput alternative phenotyping methods such as plate based assays[253,254].

Moreover, functional metagenomics can be a very useful tool for identifying candidate novel antibiotic resistance genes[139,141,169,176,255,256], but requires optimization if used to study individual isolates. Putative ARGs are often annotated as such based on nucleotide or amino acid similarity to known antibiotic resistance genes/proteins and as such it is unknown whether they will confer phenotypic resistance. Determining the level of resistance that can be conferred and the distribution of novel resistance genes should become a routine part of these experiments. Since understanding antibiotic resistance genes among communities of bacteria and in individual, uncharacterized bacteria is difficult, perhaps prioritising bacteria/antibiotics of special interest – such as clinical or ecological relevance – is needed to focus the efforts of novel antibiotic resistance gene discovery.

Overall, this chapter shows that phenotypic antibiotic resistance in gut microbiota can vary between closely related isolates of commensal gut bacteria, much like in pathogenic bacterial species. Moreover, the Bacteroidetes and Firmicutes also demonstrate extensive phenotypic resistance, despite Proteobacteria appearing enriched for clinically relevant genetic determinants of antibiotic resistance in the previous chapter. These results further emphasise the role of the human gut microbiome as a reservoir for antibiotic resistance in terms of its occurrence and prevalence, but also that the extent of this is not yet fully known. In the next chapter, I will investigate the dynamics of antibiotic resistance in human gut microbiota and how commensal gut bacteria can evolve and spread antibiotic resistance.

Chapter 5. Modelling the development of antibiotic resistance *in vivo*

5.1 Introduction

In the previous two chapters antibiotic resistance genotypes and phenotypes in commensal human gut bacteria were investigated. In addition to determining the taxonomic distribution of genetic determinants of resistance and comparing antibiotic resistance genotype with phenotype, it is also important to understand the dynamics of how gut bacteria develop and spread antibiotic resistance in the human microbiota. Several *in vitro* studies have investigated the horizontal transfer of antibiotic resistance genes between closely related bacterial species, such as plasmids from *Salmonella* into *E. coli*[257] or from *Lactococcus* spp. into *Bacteroides*, *Bifidobacterium* and *Enterococcus* spp.[258], but results from *in vitro* experiments cannot easily be extrapolated to *in vivo* situations. *In silico* analysis of publicly available bacterial genomes has identified antibiotic resistance genes with high sequence similarity between human and animal gut bacteria and pathogens[152,259], indicating horizontal transfer, but this is indirect evidence. For direct evidence, we need to study a baseline level of antibiotic resistance in the human gut microbiome and observe how it can change in response to selective pressure caused by antibiotic therapy.

Studies with this goal have been performed before, such as those that use 16S rRNA gene sequencing to assess impact of antibiotics on the overall community[178,260-262]. For example, a study of combined amoxicillin and metronidazole on wild-type mice found that approximately 70 % of 16S rRNA gene sequences in antibiotic treated mice were Proteobacteria, compared to just 1 % in control non-treated mice[179]. In addition, prolonged reductions in overall diversity were observed even after antibiotic treatment was stopped[179]. However, 16S rRNA gene

profiling studies do not allow for high-resolution taxonomic analysis, as often species share similar 16S rRNA gene sequences. Because of this, strains and some species (such as *E. coli* and *Shigella*) cannot be distinguished with this method[263]. Moreover, these studies rely on databases of characterised 16S rRNA gene sequences. As discussed previously, there are still uncharacterised human gut bacteria without reference genomes available, meaning not all 16S rRNA gene sequences can be classified taxonomically and are therefore difficult to study. This is also true of whole genome shotgun metagenomics, which relies on databases of published reference bacterial genomes: previous studies using whole genome shotgun metagenomics have identified high levels of unclassified sequences[264]. Reference genome based metagenomics, where a custom database of reference genomes is curated and tailored for metagenomic analysis relating to the study being performed, is an emerging method to circumnavigate these problems[151,265,266]. This involves culturing bacteria present in the sample to be studied, identifying novel bacteria and generating new reference genomes for these organisms, and combining these into a database of reference genomes[146,151]. Reference genome based metagenomics was implemented in the present study to provide this tailored, high-resolution analysis of antibiotic resistance dynamics in gut microbiota. The culturing and WGS component of reference genome based metagenomics also assists strain-level resolution, which has typically been difficult with whole genome shotgun metagenomics due to limitations with sequencing depth and reference databases[267].

Studies on humans are often confounded by factors such as different genetics, diets, lifestyles, health condition, infection state and being limited to relatively few participants. Experiments in mice therefore offer a much higher level of control, as we can use mice with the same genetic background, from the same litter, that are fed the same diet and kept in controlled living conditions, and it is possible to study large numbers of individuals in these controlled

conditions. Previous models have used the di-associated mouse, in which a recipient and donor strain of interest are introduced into GF mice to assess plasmid conjugation frequency in the gut environment. This model has indicated that inter-species transfer is possible through the transfer of tetracycline and erythromycin resistance from *Lactobacillus plantarum* to *Enterococcus faecalis*[268]. Conventional mouse models have helped identify that transfer of plasmids bearing antibiotic resistance genes (ARGs) can occur at high frequency in the gut[269], even when colonisation resistance is present. However, the di-associated model is considered the "worst case scenario" model due to the lack of colonisation resistance from indigenous microbiota and so does not model the typical gut environment[270]. Moreover, studies in wild-type mice are not fully representative of biological processes in humans as mice have gut microbiomes that differ from humans[271]. Our lab has established methods to colonize mice with human-derived gut microbiota: these are GF mice colonised by gut bacteria of healthy humans following faecal transplant of human stool into the GF mice. This provides a very powerful tool for controlled experiments regarding the gut microbiome in a manner that is more representative of processes in humans. In addition, by culturing and sequencing whole genomes of bacteria isolated before and after antibiotic exposure, there is tremendous power to discriminate ARG-HGT events at a level not previously possible.

In this chapter, I use mice with humanised gut microbiota to model dynamics of antibiotic resistance in the human gut bacterial community and individual strain evolution following antibiotic therapy. Amoxicillin was used as the antibiotic providing selective pressure on the gut bacteria of these humanised microbiota mice. Amoxicillin is considered a relatively broad-spectrum penicillin, and is one of the most prescribed drugs globally[18]. Moreover, amoxicillin is often administered orally, and so can directly impact on our indigenous gut microbiota as well as the infections it is prescribed to treat. As we are still discovering novel and

uncharacterised members of the human gut microbiota, the full impact of broad-spectrum antibiotics such as amoxicillin on commensal gut bacteria remains to be defined. To investigate the comprehensive impact of amoxicillin on human gut bacteria including those considered uncharacterised, I combined large-scale culturing, reference genome based shotgun metagenomics and whole genome sequencing with experiments in humanised microbiota mice. Large-scale culturing of thousands of strains of human gut bacteria allows the type and extent of any genomic and phenotypic changes caused by amoxicillin therapy to be investigated by actively tracking the movement of genes conferring AMR. Therefore, this chapter aims to assess the impact of amoxicillin therapy on amoxicillin resistance in gut bacteria of mice colonised with human-derived intestinal microbiota, uniquely at both the community level and genome level of individual strains.

5.2 Results

5.2.1 Overview of mouse models

To model the impact of amoxicillin on antibiotic resistance in the human gut microbiome, I utilised two mouse lines with human-derived gut microbiota that had been established in our laboratory prior to the start of my PhD. The mouse lines were established by performing faecal microbiota transplant (FMT) of homogenised stool from "healthy" human donors ("Donor 2" and "Donor 7", who also contributed faecal samples for developing the HBC[143]) into germ free mice. Donors were considered healthy if they had not taken antibiotics in the six months prior to donating stool, suffered any gastrointestinal conditions or taken oral medications. FMTs were performed via oral gavage to each GF mouse weekly for three weeks to establish Donor 2 and Donor 7 specific cohorts. This was repeated in several GF mice, which were then bred (e.g. Donor 2 mouse with Donor 2 mouse or Donor 7 mouse with Donor 7 mouse) to produce

two separate mouse lines. Culturing has been performed extensively on samples from human Donor 2 and Donor 7[146,151]. The gut microbiota of Donor 2 mice differs from Donor 7 mice in the following ways: different community composition (i.e., different bacteria, see Appendix 3, Fig. A3.1); different antibiotic resistance potential, including amoxicillin-resistant organisms and antibiotic-resistance genes present in each community (Appendix 3, Figure A3.2 & A3.3); colonisation resistance.

Mice from each line were given amoxicillin orally via drinking water for seven days. Faecal samples were collected immediately before and after amoxicillin treatment. The amoxicillin was provided at a theoretically therapeutic dose (approximately 45 mg/kg/day), based on the concentration required to adequately exceed the minimum inhibitory concentration (MIC) of sensitive organisms in otitis media ear infections[206]. The quantity and concentration of drinking water containing amoxicillin was determined using the average weight of a mouse (30 g) and approximate volume of water consumed per mouse per day (5 ml). I performed deep bacterial culturing with three conditions (aerobic, anaerobic, and targeted for spore-formers) on YCFA agar plates. YCFA is considered a medium able to culture a broad range of bacteria. Metascrape samples (metagenomic sequencing of the total bacterial growth on a culture plate) from vegetative bacterial growth on YCFA plates in anaerobic conditions have been shown to be representative of the community in the original stool sample that was cultured from[146]. Thus, we can use anaerobic vegetative metascrape samples to represent the overall gut bacteria community, as well as targeted metascrapes for particular phenotypes such as antibiotic resistance. Each condition was plated in duplicate: with and without the addition of amoxicillin at a concentration of 8 mg/L in the agar plates. This concentration represents a MIC level where if some pathogenic bacteria are still able to grow in its presence they are considered clinically resistant according to EUCAST and CLSI guidelines; thus, this

concentration is used to select what can be considered amoxicillin-resistant bacteria. Details of each experiment performed are described and summarised below (Fig. 5.1, Table 5.1).

In the initial experiment ("D7AMX1"), eight Donor 7 humanised microbiota mice were divided into two cages. Both cages received a therapeutic dose of amoxicillin as described above. Two faecal pellets were collected from each mouse on experimental days 0 (immediately before amoxicillin regime started), 3, 7 (amoxicillin regime ended), 10, 14 (seven day recovery) and 35 (28 day recovery). Both pellets per mouse were weighed, then one was immediately frozen at -80°C and the remaining pellets were pooled per cage. Pooled pellets were homogenised 100 mg/ml in sterile PBS and serially diluted 1 in 10, from 10^{-1} to 10^{-7}. Dilutions were plated as described above (Fig. 5.1). In this experiment, individual colonies were isolated (Table 5.1).

The Donor 7 humanised microbiota mouse experiment was repeated ("LJP01"), but instead of isolating individual colonies metascrapes were collected (Table 5.1). In this experiment, three cages each containing six Donor 7 humanised microbiota mice were treated with amoxicillin as before. Faecal samples were collected and processed as described for D7AMX1. Colonies were counted for all samples, conditions and time points. Bacterial growth on culture plates was mixed with 0.5 ml sterile PBS, scraped off each plate for Day 0, 7, 14 and 35 and collected for whole genome shotgun metagenomic extraction and sequencing.

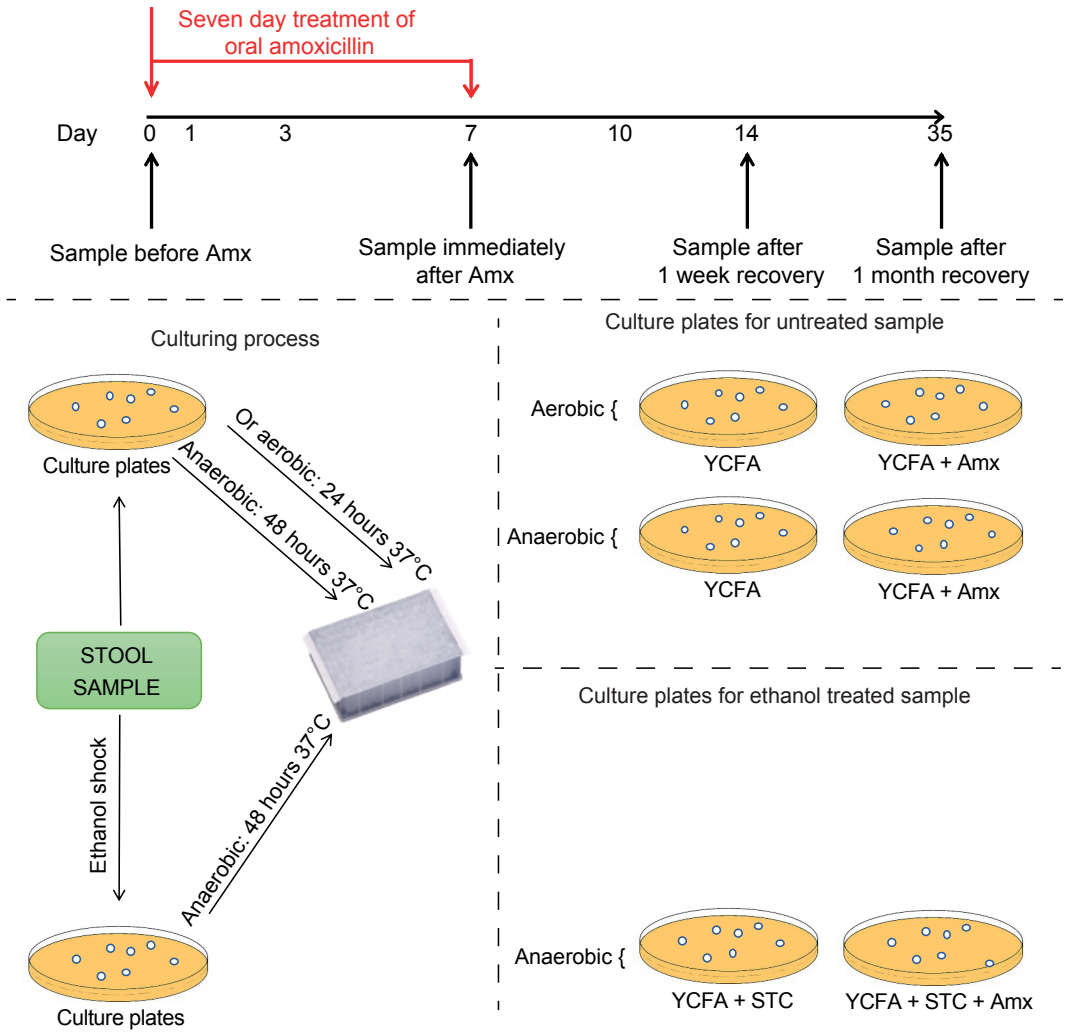


Figure 5.1. Diagram of experiments assessing the impact of amoxicillin on mice with human-derived gut microbiota. The timeline illustrates the sampling days of the experiments and the duration of amoxicillin treatment. The experiments were performed in both Donor 2 and Donor 7 humanised microbiota mice. The lower diagram illustrates the sample processing and different culture conditions and culture plate set ups: faecal pellets were collected from each mouse at each experimental time point. Faecal pellets were weighed and homogenised 100 mg/ml in sterile PBS, then pooled in equal amounts per cage. Pooled faecal homogenates were diluted 1:10 from 10^{-1} to 10^{-7} and appropriate dilutions plated on YCFA agar plates, without and with amoxicillin added (final concentration 8 mg/L amoxicillin). Plates were incubated for 24 hours at 37°C in aerobic conditions, or 48 hours at 37°C in anaerobic conditions. In addition, aliquots of the pooled faecal homogenates were ethanol shocked (diluted 1:4 in 70 % ethanol for 30 minutes) to kill vegetative bacteria and select for ethanol-resistant spores. Ethanol shocked samples were plated on YCFA plus sodium taurocholate (STC, a bile salt to promote spore germination) and incubated at 37°C in anaerobic conditions. All conditions were plated in duplicate to allow metascrape collection (collection of the total bacterial growth on agar plate for total DNA extraction and metagenomic sequencing) and the isolation of individual colonies. Individual colonies were picked into 96 well plates (one isolate per well) containing YCFA broth and incubated for 24 hours (aerobic) and 48 hours (anaerobic) at 37°C. The full length 16S rRNA gene sequence was amplified by PCR from each isolated colony in broth and sequenced by Sanger sequencing to assign species- or genus-level taxonomy based on similarity to reference 16S rRNA gene sequences.

A similar mouse model experiment was performed using three cages each containing six Donor 2 humanised microbiota mice (LJP02). The experiment set up was identical to that of LJP01 (Fig. 5.1) except that Day 0 samples were collected from all cages. In the LJP02 experiment all culturing plates were duplicated to allow both metascrapes to be collected and individual colonies to be isolated (Table 5.1).

Mouse line	No. mice	Sex	Colony counts	Isolated colonies for WGS	Metagenomics
Donor 2	18 across 3 cages	F	LJP02	LJP02	LJP02
Donor 7	18 across 3 cages	F	LJP01	×	LJP01
Donor 7	8 across 2 cages	M(4), F(4)	×	D7AMX1	×

Table 5.1. Summary of experiments assessing the impact of amoxicillin on mice with human-derived gut microbiota and samples or data generated. The table details the mice groups included in each experiment and rationale. Sex: F = female, M = male (brackets indicate number of each). The last three columns indicate what data was collected from which experiments (D7AMX, LJP01 or LJP02). × indicates not generated as part of that experiment.

5.2.2 Impact of amoxicillin on the bacterial load in mice with humanised gut microbiota

The total number of colonies growing on each plate for all pooled homogenates, conditions and time points were counted to determine changes in bacterial load over the time course of the LJP01 and LJP02 mouse experiments (Fig. 5.3). Raw colony counts were converted to colony forming units of bacteria per gram (CFU/g) of stool, representing bacterial load. The data was combined across the Donor 2 and Donor 7 experiments to look for general trends in bacterial load. The CFU/g data was then tested for normal or log-normal distribution using the Anderson-Darling[272], D'Agostino and Pearson[273], Shapiro-Wilk[274], and Kolmogorov- Smirnov tests[275]. As many of the colony count datasets did not pass the normality tests, a

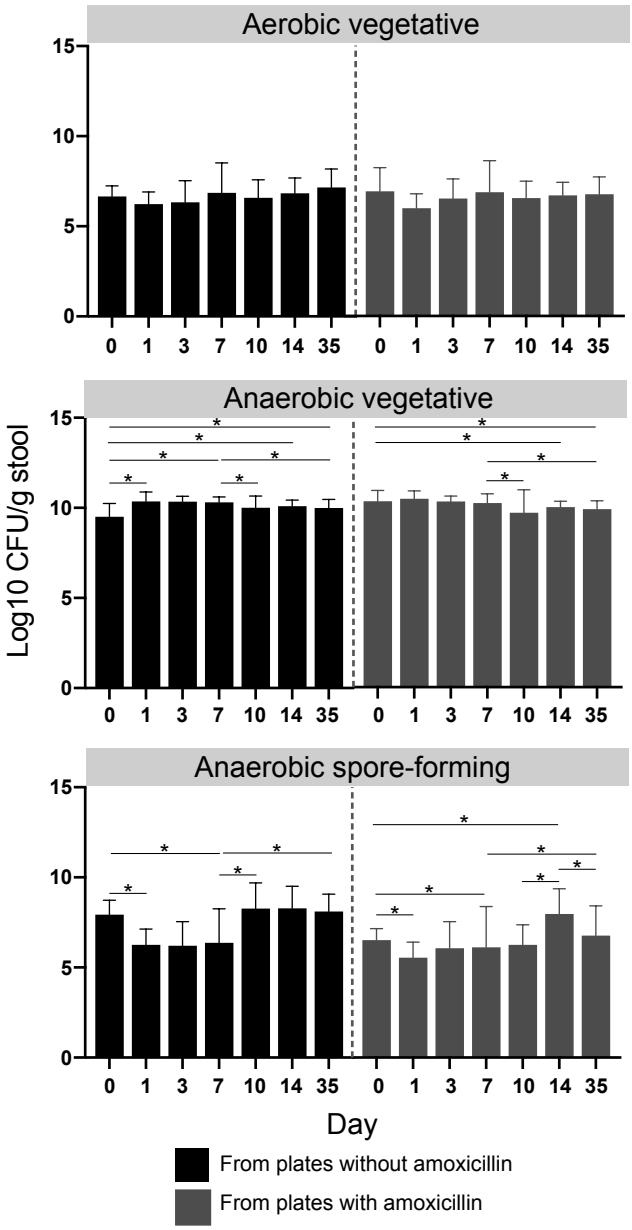


Figure 5.2. Bacterial load over the course of humanised microbiota mouse amoxicillin model experiments. Donor 2 and Donor 7 humanised microbiota mice were given a therapeutic dose of amoxicillin orally via drinking water for seven days; faecal pellets were collected per mouse, homogenized and pooled per cage per time point. The pooled homogenates were diluted and spread on agar plates in a variety of conditions: YCFA agar and anaerobic; YCFA agar plus 8 mg/L amoxicillin and anaerobic; YCFA agar, and aerobic; YCFA agar plus 8 mg/L amoxicillin and aerobic. Aliquots of each pooled homogenate was treated with 70 % ethanol for 30 minutes to select for spores, which were plated on YCFA agar + 1 % sodium taurocholate (STC) with and without 8 mg/L amoxicillin and grown anaerobically. Colonies were counted after 24 hours (aerobic growth) or 48 hours (anaerobic growth) and used to calculate the average bacterial load (colony forming units (CFU)/g stool) and standard deviation per cohort. A non-normal distribution was assumed for performing One-Way Anova (Kruskal-Wallis) tests to determine statistically significant changes in bacterial load over the time course of the mouse experiments (comparisons were performed between consecutive time points for all time points e.g. Day 0 vs. Day 1, Day 1 vs. Day 3, etc. and between major time points i.e., Day 0 vs. Day 7, Day 7 vs. Day 14, Day 14 vs. Day 35, Day 0 vs. 14, Day 0 vs. 35 and Day 7 vs. 35. The two-stage step-up method of Benjamini, Krieger and Yekutieli[193] was used to correct for multiple comparisons by controlling the false discovery rate. * marks where statistical significance ($q < 0.05$) was observed between marked time points.

non-normal distribution was assumed for performing Kruskal-Wallis (One-way ANOVA of non-parametric data) tests. Comparisons were performed between consecutive time points for all time points e.g. Day 0 vs. Day 1, Day 1 vs. Day 3, etc. and between major time points e.g. Day 0 vs. Day 7, Day 7 vs. Day 14, Day 14 vs. Day 35, Day 0 vs. Day 14, Day 0 vs. Day 35, Day 7 vs. Day 35. The two-stage step-up method of Benjamini, Krieger and Yekutieli[193] was used to correct for multiple comparisons by controlling the false discovery rate (q value significant < 0.05). There were no significant changes to aerobic bacterial load.

There was a significant increase in anaerobic vegetative bacterial load from plates without amoxicillin after the first day of treatment (Day 0 vs. Day 1, q value < 0.0001) and over the course of the amoxicillin therapy (Day 0 vs. Day 7, q value <0.0001). In addition, despite a significant reduction between Day 7 and Day 10 (q value 0.0124), the anaerobic vegetative load at recovery time points remained significantly higher than Day 0 (vs. Day 14 and Day 35, q values 0.0087 and 0.0201 respectively). However, Day 35 was significantly lower than Day 7 (0.0087). This indicates that the total anaerobic vegetative load increased after amoxicillin treatment began and then decreased after the treatment was stopped, but not back to its original level. There was also a significant decrease in anaerobic vegetative bacterial load from plates containing amoxicillin between Day 7 and Day 10 (q value 0.0138), and overall between Day 7 and Day 35 (q value 0.0026). There was an overall significant decrease in anaerobic vegetative bacterial load from plates containing amoxicillin (Day 0 vs. Day 14 and Day 35, q values 0.0010 and 0.0002 respectively). This indicates an overall decline in the load of amoxicillin-resistant anaerobic vegetative bacteria during the experiment.

The load of anaerobic spore-forming bacteria cultured on both plates with and without amoxicillin significantly decreased between Day 0 and Day 1 (q values 0.0003 and < 0.0001 respectively). In addition, both sets of anaerobic spore-forming bacteria showed significant

decreases in load between Day 0 and Day 7 (q values 0.0019 and < 0.0001 respectively). There was a subsequent significant increase in anaerobic spore-forming bacteria from plates with without amoxicillin after treatment stopped (Day 7 vs. Day 10, q value < 0.0001). The load of anaerobic spore-forming bacteria from plates without amoxicillin were significantly higher at Day 35 than Day 7 (q value <0.0001), but not significantly different to Day 0. Anaerobic spore-formers from plates with amoxicillin significantly increased between Day 10 and Day 14 (q value 0.0003), and between Day 14 and Day 35 (q value 0.0076). The Day 14 level was significantly higher than Day 0 (q value 0.00206) but whilst Day 35 was significantly higher than Day 7 (0.0163), it was not significantly different to Day 0. This indicates that the amoxicillin had a significant impact on reducing the load of spore-forming bacteria during the antibiotic treatment but that this community recovered in terms of CFU/g after the treatment stopped.

5.2.3 Deep culturing and whole genome sequencing to improve taxonomic classification of metagenomic data

Having studied the gross impact on bacterial load, the impact of amoxicillin on the diversity of the gut bacterial communities in these mice was investigated using high resolution metagenomics that can resolve to species and subspecies level. I first classified the metagenomic sequence reads from the metascrapes (total bacterial growth collected from culture plates) using Kraken[211], comparing them to a database of publicly available reference genomes of gut bacteria (including the 737 from the HBC[151]; this database was created in December 2017 by Dr Sam Forster). To determine the taxonomic resolution, I calculated the percentage of reads in each sample that could be classified to species level using this Kraken

database and determined the average proportion of classified sequence reads for each culture condition (aerobic, anaerobic vegetative and anaerobic spore-forming; Figure 5.4). Although the aerobic and anaerobic vegetative metascrapes were relatively well classified (> 95 %), lower levels of classified sequence reads were observed for spore-forming metascrapes (71.9 %). This indicates that this database is not optimal for classifying all of the metagenomic data generated in this study and I therefore sought to culture and whole genome sequence additional isolates to improve the taxonomic classification and downstream analysis.

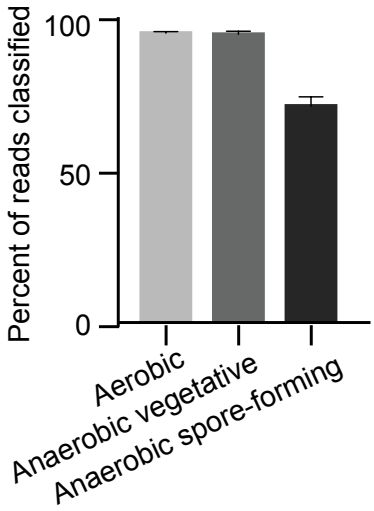


Figure 5.3. Proportion of classified metagenomic sequences reads from metascrapes of cultured faecal bacteria from mice with humanised gut microbiota. Sequence reads were assigned taxonomic classification using Kraken and a database containing publicly available gut bacterial and archaeal genomes, including those in the HBC[151]. The average percentage of classified reads for each culture condition was determined. Bars represent Standard Error of the Mean. This shows that this database was not capable of assigning taxa to all sequence reads, especially for spore-forming bacteria.

To identify candidates of novel reference genomes that might explain some of the unclassified sequence reads, I assembled contigs *de novo* from unclassified reads in one metascrape sample from the Donor 7 and one from the Donor 2 mouse experiments. I then performed a BLASTn analysis of those contigs against the NCBI RefSeq database (as of April 2017). Overall, the majority of unclassified contigs from the Donor 7 mouse sample were most similar to sequences belonging to Firmicutes, particularly the Clostridiales order. Over a quarter (27 %)

of the contigs had high similarity to a *Turicibacter* genome (Fig. 5.4A); this was also true of the Donor 2 mouse unclassified contigs (Fig. 5.4B). This indicated a *Turicibacter*-like organism could explain some of the unclassified data and would be a good candidate for a novel reference genome to add to the Kraken database.

Subsequently, I isolated approximately 11,000 colonies from across all the culture conditions to identify previously uncultured bacteria, including this candidate *Turicibacter*-like organism, and obtain as much diversity from the gut microbiota of these mice as possible. Capillary sequencing of the 16S rRNA gene was performed for all isolates. This resulted in 8838 full length 16S rRNA gene sequences being returned; the reduction in numbers of isolates to 16S rRNA gene sequences is likely caused by the failure to grow in broth or contamination of isolates. The 16S rRNA gene sequences were analysed and assigned OTUs with Mothur[208] using a 99 % identity cut off (i.e., sequences 99 % similar to each other were grouped into one OTU) in terms of similarity to the HBC[143] 16S rRNA gene sequences. In total, 367 OTUs were observed.

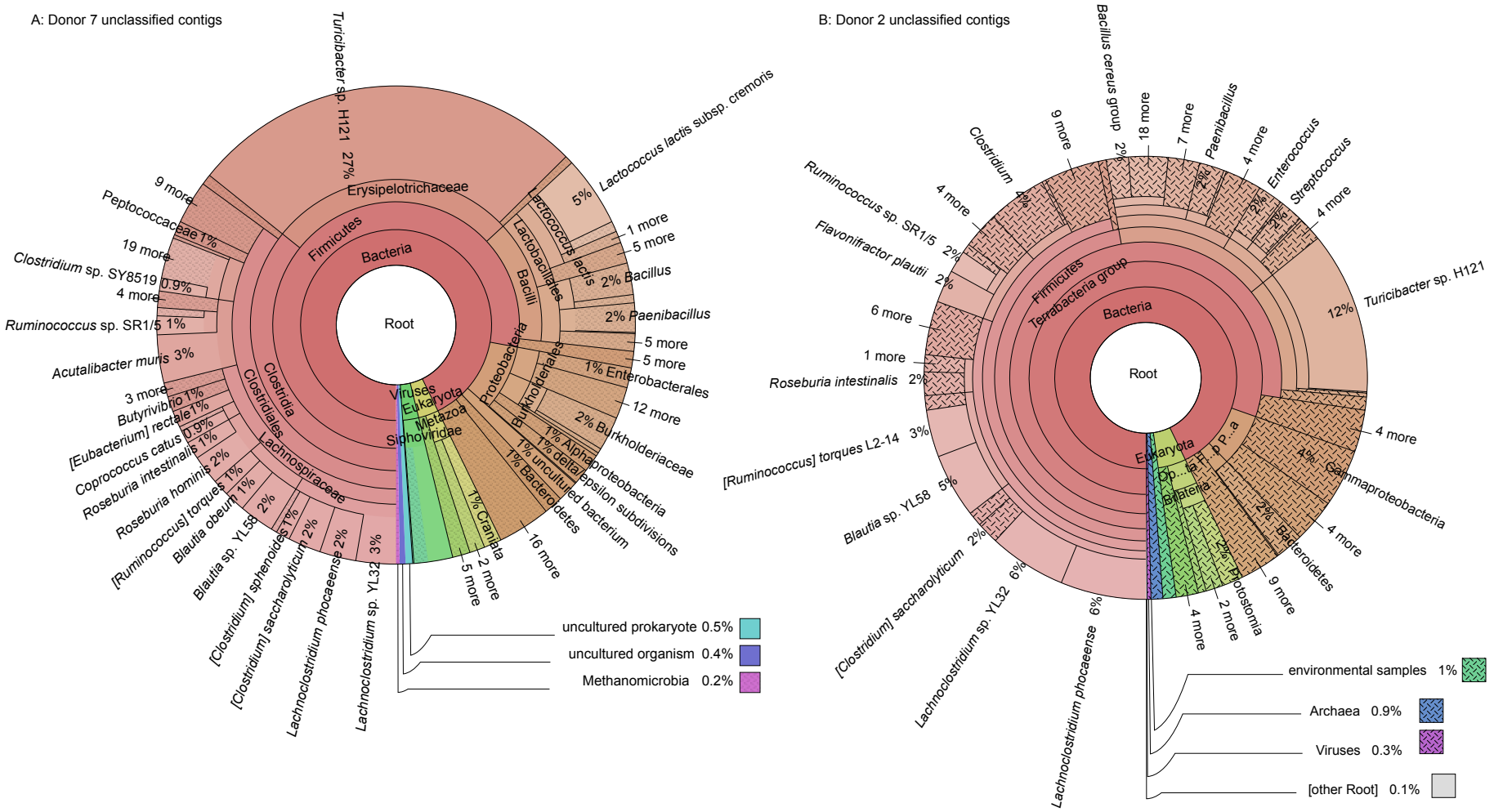


Figure 5.4. Diversity of the contigs assembled de novo from unclassified metagenomic sequence reads. *De novo* assembly was performed on the unclassified reads from a single metagenomic sample, with a high level of unclassified sequence reads, and the resulting contigs analysed using the NCBI RefSeq BLAST database containing all published genomes (as of April 2017). The chart shows the proportion of sequence reads by their most closely related reference genome (inner most ring: lowest taxonomic level e.g. domain; outer most ring: highest taxonomic level e.g. species). In this sample, the highest proportion of unclassified reads were most closely related to *Turicibacter sanguinis*, indicating that bacteria similar to this species would be useful to culture and generate new genomes from.

A rarefaction curve of this data was plotted: as the number of colonies picked increased, the number of OTUs isolated continued to increase even past 8000 colonies (Fig. 5.5). The trend was modelled to predict that the maximum OTU richness would feature 427 OTUs across the two mouse lines, but that approximately 30,000 colonies would be needed to be picked to recover all 427 OTUs. The 8838 colonies I isolated recovered approximately 86 % of the predicted total number of OTUs in gut microbiota of these mice.

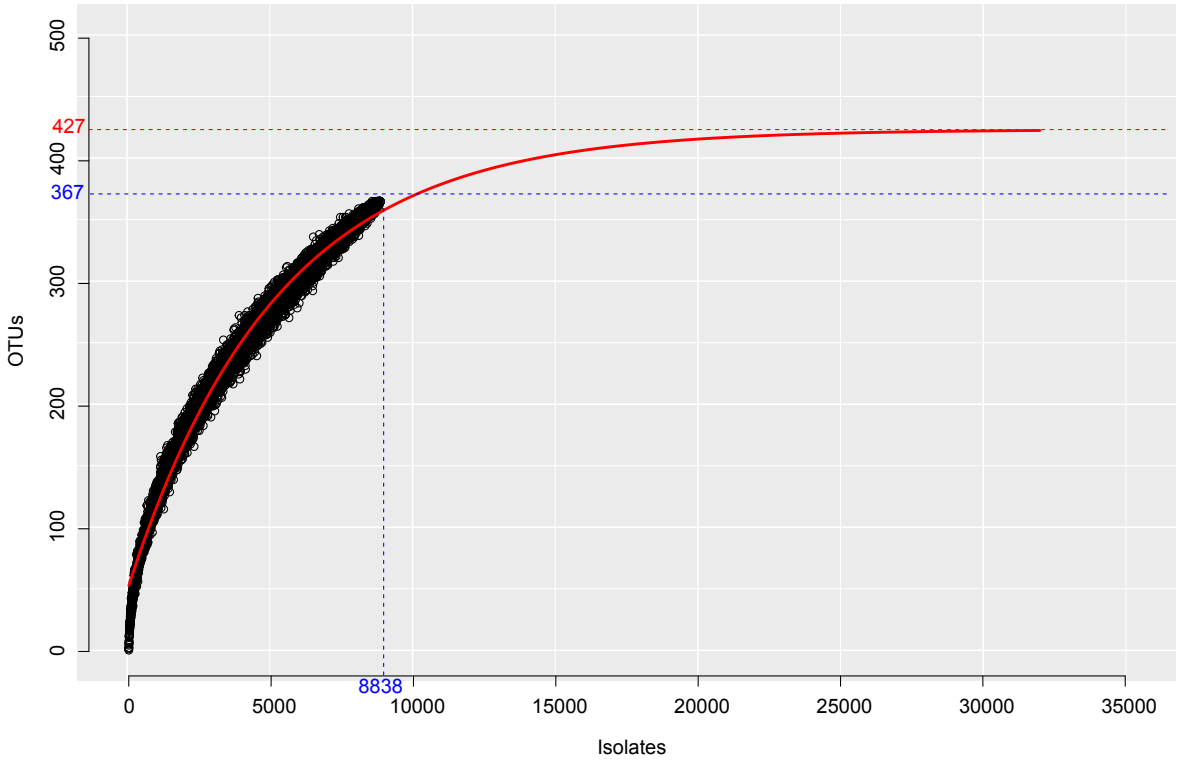


Figure 5.5. A rarefaction curve of the number of OTUs observed against the number of colonies picked. 11,000 individual bacterial colonies were isolated from mice with humanised gut microbiota before and after amoxicillin treatment. 8838 full length 16S rRNA gene sequences were generated and grouped into operational taxonomic units (OTUs) using a 99 % similarity cut off. The number of new OTUs observed as the number of colonies picked increases was plotted, showing that even as we reach almost 9000 colonies new OTUs are still being obtained. The blue dotted lines mark the maximum numbers of isolates picked and number of OTUs observed. The trend in part A was modelled using nonlinear regression to produce a rarefaction curve (solid red line) and estimate the maximum number of OTUs likely to be found across the two mouse lines with humanised gut microbiota: the red dotted line represents this predicted maximum of 427 OTUs. It would be necessary to pick approximately 30,000 colonies to recover the 427 OTUs. Dr Hilary Browne, Dr Sam Forster, Dr B. Anne Neville, Mr Mark Stares, Dr Elisa Viciani and Dr Ana Zhu all helped pick colonies from culture plates into broth cultures.

The consensus 16S rRNA gene sequence for each OTU was compared to 16S rRNA gene sequences in NCBI RefSeq (April 2017); the majority belong to Firmicutes (54.8%, Fig. 5.6) and Bacteroidetes (33.8%), with just a few Proteobacteria (11.4%) and Actinobacteria (1.36%).

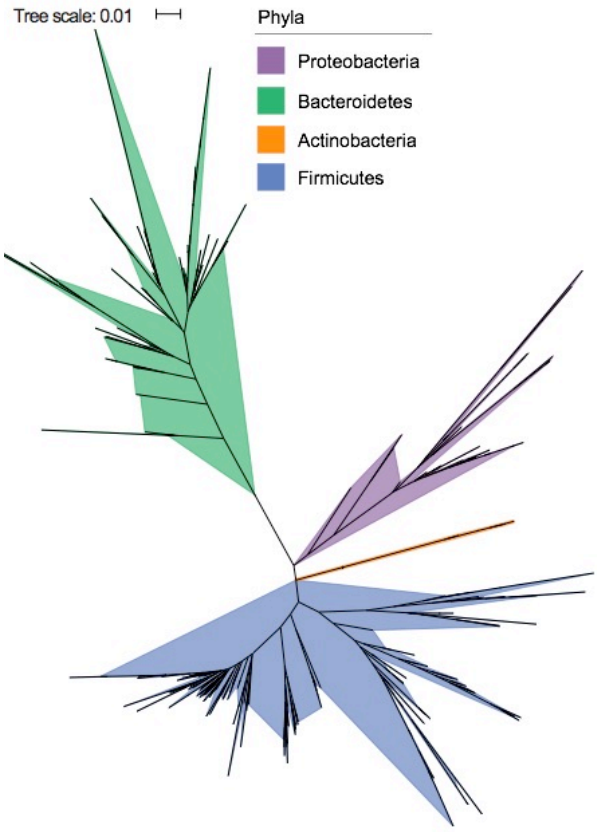


Figure 5.6. A phylogeny of the consensus sequences of 367 OTUs identified from 8838 full length 16S sequences generated in this study. 11,000 individual bacterial colonies were isolated from mice with humanised gut microbiota before and after amoxicillin treatment. 8838 full length 16S sequences were generated and grouped into operational taxonomic units (OTUs) using a 99 % similarity cut off. Consensus sequences of the 367 OTUs were compared to NCBI RefSeq (April 2017). The majority of the OTUs belong to the Bacteroidetes and Firmicutes phyla.

Looking at the proportion of colonies picked per OTU shows that approximately 1100, or 13 %, of the 8838 16S rRNA gene sequences were clustered in an OTU with 97 % nucleotide identity to *Turicibacter sanguinis* (Figure 5.7A). 16S rRNA gene sequences most similar to *Turicibacter* also demonstrated variability (Figure 5.7B), showing that these isolates are not all identical. Three isolates related to *Turicibacter* and 49 others with less than 97 % identity

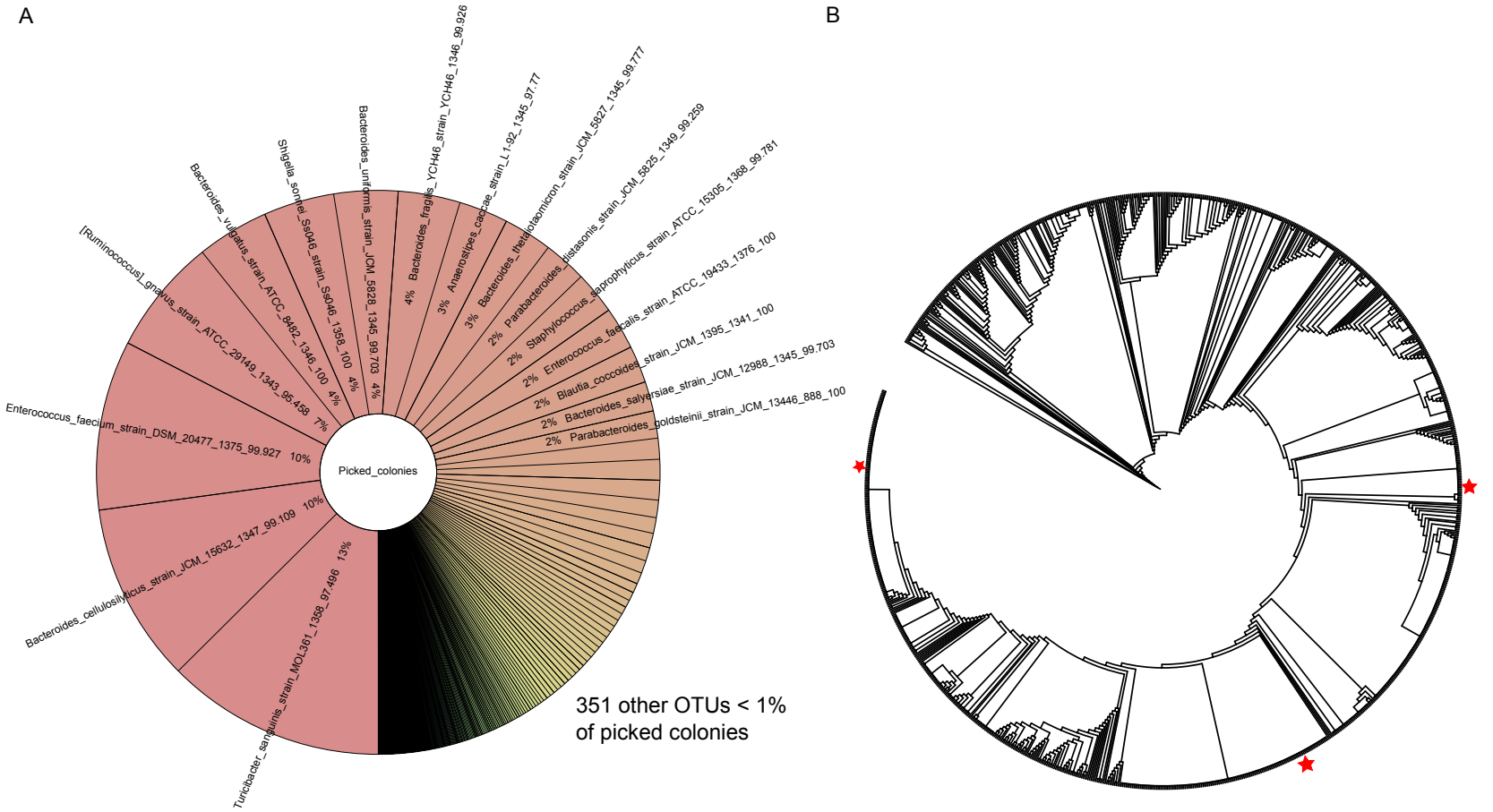


Figure 5.7. Diversity of OTUs isolated by culture from mice with human-derived microbiota. A) 16S rRNA gene amplification and Sanger sequencing was performed for all bacteria isolated in the mouse experiments described in this chapter, resulting in 8838 16S full length sequences. These were grouped into OTUs using Mothur and a 99 % cut off (i.e., 16S rRNA gene sequences >99 % similar to each other were considered one OTU); consensus OTU sequences were generated as part of this process. OTUs were named after the reference genome of highest similarity plus the length of the sequenced 16S rRNA gene and the percentage nucleotide identity to the reference sequence. For example, *Turicibacter_sanguinis*_strain_MOL361_1358_97.496 is the name of an OTU whose 16S rRNA gene is 1358 bases in length and 97.496% similar to the 16S rRNA gene sequence of *Turicibacter sanguinis* strain_ OL361. The proportion of all isolates each OTU contributed was calculated: 1127, or 13 %, were clustered in an OTU with 97 % identity to *Turicibacter sanguinis.* B) A phylogenetic tree of approximately 1127 16S rRNA gene sequences with 97 % identity to the reference *Turicibacter sanguinis* strain MOL361 genome shows there is considerable diversity even within the taxon. The red stars mark three isolates that were purified and DNA extracted for WGS.

to known 16S rRNA gene sequences (and therefore potentially novel species) were purified and whole genome sequenced. In addition, 11 isolates representing very common OTUs (>20 colonies isolated per OTU) and 135 isolates representing taxa that appeared to have increased in amoxicillin resistance during the experiment were purified and whole genome sequenced. Their relationship with the HBC isolates is demonstrated in Figure 5.8.

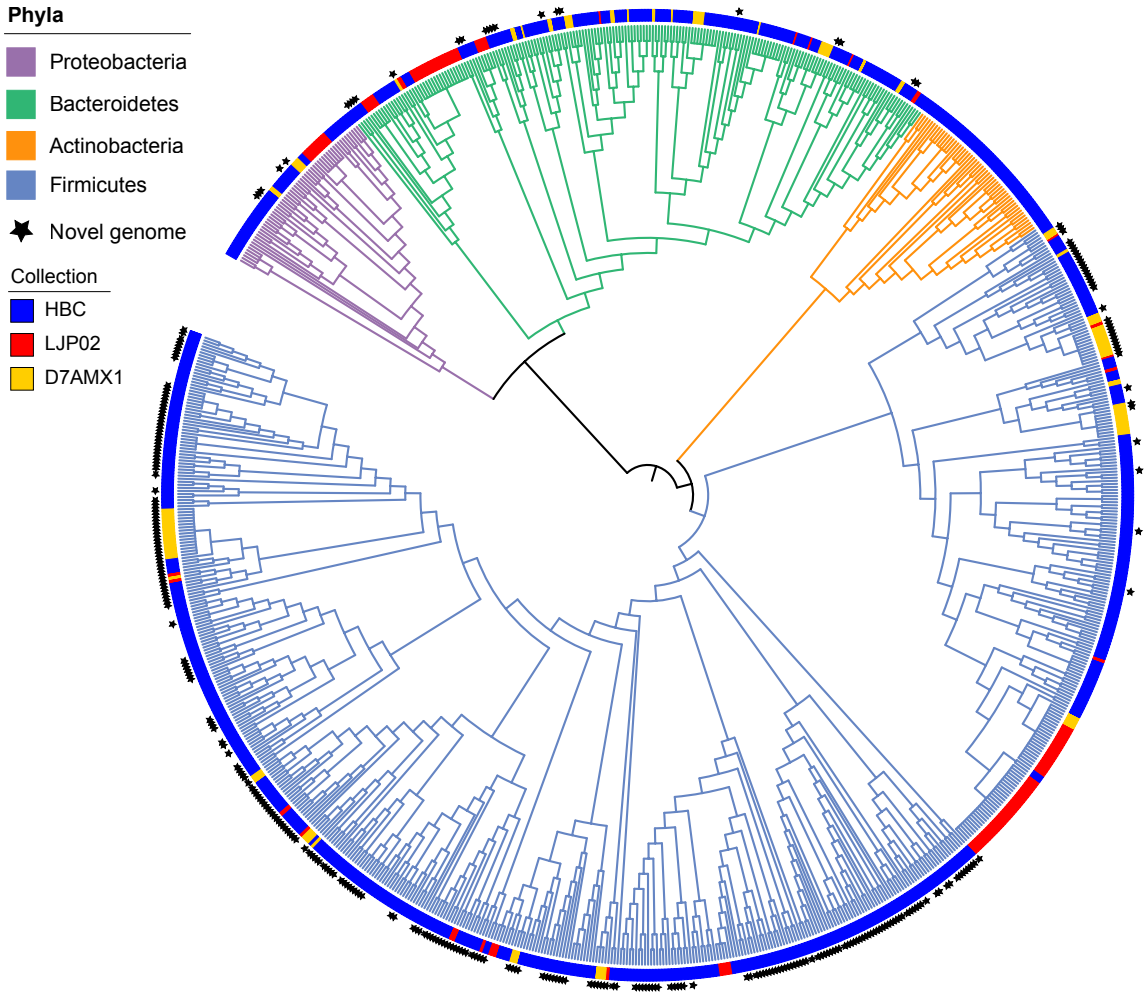


Figure 5.8. The phylogenetic relationship between the HBC and 198 new genomes generated in this study. The Human Gut Bacteria Culture Collection currently contains 737 isolates of human gut bacteria. The amino acid sequences of 40 core genes were extracted from all genomes and used to infer a phylogeny to illustrate the taxonomic diversity of the HBC and 198 genomes generated in this thesis. The stars mark which genomes are considered 'novel' based on the similarity of their 16S rRNA gene sequence to known 16S rRNA gene sequences in RefSeq. Mr Mark Stares performed genomic DNA extractions for the 198 D7AMX1 or LJP02 isolates.

All 198 new genomes were added to the Kraken database used earlier for reclassification of the metagenomic sequence reads from the metascrape data. The custom database contains publicly available gut bacteria genomes plus the 198 genome assemblies generated in this study (of which 52 represent isolates of potentially novel new taxa (16S ID < 98.7 %). The classification of metagenomic sequence reads to species level was improved from 95.6 % to 99.1 % in aerobic metascrapes and 95.3 % to 96.5 % in anaerobic vegetative metascrapes (Figure 5.9). The spore-forming samples were still less well classified at 86.6 % but also the most improved (up from 71.9 %). The resolution of downstream metagenomic data analysis was higher as a consequence.

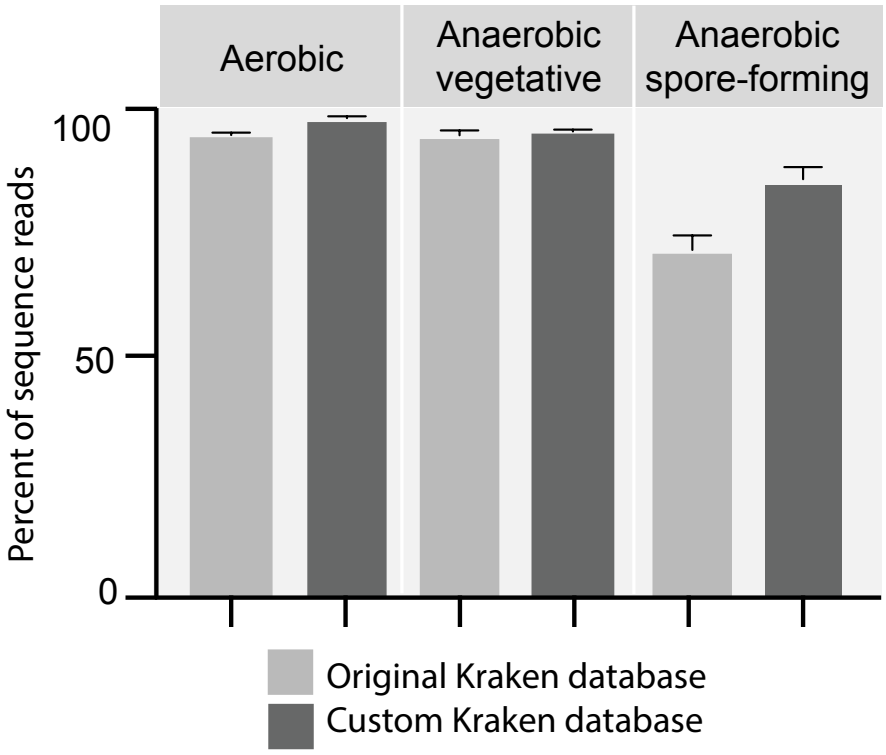


Figure 5.9. Comparisons of proportion of classified metagenomic sequences reads from metascrapes of cultured faecal bacteria from mice with humanised gut microbiota using different databases of reference bacterial genomes. Sequence reads were assigned taxonomic classification using Kraken and a database containing publicly available gut bacterial and archaeal genomes, including those published by Forster *et al.* 2019[151] (the "original" database), plus the same collection with new genomes generated in this thesis added ("custom"). The average percentage of classified reads for each condition was determined. Bars represent Standard Error of the Mean. This shows improved classification rates for all conditions when using the custom database compared to using the original database.

5.2.4 Impact of amoxicillin on the amoxicillin-resistant community

The updated custom Kraken database was used to assign taxonomy to sequence reads and assess the impact of amoxicillin on the diversity of the amoxicillin-resistant community in humanised microbiota mice. Firstly, the raw read counts assigned to each species in the Kraken database were used to determine the Shannon index of alpha diversity. Alpha diversity indices were averaged over metascrape samples from plates containing 8 mg/L amoxicillin, combined across both the Donor 2 and Donor 7 experiments, for each of the three culture conditions (i.e., aerobic, anaerobic vegetative and anaerobic spore-forming from plates containing amoxicillin).

Mean alpha diversity increased in aerobic and anaerobic vegetative metascrapes from plates with amoxicillin (Fig. 5.10) between Day 0 and Day 7, then decreased again after amoxicillin treatment had stopped, though remained higher than initial levels. This may indicate that the amoxicillin caused additional species to become resistant to amoxicillin within an individual, although this was not maintained after the treatment stopped. Alternatively, it may be that amoxicillin-resistant bacteria expanded in abundance from below detection level to above detection level, or that certain species of the same relative abundance were harmed by the amoxicillin, allowing the remaining species of that initial relative abundance to become detectable. In contrast, anaerobic spore-forming metascrapes from plates with amoxicillin, alpha diversity increased over all time points. This indicates that more spore-forming species became resistant over the experiment, or became more detectable, even after the treatment was stopped. However, there was no significant difference between any consecutive time points (i.e., Day 0 vs. Day 7 , Day 7 vs. 14, Day 14 vs. 35), nor between the start and end of the experiment (Day 0 vs. 35) for any culture condition (statistical significance determined by

Mann-Whitney U tests[212], adjusted for multiple-testing with the Benjamini, Hochberg, and Yekutieli method[193]; q value significant < 0.05).

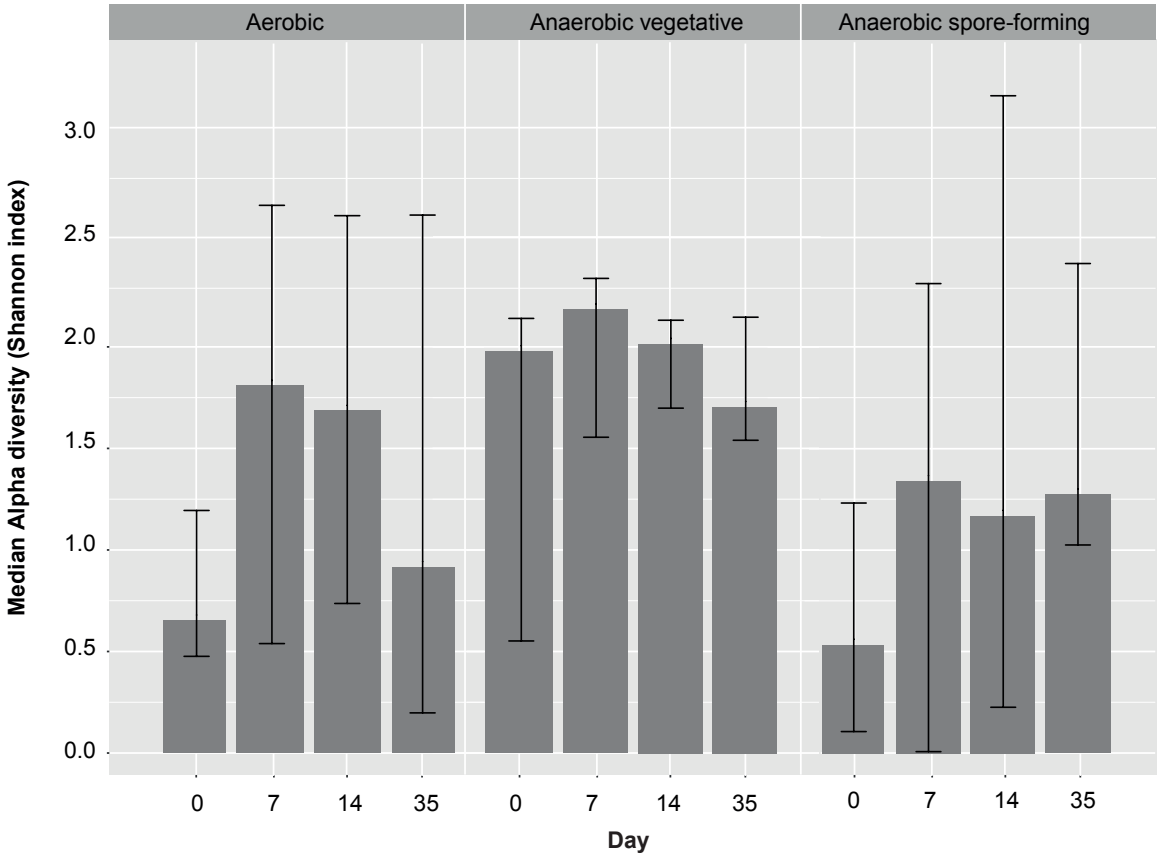


Figure 5.10. Changes in alpha diversity in metascrapes of cultured faecal bacteria from mice with humanised gut microbiota treated with amoxicillin. Sequence reads were assigned taxonomic classification using Kraken and a database containing publicly available gut bacterial and archaeal genomes plus genomes generated in this study. Raw read counts per species were used to determine alpha diversity at each experimental time point (Day 0, 7, 14, 35) for three culture conditions (aerobic, anaerobic vegetative and anaerobic spore-forming), on agar plates with 8 mg/L amoxicillin added. The median alpha diversity increased in aerobic metascrapes from plates with amoxicillin during the amoxicillin treatment (between Day 0 and Day 7) and reduced after amoxicillin treatment had stopped (Day 14 and Day 35), but remained higher than its initial level. A similar pattern was seen in anaerobic vegetative metascrapes from plates with amoxicillin but to a lesser extent. In anaerobic spore-forming metascrapes from plates with amoxicillin, where alpha diversity increased over all time points. There were no significant differences between consecutive time points (i.e., Day 0- vs. , Day 7 vs. 14, Day 14 vs. 35), nor between the start and end of the experiment (Day 0 vs.35) for any culture condition. Statistical significance determined by Mann-Whitney U tests[212], adjusted for multiple-testing with the Benjamini, Hochberg, and Yekutieli method[193]; q value significant < 0.05. Error bars represent the range. Dr Kevin Vervier wrote the R scripts that measured and compared alpha diversity.

Beta diversity (Fig. 5.11) was also determined to identify whether individual samples (i.e., metascrapes from a particular cage of mice) changed in similar ways during the experiment. Bray-Curtis dissimilarity measures were determined from raw read counts and averaged over metascrape samples from plates containing 8 mg/L amoxicillin, across both experiments, for each of the three culture conditions (i.e., aerobic, anaerobic vegetative and anaerobic spore-forming from plates containing amoxicillin). Although in all culture conditions beta diversity increased over the time course of the experiment, indicating that the amoxicillin-resistant communities belonging to different cages of mice became slightly more different from each other following amoxicillin therapy (i.e., there was a cage-specific effect), again, there were no significant differences between consecutive time points (i.e., Day 0 vs. Day 7, Day 7 vs. 14, Day 14 vs. 35), nor between the start and end of the experiment (Day 0 vs. 35) for any culture condition (determined as described for alpha diversity). This indicates that overall, the amoxicillin-resistant communities belonging to different cages did not become significantly different from each other during or after amoxicillin treatment.

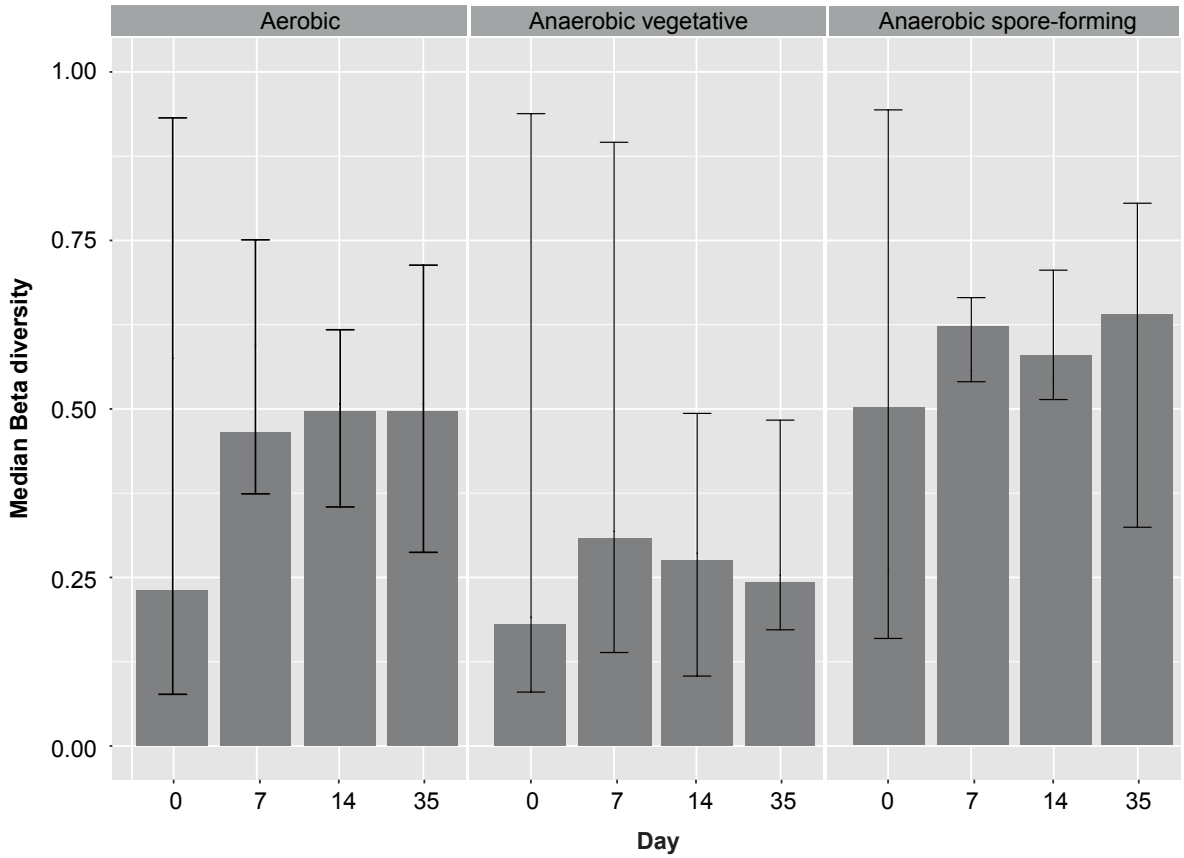


Figure 5.11. Changes in beta diversity in metascrapes of cultured faecal bacteria from mice with humanised gut microbiota treated with amoxicillin. Sequence reads were assigned taxonomic classification using Kraken and a database containing publicly available gut bacterial and archaeal genomes plus genomes generated in this study. Raw read counts per species were used to determine beta diversity at each experimental time point (Day 0, 7, 14, 35) for three culture conditions (aerobic, anaerobic vegetative and anaerobic spore-forming), on agar plates with 8 mg/L amoxicillin added. Median beta diversity increased in aerobic metascrapes over all time points. In both anaerobic vegetative and anaerobic spore forming metascrapes from plates without amoxicillin, beta diversity increased between Day 0 and Day 7, dropped slightly at Day 14 then increased again at Day 35. Anaerobic vegetative and anaerobic spore-forming metascrapes from plates with amoxicillin also show an increase in beta diversity from Day 0 to Day 7 which then decreases at both Day 14 and Day 35. There were no significant differences between consecutive time points (i.e., Day 0- vs. Day 7, Day 7 vs. 14, Day 14 vs. 35), nor between the start and end of the experiment (Day 0 vs.35). for any culture condition. Statistical significance determined by Mann-Whitney U test[212], p value significant < 0.05. Error bars represent the range. Dr Kevin Vervier wrote the R scripts that measured and compared beta diversity.

Despite the overall changes in diversity not being significant, it is possible that there were significant changes occurring in individual taxa. To assess this, the relative abundance of species in the amoxicillin-resistant communities for each culture condition (aerobic, anaerobic vegetative and anaerobic spore-forming metascrapes from plates containing amoxicillin) was determined. Read counts were normalised per sample and then the normalised values were averaged across all samples from the Donor 2 and Donor 7 experiments for each of the three

conditions. The species were then ranked by their mean relative abundance in each culture condition and the mean relative abundance of the top ten most abundant species plotted, with the relative abundance of the remaining species grouped as “Other”. This “Other” category formed the majority of (>92 %) the relative abundance of species in the anaerobic vegetative and spore-forming metascrapes from plates with amoxicillin (Fig. 5.12) at all time points.

For the aerobic amoxicillin-resistant community, *Enterococcus faecalis* dominated at Day 0 before amoxicillin therapy (mean relative abundance 78.5 % of sequence reads per sample) and was significantly reduced at Day 7 (0.88 % mean relative abundance, q value < 0.001; determined by unpaired t-tests with Welch’s correction[276], adjusted for multiple-testing with the Benjamini, Hochberg, and Yekutieli method[193] and significant < 0.05). The mean relative abundance of *E. faecalis* then increased by Day 14 (to 8.19 %), and was dominant again at Day 35 (80.0 %, significantly increased from Day 14; q value < 0.001). Overall, there was no significant difference between the initial (Day 0, 78.5 %) and final measured mean relative abundance of *E. faecalis* (Day 35, 80.0 %), showing that although this species reduced in relative abundance following amoxicillin therapy it recovered to its initial level after treatment stopped.

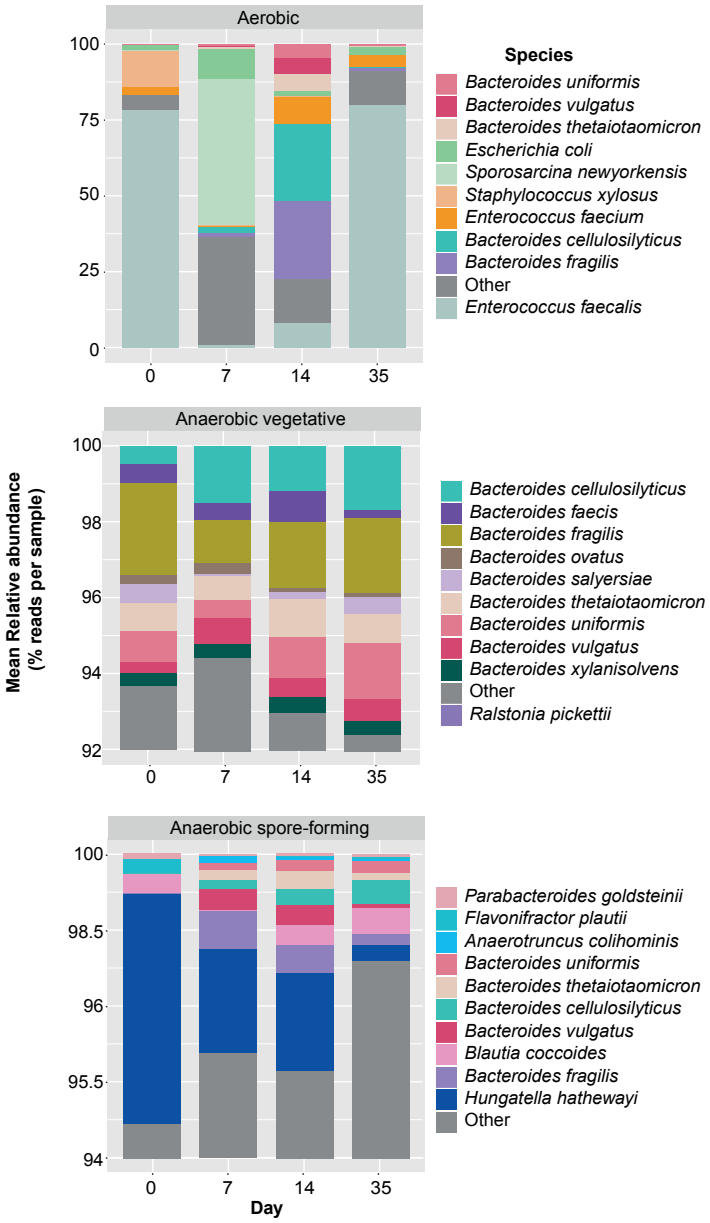


Figure 5.12. Relative abundance of species in metascrape samples cultured from mice following amoxicillin therapy. Sequence reads were assigned taxonomic classification using Kraken and a database containing publicly available gut bacterial and archaeal genomes plus genomes generated in this study. Read counts were normalized per sample and the relative abundance of each species averaged across all samples for each culture condition. Species were ranked from highest to lowest mean relative abundance per culture condition and the top 10 species for each are shown in these bar plots. All other species were grouped as "Other"; this category made up over 92 % of the total relative abundance of bacteria in the anaerobic vegetative and anaerobic spore-forming samples and so the y axis for these two conditions is limited to show the ten most abundant species more closely. Only metascrapes from plates containing amoxicillin are shown to represent the amoxicillin-resistant community. The aerobic amoxicillin resistant community was dominated by *Enterococcus faecalis* at the start of the experiment and end of the experiment (Day 0 and Day 35) but was much less relatively abundant at days 7 and 14, which were dominated by *Sporosarcina newyorkensis* or *Bacteroides fragilis* and *B. cellulosilyticus* respectively. The anaerobic vegetative amoxicillin resistant community most obviously shows a reduction in *Ralstonia picketti* and expansion of *B. cellulosilyticus* from Day 0 over all other time points, plus an increase in "Other" species at Day 7 that was not maintained at days 14 or 35. Similarly, the anaerobic spore-forming amoxicillin resistant community shows continual increases in "Other" species as well as a reduction in *Hungatella hathewayi*.

Whilst *E. faecalis* was less relatively abundant at Day 7 than Day 0, the amoxicillin-resistant aerobic community at Day 7 was dominated by *Sporosarcina newyorkensis: S. newyorkensis* was significantly higher at Day 7 (48.2 %) than Day 0 (0.17 %, q value < 0.001), but had reduced again by Day 14 (0.003 %, q value < 0.001). The relative abundance of *S. newyorkensis* was lowest at the end of the experiment (Day 35, 0.0005 %) but this was not significantly different from Day 14 or Day 0. Again, this shows that although this species increased in relative abundance following amoxicillin therapy it also returned to its initial level after treatment stopped. *Bacteroides fragilis* and *B. cellulosilyticus* were the dominant species at Day 14 (mean relative abundance 25.5 % and 25.3 % respectively). These two species were both significantly more relatively abundant at Day 14 than Day 7 (*B. fragilis* 1.2 %, *B. cellulosilyticus* 2.19 %) or Day 35 (*B. fragilis* 0.99 %, *B. cellulosilyticus* 0.47 %; q values all < 0.001). There was no significant difference between the initial and final relative abundance of *B. fragilis* and *B. cellulosilyticus* (Day 0: 0.16 % and 0.06 % respectively; Day 35: 0.99 % and 0.47 % respectively). This shows that these *Bacteroides* species became dominant in the aerobic amoxicillin-resistant community in the initial seven day recovery period following treatment, but had returned to their initial level by the end of the experiment. Finally, the relative abundance of "Other" species increased significantly from Day 0 (4.40 %) to Day 7 (29.6 %, q value < 0.0001) and significantly decreased again between Day 7 and Day 14 (4.97 %, q value < 0.0001). There was a slight increase in the relative abundance of "Other" species at Day 35 (9.96 %), but this was not significantly different from Day 14 or Day 0. This indicates that less abundant amoxicillin-resistant aerobic species became more abundant during the amoxicillin therapy and then returned to their initial level after the treatment ended. Overall, this suggests the amoxicillin therapy caused significant changes to the relative abundance of several amoxicillin-resistant species capable of growing in aerobic conditions, but that the final overall composition of this community was similar to its initial state.

The amoxicillin-resistant anaerobic vegetative community was less dominated by particular species: the mean relative abundance of "Other" species was consistently high (>92 % of sequence reads), though there was a slight increase at Day 7 (94.0 % compared to 92.2 % at Day 0) that was not maintained by Day 14 (92.6 %) or 35 (92.3 %). *Bacteroides fragilis* was the most relatively abundant species at 2.38 % of sequence reads per sample at Day 0; this reduced to 1.14 % at Day 7, then increased to 1.73 % at Day 14 and 2.00 % at Day 35. There was also a reduction in *Ralstonia picketti* (1.24 % at Day 0 to < 0.0001 % at Day 7, 14 and 35) and an increase of *B. cellulosilyticus* from Day 0 (0.45 %) over all other time points (1.47 % at Day 7, 1.16 % at Day 14, 1.67 % at Day 35). However, none of these changes were statistically significant. This shows that the relative abundance of individual species in the amoxicillin-resistant anaerobic vegetative community was not significantly affected by amoxicillin therapy.

In the amoxicillin-resistant anaerobic spore-forming community, *Hungatella hathewayi* was the most relatively abundant species at Day 0 (4.55 %) that steadily reduced at each time point (2.04 % at Day 7, 1.92 % at Day 14 and 0.33 % at Day 35). There was also a slight increase in "Other" species overall (94.6 % at Day 0, 96.0 % at Day 7, 95.5 % at Day 14 and 97.8 % at Day 35). However, none of these changes were statistically significant: the amoxicillin treatment did not have a major impact on the relative abundance of amoxicillin-resistant anaerobic spore-forming species. Overall, this suggests the amoxicillin therapy had more of an impact on the relative abundance of species in the amoxicillin-resistant aerobic community than the anaerobic vegetative or anaerobic spore-forming communities.

5.2.5 *Characterisation of strain- and sequence-level changes in gut microbiota following exposure to amoxicillin*

While metagenomics provides interesting insights of particular taxa within a community, individual isolates provide the opportunity to link genotype to experimentally validated phenotype. As earlier discussed, 8838 isolates with 16S rRNA gene sequences were obtained in these experiments and assigned to 367 OTUs. A record was kept of which culture plates and thus conditions, experimental time point and mouse line that each colony was picked from. This allowed the investigation of the impact of amoxicillin therapy on particular OTUs. The 16S rRNA gene sequence data was used to determine which OTUs were picked from agar plates containing 8 mg/L amoxicillin after the mice received antibiotic treatment (Day 7, 14 or 35) but not before (Day 0). OTUs that fit this criteria therefore represent lineages of bacteria that have either increased in relative abundance to above the detection limit or have increased in amoxicillin resistance (from under to over 8mg/L) during the course of the experiment. 14 OTUs fitted this criteria of only being picked from the amoxicillin agar plates after the mice received treatment. Isolates from before and after amoxicillin treatment in the mice of the 14 candidate OTUs of increased amoxicillin resistance were purified and whole genome sequenced (Table 5.2). These whole genome sequences were then searched for in the metascrape data from plates containing amoxicillin at Day 0 using Mash[213] to confirm the observations based on isolated colonies. All OTUs except *Bacteroides stercoris* 99.703 % and *Shigella sonnei* 99.757 % still fitted the criteria. For this reason, these two OTUs plus *Alistipes senegalensis* 98.49 %, *Alistipes shahii* 99.703 % and *Anaerostipes caccae* 99.773 % (isolates of which could not be recovered sufficiently for whole genome sequencing) are excluded from downstream analyses.

Table 5.2. Candidate OTUs of potential increase in amoxicillin resistance following *in vivo* exposure to amoxicillin. OTUs were considered candidates of increased amoxicillin resistance if they were only isolated from agar containing amoxicillin after the mice received amoxicillin (Day 7/14/35) but not before (Day 0). The number of isolates in each OTU at Day 0 (plates without amoxicillin), Day 0 (plates with amoxicillin) and the later time points (Day 7, 14, 35 with amoxicillin) were calculated. Isolates of these OTUs were purified for whole genome sequencing to assess the impact of amoxicillin exposure on the genomes of these particular taxa. Beta-lactam resistance genes and mutations were predicted using ARIBA and CARD (none: no beta-lactam determinants). Phylum key: B: Bacteroidetes, F: Firmicutes; P: Proteobacteria. No WGS indicates where isolates could not be recovered or purified sufficiently for genomic DNA extractions: either the glycerol stocks were contaminated or unviable. Excluded indicates OTUs that did not pass the Mash analysis and were found with 99 % identity in Day 0 metascrapes from plates containing amoxicillin. % in OTU column represents the similarity of the consensus 16S rRNA gene sequence of that OTU to reference 16S rRNA gene sequences for that species.

			Number of isolated colonies			Beta-lactam resistance genotype	
Mouse line	**OTU**	**Phylum**	**Day 0 – Amox**	**Day 0 + Amox**	**Day 7/14/35 + Amox**	**Before**	**After**
7	*Alistipes senegalensis* 98.49 %	B	1	0	1	No WGS	No WGS
7	*Bacteroides stercoris* 99.703 %	B	1	0	9	Excluded	Excluded
7	*[Clostridium] populeti* 94.435 %	F	2	0	2	None	None
7	*Coprobacillus cateniformis* 99.270 %	F	1	0	1	None	None
7	*Coprobacillus cateniformis* 92.211 %	F	6	0	2	None	None
7	*Marvinbryantia formatexigens* 92.942 %	F	0	0	8	None	None
7	*Sporosarcina newyorkensis* 99.710 %	F	9	0	1	None	None
2	*Alistipes finegoldii* 99.851 %	B	4	0	1	None	None
2	*Alistipes shahii* 99.703 %	B	1	0	17	No WGS	*cfxA*
2	*Anaerostipes caccae* 99.773 %	B	2	0	2	None	No WGS
2	*Enterococcus faecium* 99.791 %	F	19	0	48	None	None
2	*Flavonifractor plautii* 99.702 %	F	1	0	22	None	None
2	*Odoribacter splanchnicus* 99.556 %	B	2	0	2	None	*cfxA*
2	*Shigella sonnei* 99.757 %	P	2	0	7	Excluded	Excluded

To assess strain-level genomic changes such as antibiotic resistance gene acquisition, the presence of known, clinically relevant beta-lactam resistance genes and mutations in these genomes was determined using the ARIBA and CARD as in Chapter 3. Table 5.2 summarises the different isolates that were whole genome sequenced for this purpose and whether or not and they were predicted to harbour beta-lactam resistance genes. Genetic determinants of beta-lactam resistance were observed in the genomes isolated after amoxicillin treatment in the mice, but not before, in one candidate OTU (*Odoribacter splanchnicus* 99.56 %). Specifically, isolates of this OTU did not appear to contain the *cfxA* beta-lactamase before amoxicillin treatment, but did afterwards. The 198 genome sequences generated in these mouse experiments were searched for the *cfxA* gene, identifying 46 observations of this gene. All *cfxA* genes from bacteria cultured from Donor 7 mice were identical (Fig. 5.13), and all *cfxA* genes from bacteria cultured from Donor 2 mice were identical, with the only difference between the two groups being a single nucleotide polymorphism at position 775 (A in Donor 7 mouse *cfxA* genes, T in Donor 2 mouse *cfxA* genes). This suggests the *O. splanchnicus* 99.56 % isolates may have acquired the *cfxA beta-lactamase* during the course of the experiment from *Alistipes shahii* 99.703 %, *Bacteroides vulgatus* 88.703 %, or *Bacteroides ovatus* 99.48 %.

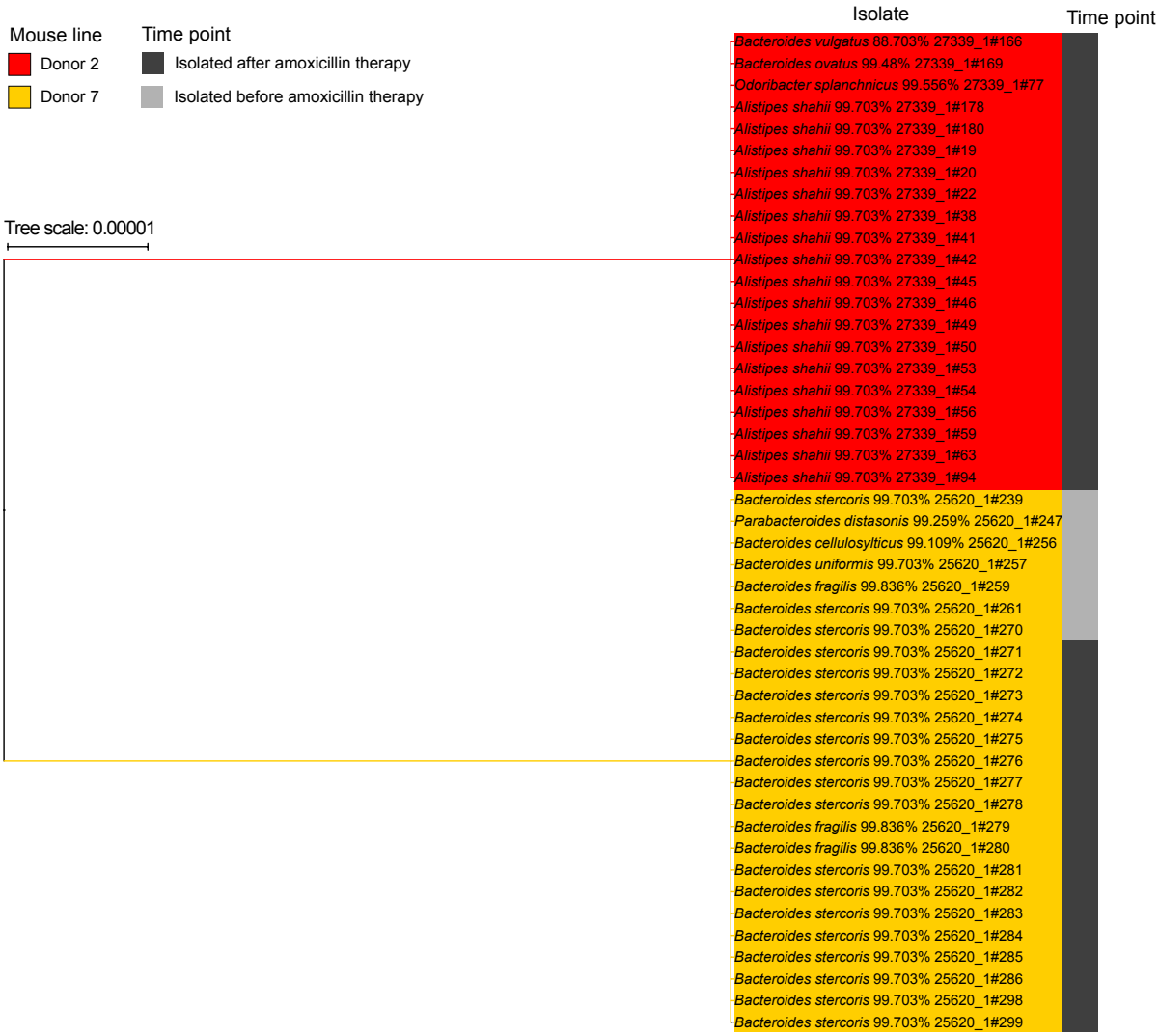


Figure 5.13. A phylogeny of the *cfxA* gene from 46 genomes from isolates cultured from mice with human-derived microbiota. Beta-lactam resistance genes and mutations were predicted using ARIBA[148] with CARD[151,152] in genomes of OTUs considered candidates of increased amoxicillin resistance (only isolated from agar containing amoxicillin after the mice received amoxicillin (Day 7/14/35) but not before (Day 0)) and other genomes generated during this study. This identified *cfxA* genes in 46 isolates; the nucleotide genes were extracted and used to infer a phylogeny, showing that *cfxA* genes were identical across genomes from a particular mouse line. No isolates with the *cfxA* gene were whole genome sequenced from the Donor 2 mouse before the mice received amoxicillin. In the Donor 7 mice, there were a number of isolates with *cfxA* sequenced from before and after the mice received amoxicillin.

The four *O. splanchnicus* isolates were cultured and three well-isolated colonies of each were used for amoxicillin MIC measurement with Etests as described previously. The two *O. splanchnicus* isolates from before amoxicillin treatment both had an average amoxicillin MIC

of 13.3 mg/L; the two isolates from after amoxicillin treatment both had an average amoxicillin of > 256 mg/L (i.e., all biological replicates of these isolates grew at the maximum amoxicillin concentration of the Etest gradient). This suggests the acquisition of the *cfxA* beta-lactamase caused an approximately 19-fold increase in amoxicillin resistance in *O. splanchnicus* 99.56 %. *O. splanchnicus* is considered strictly anaerobic and non-spore-forming[277], therefore the relative abundance of *O. splanchnicus* in the anaerobic vegetative metascrape data from plates containing amoxicillin was determined (Fig. 5.14). This species was the 30th most abundant across these samples, and the mean relative abundance increased from 0.0033 % at Day 0 to 0.022 % at the end of the experiment (Day 35). However, there were no significant differences between the mean relative abundance values of O*. splanchnicus* between consecutive time points (Day 0 vs. Day 7, Day 7 vs. Day 14, Day 14 vs. Day 35), or after either recovery periods compared to the beginning (Day 0 vs. Day 14 and 35, q values determined by unpaired t-tests with Welch's correction[276] and adjusted for multiple-testing with the Benjamini, Hochberg, and Yekutieli method[193]; q value significant < 0.05).

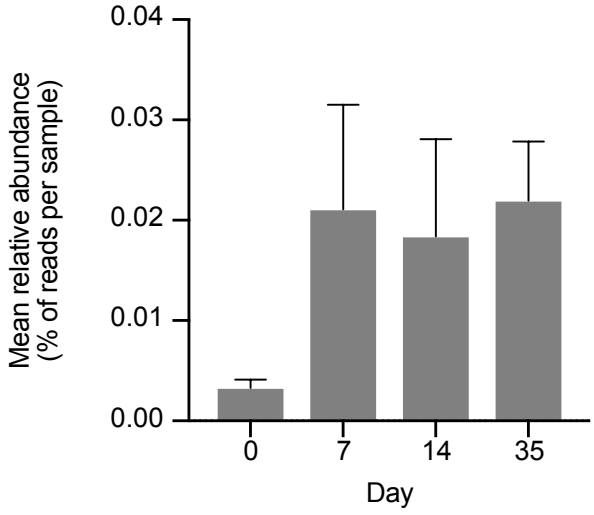


Figure 5.14. Relative abundance of *Odoribacter splanchnicus* in anaerobic vegetative metascrape samples. The relative abundance of the "strictly anaerobic, non-spore-forming" *Odoribacter splanchnicus* species was determined in the anaerobic vegetative metascrapes from plates containing amoxicillin. The relative abundance of this species increased between Day 0 and Day 7 following amoxicillin therapy in mice with human-derived microbiota, decreased slightly at Day 14 and slightly increased again by Day 35. However, these differences were not significant. Q values were determined by unpaired t-tests with Welch's correction[276], adjusted for multiple-testing with the Benjamini, Hochberg, and Yekutieli method[193]; q value significant < 0.05. Bars represent Standard Error of the Mean.

Eight of the candidate OTUs that had whole genome sequences for before and after amoxicillin therapy and were not excluded following the Mash[213] analysis did not appear to have acquired a beta-lactamase that would explain the described observations. Consequently, a comparative genomics approach as described in Chapter 4 was applied to identify potentially novel beta-lactamases. Genes in each OTU that were shared by all isolates cultured on plates containing amoxicillin after the mice were treated with amoxicillin, but absent from isolates of the same OTU cultured on plates not containing amoxicillin from before the mice received therapy, were identified. Genes fitting these criteria were searched for 90 % similarity to beta-lactamase amino acid markers as generated and described in Chapter 4. No genes were found to contain these markers with 90 % identity.

Finally, single nucleotide polymorphisms (SNPs) were searched for in the same eight OTUs as above, plus *Odoribacter splanchnicus* 99.56 %. Paired-end sequencing reads from bacteria isolated after amoxicillin treatment were mapped against assembled contigs from an isolate cultured before amoxicillin treatment. Variant bcf files were produced and used to identify SNPs in coding sequences (open reading frames (ORFs) annotated as such if over 100 bases) and non-coding regions. ORFs and non-coding regions containing SNPs were compared to known protein sequences using BLASTx or to the nr/nt database using BLASTn (Table 5.3). Twenty-two ORFs containing SNPs were most similar to enzymes involved in metabolic processes such as kinases or reductase and three ORFs were most similar to restriction endonucleases, plus one ORF was annotated as an ABC transporter ATP binding protein. However, ten ORFs were annotated as uncharacterised or hypothetical proteins.

Table 5.3. Mutations in open reading frames and non-coding regions of gut bacteria isolated after amoxicillin therapy in mice with human-derived gut microbiota. Raw sequencing reads from bacteria cultured from mice with human-derived gut microbiota after amoxicillin therapy (query samples) were mapped against reference genomes of bacteria from the same OTU that had been isolated before amoxicillin therapy. Variants were identified in open reading frames (ORFs) of more than 100 bases and non-coding regions: NS nonsynonymous; S synonymous; I insertion; SNP single nucleotide polymorphism. The sequence of ORFs and non-coding regions containing mutations were analysed with BLASTx (nr database) to identify the most similar proteins or with BLASTn (nr/nt database) to identify the most similar nucleotide sequences.

O.T.U.	Reference	Query	Feature	Mutations	Top BLAST hit
***[Clostridium] populeti* 94.435 %**	25620_1#243	25620_1#295	ORF	1 S	Sensor histidine kinase [*Eubacterium ramulus*]
			ORF	1 S	Sensor histidine kinase [*Eubacterium ramulus*]
			ORF	1 NS	TPA: beta-glucosidase [Lachnospiraceae bacterium]
			ORF	1 NS	TPA: beta-glucosidase [Lachnospiraceae bacterium]
			ORF	1 NS	Putative uncharacterized protein [*Clostridium* sp. CAG:352]
			Non-coding	1 SNP	No similarity BLASTn nr/nt or BLASTx
			ORF	3 NS	Nitroreductase family protein [*Beduini massiliensis*]
			ORF	4 NS; 1 S	Uncharacterised protein RAG0 10061 [*Rhynchosporium agropyri*]
			ORF	1 S	TPA: phosphoenolpyruvate carboxykinase (ATP) [*Blautia* sp.]
			ORF	1 NS	Phosphoenolpyruvate carboxykinase [ATP] [[*Clostridium*] *clostridioforme* 90A7]
			ORF	1 S	MULTISPECIES: phosphoenolpyruvate carboxykinase (ATP) [Clostridiales]
			ORF	1 S	ABC transporter ATP-binding protein [*Flavonifractor plautii*]
			Non-coding	1 SNP	No similarity BLASTn nr/nt or BLASTx
			ORF	5 NS; 2 S	Gfo/Idh/MocA family oxidoreductase [*Anaerobium acetethylicum*]
			ORF	6 NS; 1 S	Gfo/Idh/MocA family oxidoreductase [*Anaerobium acetethylicum*]
			ORF	1 NS	Glucan 1,6-alpha-glucosidase [*Clostridium* sp. KLE 1755]
			ORF	1 NS	Alpha-glucosidase [*Eisenbergiella massiliensis*]
***Coprobacillus cateniformis* 99.270 %**	25620_1#240	25620_1#297	Non-coding	1 SNP	Helix-turn-helix domain-containing protein [*Massilimicrobiota timonensis*]
***Coprobacillus cateniformis* 99.211 %**	25620_1#293	25620_1#300	ORF	1 S	Hypothetical protein [*Coprobacillus cateniformis*]
***Marvinbryantia formatexigens* 92.942 %**	25620_1#287	25620_1#309	ORF	1 NS	Peptidase [*Anaerotruncus* sp. 1XD22-93]

O.T.U.	Reference	Query	Feature	Mutations	Top BLAST hit
			ORF	1 NS	Peptidase [Lachnospiraceae bacterium] or Penicillin-binding protein A [uncultured *Clostridium* sp.]
***Sporosarcina newyorkensis* 99.710 %**	25620_1#231	25620_1#325	ORF	1 NS	Preprotein translocase subunit SecA [*Sporosarcina newyorkensis*]
			ORF	1 NS; 1 S	TrkH family potassium uptake protein [*Sporosarcina newyorkensis*]
***Sporosarcina newyorkensis* 99.710 %**	25620_1#232	25620_1#325	ORF	1 NS; 1 S	TrkH family potassium uptake protein [*Sporosarcina newyorkensis*]
***Alistipes finegoldii* 99.851 %**	27261_7#7	27339_1#174	ORF	1 NS	Hypothetical protein/restriction endonuclease
			ORF	1 NS	Restriction endonuclease
			ORF	1 I	Leucine-rich repeat domain-containing protein [*Alistipes finegoldii*]
			ORF	1 I	Leucine-rich repeat domain-containing protein [*Alistipes finegoldii*]
			ORF	1 I	Leucine-rich repeat domain-containing protein [*Alistipes finegoldii*]
			ORF	1 I	Uncharacterised protein BN754_01505 [*Alistipes finegoldii* CAG:68]
			ORF	2 I	Glycosyl transferase
			ORF	3 NS; 5 S	Restriction endonuclease
			ORF	2 NS: 5 S	Restriction endonuclease
***Enterococcus faecium* 99.791 %**	27339_1#100	27339_1#138	ORF	1 NS	Heavy metal translocating P-type ATPase [*Enterococcus faecium*]
			ORF	1 NS	MULTISPECIES: aldo/keto reductase [*Enterococcus*]
***Flavonifractor plautii* 99.702 %**	27339_1#70	27339_1#26	ORF	1 NS	No similarity BLASTx; BLASTn similarity to *Flavonifractor* genome
			ORF	1 NS	Acetaldehyde dehydrogenase [*Flavonifractor plautii* ATCC 29863]
			ORF	1 S	Acetaldehyde dehydrogenase [*Flavonifractor plautii* ATCC 29863]

5.3 Discussion

In this chapter I have studied the impact of amoxicillin therapy on amoxicillin resistance in mice with human-derived gut microbiota. Reference genome based metagenomics was used to observe changes to the pre-existing amoxicillin-resistant community that lasted beyond the end of the antibiotic treatment. A recent study on amoxicillin combined with clavulanate therapy in healthy adult humans also found gut microbiota compositional changes immediately after one week's treatment[278]. Two weeks after treatment stopped, these changes were no longer apparent and the gut microbiota communities had reverted back to pre-treatment profiles[278]. However, that study mainly used 16S rRNA sequencing for profiling the gut community and only looked at relative abundances of bacterial families, whereas I have assessed the species level. In addition, the combination of culturing and WGS to generate study-specific reference genome databases for taxonomic classifications means my analysis in this thesis is tailored to the gut microbiomes being studied, providing higher taxonomic resolution. This may explain the disparity between the two studies.

The amoxicillin-resistant community was studied in mice with human-derived microbiota by culturing faecal samples on agar plates, including plates containing amoxicillin at 8 mg/L amoxicillin, a concentration above which pathogenic bacteria can be considered resistant to amoxicillin. This was further divided into the aerobic, anaerobic vegetative and anaerobic spore-forming communities. The key findings from these humanised microbiota mouse experiments are illustrated in Figure 5.15. The amoxicillin therapy appeared to have an impact on the bacterial load of each of these communities, including the reduction of the amount of anaerobic spore-forming bacteria. An overall reduction in the total amount of gut bacteria following amoxicillin has been observed in previous studies[179]. However, in this study

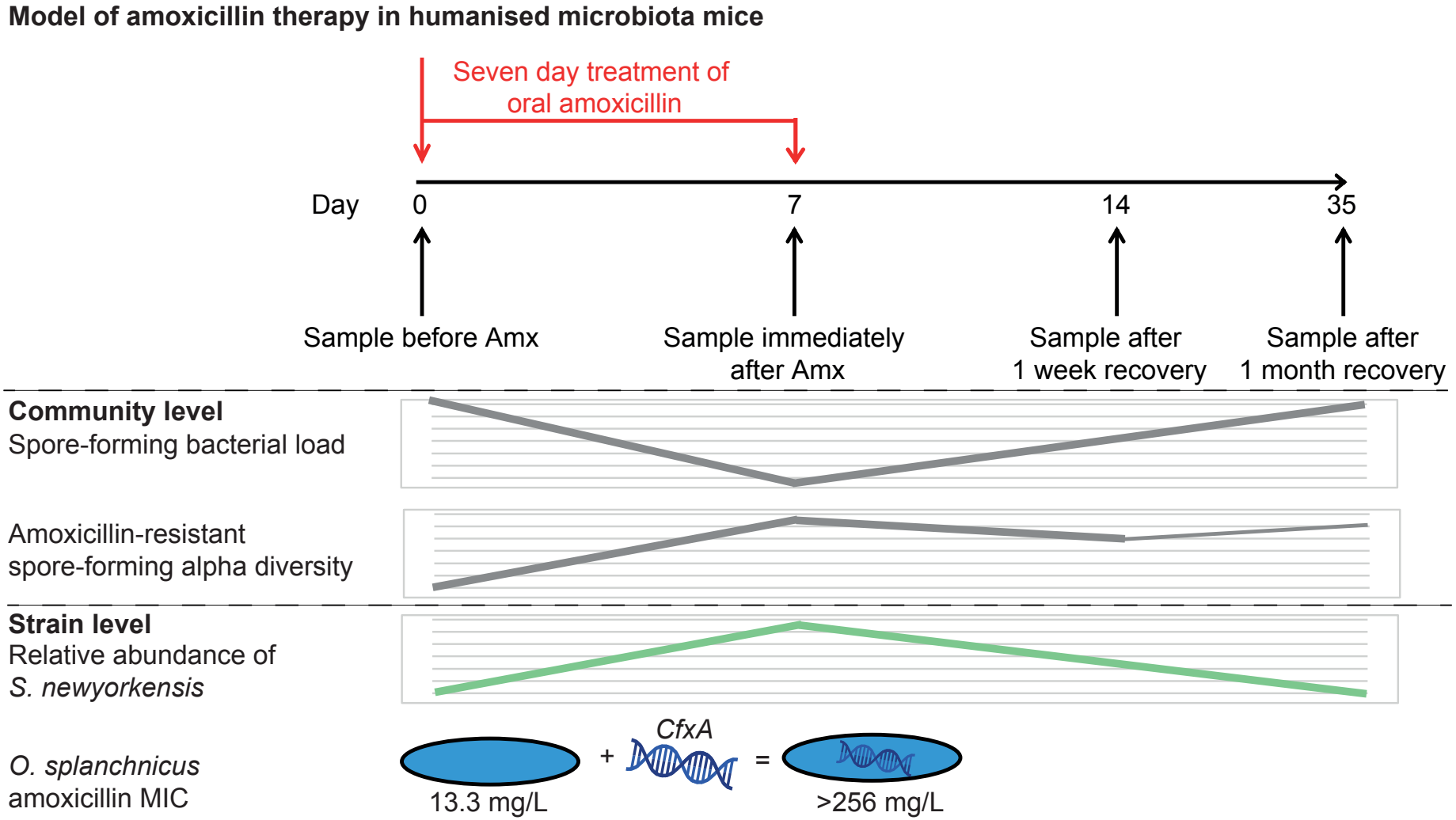


Figure 5.15. The key findings from the humanised microbiota mouse experiments. An illustration of the changes to the amoxicillin-resistant spore-forming community (bacterial load and alpha diversity) and individual strains (relative abundance of *Sporosarcina newyorkensis* and acquisition of *cfxA* by *Odoribacter splanchnicus*). Mice with humanised gut microbiota were given amoxicillin therapy for seven days at a therapeutic concentration. Faecal pellets were collected before and after therapy to allow culturing of individual isolates, WGS and whole genome shotgun metagenomic sequencing of metascrape samples.

there was a slight increase in anaerobic vegetative bacteria – perhaps utilising the space that had been made available by reductions in other bacteria. The amoxicillin-resistant anaerobic spore-forming bacteria also showed significant reductions in load; this indicates that the recommended therapeutic dose of amoxicillin[206] is sufficient to kill some bacteria considered resistant. This is in line with what is expected for amoxicillin – even though it is considered broad spectrum, it targets more Gram-negatives than Gram-positives[31].

That spore-formers, a major but comparatively understudied part of the gut microbiome, are significantly reduced by amoxicillin highlights that the impact of antibiotics on the gut microbiota is not yet fully understood; this result is unexpected given the described amoxicillin spectrum of activity. To my knowledge, this is the first time the spore-forming community has specifically been studied for its response to antibiotics, by using a combination of culturing, WGS and reference genome based metagenomics. Antibiotic treatment can promote spore formation[279] and the assumption could be made that spore-forming bacteria would therefore be protected against significant effects to their total abundance in the gut. However, if the spores re-enter the vegetative state they may not be protected and be impacted by the amoxicillin treatment. This could result in fewer spore-forming bacteria present to produce spores; during the targeted culturing there would then be fewer spores present. After the amoxicillin therapy ends, the spore-formers are able to recover and produce spores again, hypothetically leading to the findings observed in this chapter. In the future, quantitative methods for measuring bacterial load may provide more precise and insightful information[280].

Comparing alpha diversity in the amoxicillin-resistant spore-formers appeared to suggest that the richness of this community increased following amoxicillin treatment. This could be explained by previously susceptible species developing resistance. However, the amoxicillin-resistant anaerobic vegetative and spore-forming communities comprised mostly lowly

abundant or rare species (> 90 % cumulative relative abundance), compared to the amoxicillin-resistant aerobic community. The increase in richness could be due to a reduction in species such as *E. faecalis*, allowing lowly-abundant ones such as *S. newyorkensis* to reach the detectable limit in the metagenomic sequencing analyses and appear more relatively abundant.

The amoxicillin-resistant aerobic community was initially dominated by *E. faecalis*, a Gram-positive facultative anaerobe that is common in the gut[281]. This species was significantly less relatively abundant following amoxicillin therapy. In previous studies of antibiotics on the gut microbiome, enterococci have been observed to expand in the gut following antibiotic therapy[282]; however, this is dependent on the type of antibiotic. Amoxicillin has previously been shown to reduce Gram-positive cocci such as enterococci and increase enterobacteria instead[283,284]. Here, there was a small but insignificant increase in the relative abundance Enterobacteriaceae member *E. coli*. Two *Bacteroides* species (*B. fragilis* and *B. cellulosilyticus*) were dominant at Day 14, and three other *Bacteroides* species (*B. thetaiotaomicron*, *B. uniformis* and *B. vulgatus*) were also in the top ten most relatively abundant species at this time point in aerobic metascrapes. This is surprising given *Bacteroides* species are thought to be obligate anaerobes[148]. Colonies of these five *Bacteroides* species plus *B. faecis*, *B. stercoris* and *B. xylanisolvens* were all isolated from aerobic culture plates, indicating that this was not a "false positive" result in the metagenomic data. *B. fragilis* has been reported as growing in microaerophilic environments (≤ 0.05 % oxygen)[285], and specific DNA mutations have resulted in strains being able to grow in up to 2 % oxygen. In addition, *B. thetaiotaomicron* has been shown to express scavenging enzymes when exposed to oxygen to degrade oxygen compounds and reduce their toxicity[286]. The *Bacteroides* strains growing under aerobic conditions here may have contained or developed mutations that were selected for, allowing

their growth under aerobic conditions. Whether or not this was caused by the amoxicillin treatment or by exposure to oxygen is not determinable in this study. How *Bacteroides* might adapt to aerobic conditions would be interesting to investigate in the future, as this would provide novel insights into their biology.

The most obvious increase in relative abundance in the aerobic metascrapes was that of *S. newyorkensis*. *S. newyorkensis* is another Gram-positive cocci but one that is rare in its ability to form spores; it is also facultatively anaerobic. This may explain its presence and increase in relative abundance in the amoxicillin-resistant aerobic community (Fig. 5.15), following the reduction of *E. faecalis*. *Sporosarcina* species have only recently begun to be studied using genomics[287], with *S. newyorkensis* first being described in 2012[288]. This species was first isolated from human blood[288], which is typically considered a sterile environment and as such, the presence of bacteria in blood indicates infection. Therefore, this study provides insights into the effect of amoxicillin on newly-discovered bacteria, including ones that may have clinical relevance through the potential to cause infections. Another recent study has identified an increase in the relative abundance of a spore-former following beta-lactam treatment: *Candidatus* Borkfalkia ceftriaxoniphila[289] was relatively lowly abundant in the gut prior to ceftriaxone treatment in humans and then dominated the community after ceftriaxone treatment[264]. There, *Candidatus* Borkfalkia ceftriaxoniphila was considered a potential keystone species in the recovery of the gut community following antibiotics; perhaps these temporarily dominant spore-formers aid the re-establishment of a diverse gut microbiome.

In addition to examining the impact of amoxicillin on the amoxicillin-resistant community, I studied the impact on individual strains of gut bacteria. Detecting changes in the genotypic and phenotypic resistance profiles before and after therapeutic antibiotic exposure using

broad identified the acquisition of a known beta-lactamase by a commensal strain of gut bacteria). The *cfxA* beta-lactamase gene is localised on a mobile transposon[290], and since being first described in *B. fragilis*[290], *cfxA* and related genes have also commonly been found in *Prevotella*, *Porphyromonas* and *Capnocytophaga* species[291]. *Prevotella* species are common members of the gut microbiota[292] and *cfxA* in *Prevotella* in particular is associated with increased beta-lactam resistance in periodontal infections[292]. Thus, the presence of *cfxA* in gut microbiota that can act as opportunistic pathogens has potential clinical relevance.

Here, *cfxA* appears to have been acquired by an *O. splanchnicus* strain following amoxicillin therapy, possibly from another Bacteroidetes species, apparently conferring a 19-fold increase in amoxicillin resistance (Fig. 5.15). *O. splanchnicus* has been described as an opportunistic pathogen[277]; if we extrapolate beyond these experiments, the ability of *O. splanchnicus* to acquire beta-lactamases from the gut microbiome therefore has direct potential clinical relevance. A recent porcine model of amoxicillin treatment showed that amoxicillin promoted the enrichment of beta-lactamases including *cfxA*[293]. However, as previously discussed, using whole genome shotgun metagenomics to assess gene prevalence does not allow for the hosts of those genes to be identified, nor can it distinguish whether an increase in gene abundance is caused by the spread of that gene to additional species or the relative increase in abundance of the original host species. Here, I have used broad culturing and WGS to discriminate new hosts of antibiotic resistance genes. Therefore, the unique combination of humanised gnotobiotic mouse models, advanced culturing and genomic analysis in this chapter provides novel and deep understanding of how antibiotic treatment changes the resistance profile of the gut microbiota community and individual strains, with implications relevant to the dissemination of antibiotic resistance and the evolution of drug-resistant infections.

There were other candidate OTUs that appeared to have acquired amoxicillin resistance over the course of the experiment in the absence of known beta-lactam resistance genes or mutations using the methods applied in this study. These OTUs could represent pre-existing amoxicillin-resistant species that have increased in relative abundance above the detection level. Alternatively, it could be due to other resistance mechanisms, such as developing mutations not listed in CARD that increased resistance. This would be in line with the issues predicting resistance in commensal gut bacteria using genomics from the previous chapter, and further highlights how improved annotation of intestinal microbiota would be beneficial. Isolates of these OTUs from after the amoxicillin treatment had SNPs compared to isolates from before the antibiotic therapy. The majority of the ORFs containing variant sites were involved with metabolic processes. One was annotated as a transporter protein, which could potentially be involved in the efflux or export of amoxicillin. In addition, one was similar to a penicillin-binding protein; the observed non-synonymous mutation could have resulted in a reduced affinity for amoxicillin. Several of the ORFs were annotated as hypothetical or uncharacterised proteins; thus their function is unclear. Potentially these may also be candidate novel amoxicillin-resistance genes – though this additionally emphasises the issues surrounding annotation of bacterial genomes and proteins.

Overall, this chapter shows that amoxicillin has significant impacts on the amoxicillin-resistant community of gut bacteria in mice with human-derived microbiota, including relatively uncharacterised spore-forming bacteria. Moreover, a commensal Bacteroidetes strain of *O. splanchnicus* acquired the *cfxA* beta-lactamase gene from the surrounding gut microbiome following exposure to amoxicillin *in vivo* in these mice. Thus, this chapter further emphasises the role of the human gut microbiome as a reservoir for antibiotic resistance in terms of its antibiotic-resistant community- and strain-level changes following antibiotic treatment.

Chapter 6: Discussion

6.1. Key messages and future work

In this thesis, I explored the capacity of the human gut microbiome as a reservoir of antibiotic resistance. I determined that the majority of bacterial members of a normal, healthy gut microbiota possess genetic antibiotic resistance determinants using genomic studies of a unique human gut bacteria culture collection. I also generated antibiotic resistance phenotypes to link to these genotypes and identified the presence of resistances to antibiotics essential in medicine that were not predicted using a range of genome-based databases and methods. To my knowledge, this is the first time that such a large phylogenetically diverse collection of human commensal gut bacteria has been tested for resistance to multiple clinically relevant classes of antibiotics using phenotypic and/or genotypic approaches. This identified multiple instances of antibiotic resistance that could not be predicted using genomics, signifying that the full extent of antibiotic resistance in the microbiome is currently unknown. In addition, I combined deep culturing with reference genome-based metagenomics to study the impact of amoxicillin on the gut bacteria in mice with humanised microbiota, identifying the acquisition of a clinically relevant beta-lactamase by a potential opportunistic gut pathogen via HGT. This unique combination of techniques therefore enabled strain-level discrimination of antibiotic-induced genetic changes. Overall, this thesis highlights the power of a combined genomic and culture-based approach to investigate the gut resistome.

The overall widespread distribution of resistance determinants in the HBC (Chapter 3) is concerning in its implications for the potential horizontal transfer of antibiotic resistance genes among resident and transient members of the gut microbiome. With approximately 52

% of commensal gut bacteria possessing ARGs there are more opportunities for these HGT events to occur than if only a small minority of gut bacteria harboured ARGs. Thus, the majority of the human microbiome is in theory capable of contributing to the spread of antibiotic resistance. Indeed, this was in part demonstrated by the acquisition of the *cfxA* beta-lactamase gene in isolates of *Odoribacter splanchnicus* following exposure to amoxicillin (Chapter 5); this transfer event conferred high-level beta-lactam resistance in the recipient *O. splanchnicus* strain. This shows how antibiotics can lead to the movement of clinically relevant ARGs and cause increased antibiotic resistance in gut bacteria. If this occurs in pathogens or opportunistic pathogens, then this could result in antibiotic-resistant infections. Though *cfxA* has not currently been detected in bacteria belonging to other phyla, possibly due to phyla HGT barriers, *O. splanchnicus* has been described as an opportunistic pathogen and thus this event has potentially direct clinical impacts. Moreover, if *cfxA* were to become located within a more promiscuous mobile element or plasmid, it could hypothetically spread to more distantly related taxa such as Enterobacteriaceae and cause highly beta-lactam resistant infections.

However, antibiotic resistance in commensal gut bacteria can also be a benefit to the gut microbiome: for example, beta-lactamase-producing resident bacteria may be protected against the effects of beta-lactam therapy. The action of the beta-lactamase to degrade the antibiotic may also protect other resident commensal bacteria. Moreover, antibiotic resistant commensals can protect against colonisation by antibiotic-resistant pathogens[294]. As antibiotic resistance in commensals can be considered both an issue and a benefit, it is important in the future to define which antibiotic resistance genes are most likely to spread to opportunistic or obligate pathogens, plus which commensal gut bacteria are most likely to cause opportunistic infections.

That isolates representing novel taxa were also found to contain known antibiotic resistance genes highlights that we cannot fully understand which bacteria carry antibiotic resistance genes until we have identified all bacterial species that inhabit the gastrointestinal tract. One approach for this would be continuing to predict the presence of antibiotic resistance genes and mutations in genomes of newly cultured human gut bacteria as they become publicly available. For example, PATRIC[194] provides a genome annotation services that includes CARD[163,164] for annotating antibiotic resistance genes; NCBI has its own AMRFinderPlus[295] tool for identifying antimicrobial resistance genes in whole genome sequences. These tools could be programmed to run automatically when new bacterial genomes are uploaded, allowing antibiotic resistance genotype to be readily available and curated. An alternative approach for determining which bacteria carry antibiotic resistance genes or mutations would be to search for known genetic antibiotic resistance determinants in metagenome-assembled genomes (MAGs); this may help to increase the phylogenetic diversity being studied. However, MAGs are limited due to the absence of isolates for phenotypic analysis plus can result from mixed-assemblies of closely related strains[296], potentially introducing spurious results. Unless this can be resolved, culturing and studying whole genome sequences remains the most accurate way of surveying the distribution of genetic determinants of antibiotic resistance.

Continuing to culture gut bacteria from stool samples remains imperative for identifying novel isolates and studying the presence of antibiotic resistance genes and mutations. In addition, this will help improve taxonomic classification of metagenomic sequencing. In Chapter 5, despite isolating approximately 11,000 individual isolates and generating study-specific reference genomes there were still metagenomic sequences identified in metascrapes derived from faecal samples that could not be assigned taxa. Isolating additional colonies could be beneficial, though I predicted that around three times as many colonies picked in this

study would need to be isolated to recover all predicted OTUs. Presently, this would require huge manual effort, as culturing anaerobic bacteria is difficult to automate due to the space restrictions inside an anaerobic cabinet. The culturomics approach[297] could be implemented, which uses multiple types of media to select for different bacteria and mass-spectrometry to rapidly identify those bacteria. Targeted culturing such as ethanol selection for spore-formers may help reduce the number of additional colonies that need to be picked; developing more advanced selective culturing may assist further. As technology improves, it will become easier to isolate all the possible bacteria residing in the gut. There is a particular need for high-throughput, broad-scale culturing of gut bacteria from diverse peoples across the globe to recover maximum diversity.

In addition to isolating and studying novel genomes, novel antibiotic resistance genes are also important to discover for improving the annotation of antibiotic resistance in bacterial genomes and metagenomes belonging to diverse microbiomes. As shown in this thesis, rule-based methods are not accurate for predicting phenotypic resistance in gut bacteria, which as discussed could have direct clinical implications. Though other studies have shown higher rates of accuracy for rule-based genomic prediction of antibiotic resistance in pathogens[240,242,248], a recent bioRxiv preprint showed that accuracy could vary between laboratories studying the same WGS datasets of carbapenem-resistant clinical isolates[240,298]. Identifying known point mutations is especially difficult as they could be masked by sequencing errors or consensus sequences[240]. In addition, it can be hard to identify mutations causing antibiotic resistance phenotypes in the first place[299], and so databases of point mutations should be considered particularly incomplete[299].

In Chapter 4, it was shown difficult to identify or validate candidate novel antibiotic resistance genes from individual gut bacteria using functional metagenomics or cloning of a specific

candidate gene. In the future, qPCR or RNASeq of candidate resistance genes or mutations may provide insights into their expression that could explain differing antibiotic resistance phenotypes been closely related bacteria. In addition, integration sites of cloning vectors (such as of the GeneArt synthesised Group2384-carrying plasmid) could be investigated to see if this affects the expression of the insert candidate gene. Alternative recipient strains instead of *E. coli* could be used; however, classic genetics techniques are more limited for gut bacteria than for pathogens and so presently few alternatives exist. Moreover, successful shotgun cloning from pure cultures of isolated bacteria is crucial for better annotating gut bacteria and should continue to be optimised. Developing genetic techniques for gut bacteria will take time but is essential if we are to better understand our gut microbiota. In particular, improving annotation of antibiotic resistance would be useful to study antibiotic resistance on a global scale: it is possible that one day commercial microbiome sequencing services could be combined with surveillance of antibiotic resistance genes to help achieve this.

Discovering novel antibiotic resistance genes and mutations will help improve the accuracy of rule-based antibiotic resistance prediction methods. It is important for these methods to be accurate: if WGS-AST becomes a routine procedure in healthcare, we need to ensure patients are receiving antibiotics that are effective but not causing unnecessary harm to their gut microbiota. False predictions of susceptibility or resistance may result in patients being treated with antibiotics that do not work or with antibiotics that cause significant impacts to their gut microbiota when a narrower-spectrum antibiotic would have been sufficient. Fortunately, rule-based methods have been shown to be more accurate for specific species of pathogenic bacteria[299]. If this type of study is expanded with additional isolates per species or isolation source, comparable to that of studies in pathogens, we may be able to increase the accuracy of genome-based antibiotic resistance predictions in gut bacteria. Studying more

isolates would allow the approach developed in this thesis to define gut bacteria resistant or susceptible to be limited to more specific taxa (such as species rather than across all four phyla studied here), which may further improve WST-AST accuracy. Measuring MIC instead of zone of inhibition would also provide more direct information and enable antibiotic concentration breakpoints to be determined for gut bacteria. Moreover, expanded databases with genetic antibiotic resistance determinants from a wide range of bacterial sources may result in less bias towards pathogenic Proteobacteria. In the future, WGS-AST could become a gold-standard method for inferring phenotypic antibiotic resistance in commensal gut bacteria, as well as clinical isolates of bacterial pathogens.

Expanding the number of bacteria (e.g., the entire HBC) and range of antibiotics (e.g. all of those on the WHO essential list) studied in this thesis would provide further insights into the impact of antibiotic spectrum on the gut microbiome. For example, in Chapter 4 gentamicin was effective against four Firmicutes bacteria considered strict anaerobes. Aminoglycoside uptake is thought to require quinones, lipid-soluble membrane electron carriers required for aerobic respiration[300] – which are typically only present in Alpha-, Beta- and Gammaproteobacteria[300]. These findings highlight how current knowledge of antibiotic spectrum is based on a limited set of bacterial species and that some antibiotics – such as gentamicin – may have under-realised impacts on the gut microbiota. This is important to understand more fully since the use of antibiotics is known to have long term impacts on the gut microbiome with potentially serious implications for our health. This knowledge could eventually be used to establish a framework to guide antibiotic selection in a clinical setting and ultimately guide microbiota-dependent personalised medicine. Furthermore, with research into gut microbiota-based therapeutics advancing rapidly, perhaps in the future a scenario will exist where when taking an antibiotic the patient is also prescribed a specific

probiotic containing bacteria likely to be harmed by that antibiotic or that aid gut microbiome recovery. Mouse models with humanised microbiota are essential for early testing of these proposed probiotics or any gut microbiota-based therapy.

Though the mouse model system described requires optimisation, it resulted in remarkable discrimination that identified the acquisition of a beta-lactamase by a commensal strain of gut bacteria. Eight other OTUs were observed to have gained several SNPs following antibiotic therapy, but these require further investigation and validation to confirm their hypothetical role in increased amoxicillin resistance. Moreover, the model could be easily altered to study other antibiotics, such as those of last resort, or gut microbiomes with different features from alternative human donors. Off-target effects of particular antibiotic therapy (e.g. increased resistance to other antibiotics that were not used as the therapy) could also be studied. These experiments will help to understand the species- and strain-level ecological dynamics of antibiotic resistance that occur in the gut following antibiotic treatment. This information could aid personalised microbiome-based medicine.

6.2. Concluding remarks

High-throughput ‘omics’ studies have been extremely useful for gaining understanding into the gut microbiota and resistome, and microbiomes more generally. In this thesis, whole genome sequencing was used to map the distribution of genetic antibiotic resistance determinants in individual isolates of phylogenetically diverse commensal gut bacteria and identify many unpredicted resistances to antibiotics considered essential in medicine. Moreover, high-throughput culturing and whole genome sequencing identified a 1 in 11,000 horizontal gene transfer event of a clinically relevant beta-lactamase. However, a return to

classical microbiology to complement these techniques would enable better resolution of high-throughput and large-scale sequencing studies, as well as increased understanding of the biology of the individual bacteria themselves. This is critical if we are to fully understand the role of each member of the gut microbiota and how they contribute to the overall microbiome function and therefore our health. Only when this is achieved will understanding the spread of antibiotic resistance, the complete gut resistome, detailed microbiome function and personalised microbiome-based medicine become possible.

Bibliography

1 Society, T. M. *What are antibiotics and how do they work?*, <https://microbiologysociety.org/education-outreach/antibiotics-unearthed/antibiotics-and-antibiotic-resistance/what-are-antibiotics-and-how-do-they-work.html> (2019).

2 MacGregor, R. R. & Graziani, A. L. Oral administration of antibiotics: a rational alternative to the parenteral route. *Clin Infect Dis* **24**, 457-467, (1997).

3 Chandra, N. & Kumar, S. in *Antibiotics and Antibiotics Resistance Genes in Soils: Monitoring, Toxicity, Risk Assessment and Management* (eds Muhammad Zaffar Hashmi, Vladimir Strezov, & Ajit Varma) 1-18 (Springer International Publishing, 2017).

4 Wiese, J. & Imhoff, J. F. Marine bacteria and fungi as promising source for new antibiotics. *Drug Dev Res* **80**, 24-27, (2019).

5 D'Costa, V. M., King, C. E., Kalan, L., Morar, M., Sung, W. W., Schwarz, C., Froese, D., Zazula, G., Calmels, F., Debruyne, R., Golding, G. B., Poinar, H. N. & Wright, G. D. Antibiotic resistance is ancient. *Nature* **477**, 457-461, (2011).

6 Romero, D., Traxler, M. F., Lopez, D. & Kolter, R. Antibiotics as signal molecules. *Chem Rev* **111**, 5492-5505, (2011).

7 Doherty, M., Buchy, P., Standaert, B., Giaquinto, C. & Prado-Cohrs, D. Vaccine impact: Benefits for human health. *Vaccine* **34**, 6707-6714, (2016).

8 Mara, D., Lane, J., Scott, B. & Trouba, D. Sanitation and health. *PLoS Med* **7**, e1000363, (2010).

9 Adedeji, W. A. The treasure called antibiotics. *Ann Ib Postgrad Med* **14**, 56-57, (2016).

10 Ehrlich, P. Über den jetzigen Stand der Chemotherapie. *Berichte der deutschen chemischen Gesellschaft* **42**, 17-47, (1909).

11 Emmerich, R. & Löw, O. Bakteriolytische Enzyme als Ursache der erworbenen Immunität und die Heilung von Infectionskrankheiten durch dieselben. *Med Microbiol Immunol* **31**, 1-65, (1899).

12 Hata, S. & Ehrlich, P. *Die experimentelle Chemotherapie der Spirillosen (Syphilis, Rückfallfieber, Hühnerspirillose, Frambösie), von Paul Ehrlich und S. Hata*. (J. Springer, 1910).

13 Aminov, R. I. Horizontal gene exchange in environmental microbiota. *Front Microbiol* **2**, 158, (2011).

14 Domagk, G. Ein beitrag zur chemotherapie der bakteriellen infektionen. *DMW-Deutsche Medizinische Wochenschrift* **61**, 250-253, (1935).

15 Aminov, R. I. A brief history of the antibiotic era: lessons learned and challenges for the future. *Front Microbiol* **1**, 134-134, (2010).

16 Silver, L. L. Challenges of antibacterial discovery. *Clin Microbiol Rev* **24**, 71-109, (2011).

17 Fleming, A. On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of *B. influenzæ. Br J Exp Pathol* **10**, 226-236, (1929).

18 The state of the world's antibiotics 2015. (Centre for Disease Dynamics, Economics and Policy (CDDEP), Washington DC, USA, 2015).

19 Schatz, A., Bugle, E. & Waksman, S. A. Streptomycin, a substance exhibiting antibiotic activity against Gram-positive and Gram-negative bacteria. *Proc Soc Exp Biol Med* **55**, 66-69, (1944).

20 *Antimicrobial resistance: global report on surveillance.* Report No. 9241564741, (World Health Organisation, France, 2014).

21 *WHO model list of essential medicines, 20th list (March 2017, amended August 2017)*. (World Health Organisation, Geneva, Switzerland, 2017).

22 Gram, C. Ueber die isolirte Farbung der Schizomyceten in Schnitt-und Trockenpraparaten. *Fortschritte der Medicin* **2**, 185-189, (1884).

23 Hentges, D. J. in *Medical Microbiology* (ed S. Baron) (University of Texas Medical Branch at Galveston, 1996).

24 Kapoor, G., Saigal, S. & Elongavan, A. Action and resistance mechanisms of antibiotics: A guide for clinicians. *J Anaesthesiol Clin Pharmacol* **33**, 300-305, (2017).

25 Shaikh, S., Fatima, J., Shakil, S., Rizvi, S. M. & Kamal, M. A. Antibiotic resistance and extended spectrum beta-lactamases: Types, epidemiology and treatment. *Saudi J Biol Sci* **22**, 90-101, (2015).

26 Krause, K. M., Serio, A. W., Kane, T. R. & Connolly, L. E. Aminoglycosides: An Overview. *Cold Spring Harb Perspect Med* **6**2016).

27 Blair, J. M., Webber, M. A., Baylay, A. J., Ogbolu, D. O. & Piddock, L. J. Molecular mechanisms of antibiotic resistance. *Nat Rev Microbiol* **13**, 42-51, (2015).

28 Epling, J. Bacterial conjunctivitis. *BMJ Clin Ed* **2012**, 0704, (2012).

29 Munita, J. M. & Arias, C. A. in *Virulence Mechanisms of Bacterial Pathogens, Fifth Edition* (eds I. T. Kudva *et al.*) (American Society of Microbiology, 2016).

30 Rao, S. *Extended spectrum beta-lactamases – a comprehensive review* (JJM Medical College, India, 2015).

31 Skarpeid, P. L. & Hoye, S. Phenoxymethylpenicillin versus amoxicillin for infections in ambulatory care: a systematic review. *Antibiotics (Basel)* **7**, 81, (2018).

32 *British National Formulary for Children: Amoxicillin*, <https://www.evidence.nhs.uk/formulary/bnfc/current/5-infections/51-antibacterial-drugs/511-penicillins/5113-broad-spectrum-penicillins/amoxicillin> (2014).

33 Rawof, S. & Upadhye, S. Antibiotics for acute otitis media: Which children are likely to benefit? *CJEM* **11**, 553-557, (2015).

34 Bermingham, A. & Derrick, J. P. The folic acid biosynthesis pathway in bacteria: evaluation of potential for antibacterial drug discovery. *Bioessays* **24**, 637-648, (2002).

35 Zhang, T., Xu, Y., Wang, C., Liu, Z., Wu, T., Cao, Y., Njire, M., Julius, M., Zhou, Y., Guo, J., Wang, B. & Liu, T. Role of *folP1* and *folP2* genes in the action of sulfamethoxazole and trimethoprim against mycobacteria. *J Microbiol Biotechnol* **25**, 1559-1567, (2015).

36 Sköld, O. Sulfonamide resistance: mechanisms and trends. *Drug Resist Updat* **3**, 155-160, (2000).

37 Lofmark, S., Edlund, C. & Nord, C. E. Metronidazole is still the drug of choice for treatment of anaerobic infections. *Clin Infect Dis* **50 Suppl 1**, S16-23, (2010).

38 Shweta & Prakash, S. K. Dental abscess: A microbiological review. *Dent Res J (Isfahan)* **10**, 585-591, (2013).

39 Huttner, A., Verhaegh, E. M., Harbarth, S., Muller, A. E., Theuretzbacher, U. & Mouton, J. W. Nitrofurantoin revisited: a systematic review and meta-analysis of controlled trials. *J Antimicrob Chemother* **70**, 2456-2464, (2015).

40 Cross, R., Ling, C., Day, N. P., McGready, R. & Paris, D. H. Revisiting doxycycline in pregnancy and early childhood--time to rebuild its reputation? *Expert Opin Drug Saf* **15**, 367-382, (2016).

41 *WHO list of critically important antimicrobials for human medicine (WHO CIA list).* (World Health Organization, Geneva, Switzerland, 2019).

42 Smith, C. J., Madinier, I., Fosse, T., Guidicelli, J. & Labia, R. What constitutes an extended-spectrum beta-lactamase? *Antimicrob Agents Chemother* **46**, 600-601, (2002).

43 Bassetti, M., Vena, A., Croxatto, A., Righi, E. & Guery, B. How to manage *Pseudomonas aeruginosa* infections. *Drugs Context* **7**, 212527, (2018).

44 Klein, N. C. & Cunha, B. A. Third-generation cephalosporins. *Med Clin North Am* **79**, 705-719, (1995).

45 Papp-Wallace, K. M., Endimiani, A., Taracila, M. A. & Bonomo, R. A. Carbapenems: past, present, and future. *Antimicrob Agents Chemother* **55**, 4943-4960, (2011).

46 Meletis, G., Exindari, M., Vavatsi, N., Sofianou, D. & Diza, E. Mechanisms responsible for the emergence of carbapenem resistance in *Pseudomonas aeruginosa*. *Hippokratia* **16**, 303-307, (2012).

47 Bradley, J. S., Jackson, M. A., Committee on Infectious Diseases & American Academy of Pediatrics. The use of systemic and topical fluoroquinolones. *Pediatrics* **128**, e1034-1045, (2011).

48 Yarlagadda, V., Manjunath, G. B., Sarkar, P., Akkapeddi, P., Paramanandham, K., Shome, B. R., Ravikumar, R. & Haldar, J. Glycopeptide antibiotic to overcome the intrinsic resistance of Gram-negative bacteria. *ACS Infect Dis* **2**, 132-139, (2016).

49 Leclercq, R. Mechanisms of resistance to macrolides and lincosamides: nature of the resistance elements and their clinical implications. *Clin Infect Dis* **34**, 482-492, (2002).

50 Smieja, M. Current indications for the use of clindamycin: A critical review. *Can J Infect Dis* **9**, 22-28, (1998).

51 Kirst, H. A. in *Macrolide Antibiotics* (eds Wolfgang Schönfeld & Herbert A. Kirst) 1-13 (Birkhäuser, 2002).

52 Giamarellou, H. Fourth generation cephalosporins in the antimicrobial chemotherapy of surgical infections. *J Chemother* **11**, 486-493, (1999).

53 Chapman, T. M. & Perry, C. M. Cefepime: a review of its use in the management of hospitalized patients with pneumonia. *Am J Respir Med* **2**, 75-107, (2003).

54 Duplessis, C. & Crum-Cianflone, N. F. Ceftaroline: A new cephalosporin with activity against Methicillin-Resistant Staphylococcus aureus (MRSA). *Clin Med Rev Ther* **3**, a2466, **(**2011).

55 Sykes, R. B. & Bonner, D. P. Aztreonam: The first monobactam. *Am J Med* **78**, 2-10, (1985).

56 Thorne, G. M. & Alder, J. Daptomycin: a novel lipopeptide antibiotic. *Clin Microbiol Newsletter* **24**, 33-40, (2002).

57 Streit, J. M., Jones, R. N. & Sader, H. S. Daptomycin activity and spectrum: a worldwide sample of 6737 clinical Gram-positive organisms. *J Antimicrob Chemother* **53**, 669-674, (2004).

58 Baltz, R. H. Daptomycin: mechanisms of action and resistance, and biosynthetic engineering. *Curr Opin Chem Biol* **13**, 144-151, (2009).

59 Tran, T. T., Munita, J. M. & Arias, C. A. Mechanisms of drug resistance: daptomycin resistance. *Ann N Y Acad Sci* **1354**, 32-53, (2015).

60 Michalopoulos, A. S., Livaditis, I. G. & Gougoutas, V. The revival of fosfomycin. *Int J Infect Dis* **15**, e732-e739, (2011).

61 Matthews, P. C., Barrett, L. K., Warren, S., Stoesser, N., Snelling, M., Scarborough, M. & Jones, N. Oral fosfomycin for treatment of urinary tract infection: a retrospective cohort study. *BMC infectious diseases* **16**, 556, (2016).

62 Falagas, M. E., Athanasaki, F., Voulgaris, G. L., Triarides, N. A. & Vardakas, K. Z. Resistance to fosfomycin: Mechanisms, frequency and clinical consequences. *Int J Antimicrob Agents* **53**, 22-28, (2019).

63 Greer, N. D. Tigecycline (Tygacil): the first in the glycylcycline class of antibiotics. *Proc (Bayl Univ Med Cent)* **19**, 155-161, (2006).

64 Pankey, G. A. Tigecycline. *J Antimicrob Chemother* **56**, 470-480, (2005).

65 Sum, P. E., Sum, F. W. & Projan, S. J. Recent developments in tetracycline antibiotics. *Current pharmaceutical design* **4**, 119-132, (1998).

66 Shinabarger, D. L., Marotti, K. R., Murray, R. W., Lin, A. H., Melchior, E. P., Swaney, S. M., Dunyak, D. S., Demyan, W. F. & Buysse, J. M. Mechanism of action of oxazolidinones: effects of linezolid and eperezolid on translation reactions. *Antimicrob Agents Chemother* **41**, 2132-2136, (1997).

67 Eliopoulos, G. M., Meka, V. G. & Gold, H. S. Antimicrobial resistance to linezolid. *Clin Infect Dis* **39**, 1010-1015, (2004).

68 Yu, Z., Qin, W., Lin, J., Fang, S. & Qiu, J. Antibacterial mechanisms of polymyxin and bacterial resistance. *Biomed Res Int* **2015**, 11 pages, (2015).

69 Morrill, H. J., Pogue, J. M., Kaye, K. S. & LaPlante, K. L. Treatment options for carbapenem-resistant Enterobacteriaceae infections. *Open Forum Infect Dis* **2**, ofv050, (2015).

70 McCarthy, M. Woman dies after infection with bacteria resistant to all antibiotics available in US. *BMJ* **356**, j254, (2017).

71 Coico, R. Gram Staining. *Curr Protoc Microbiol* **Appendix 3**, A.3C.1-A.3C.2, (2006).

72 van Hoek, A. H., Mevius, D., Guerra, B., Mullany, P., Roberts, A. P. & Aarts, H. J. Acquired antibiotic resistance genes: an overview. *Front Microbiol* **2**, 203, (2011).

73 Baquero, F., Alvarez-Ortega, C. & Martinez, J. L. Ecology and evolution of antibiotic resistance. *Environ Microbiol Rep* **1**, 469-476, (2009).

74 Clark, D. J. & Maaløe, O. DNA replication and the division cycle in *Escherichia coli*. *J Mol Biol* **23**, 99-112, (1967).

75 Llewelyn, M. J., Fitzpatrick, J. M., Darwin, E., SarahTonkin-Crine, Gorton, C., Paul, J., Peto, T. E. A., Yardley, L., Hopkins, S. & Walker, A. S. The antibiotic course has had its day. *BMJ* **358**, j3418, (2017).

76 Livanos, A. E., Greiner, T. U., Vangay, P., Pathmasiri, W., Stewart, D., McRitchie, S., Li, H., Chung, J., Sohn, J., Kim, S., Gao, Z., Barber, C., Kim, J., Ng, S., Rogers, A. B., Sumner, S., Zhang, X. S., Cadwell, K., Knights, D., Alekseyenko, A., Backhed, F. & Blaser, M. J. Antibiotic-mediated gut microbiome perturbation accelerates development of type 1 diabetes in mice. *Nat Microbiol* **1**, 16140, (2016).

77 Johnson, C. M. & Grossman, A. D. Integrative and Conjugative Elements (ICEs): What they do and how they work. *Ann Rev Genet* **49**, 577-601, (2015).

78 Modi, S. R., Collins, J. J. & Relman, D. A. Antibiotics and the gut microbiota. *J Clin Invest* **124**, 4212-4218, (2014).

79 Hastings, P. J., Rosenberg, S. M. & Slack, A. Antibiotic-induced lateral transfer of antibiotic resistance. *Trends Microbiol* **12**, 401-404, (2004).

80 O'Neill, J. Tackling drug-resistant infections globally: Final report and recommendations. (HM Government, The Wellcome Trust, London, 2016).

81 Nga do, T. T., Chuc, N. T., Hoa, N. P., Hoa, N. Q., Nguyen, N. T., Loan, H. T., Toan, T. K., Phuc, H. D., Horby, P., Van Yen, N., Van Kinh, N. & Wertheim, H. F. Antibiotic sales in rural and urban pharmacies in northern Vietnam: an observational study. *BMC Pharmacol Toxicol* **15**2014).

82 Baker, S. A return to the pre-antimicrobial era? *Science* **347**, 1064-1066, (2015).

83 Rabaa, M. A., Tue, N. T., Phuc, T. M., Carrique-Mas, J., Saylors, K., Cotten, M., Bryant, J. E., Nghia, H. D., Cuong, N. V., Pham, H. A., Berto, A., Phat, V. V., Dung, T. T., Bao, L. H., Hoa, N. T., Wertheim, H., Nadjm, B., Monagin, C., van Doorn, H. R., Rahman, M., Tra, M. P., Campbell, J. I., Boni, M. F., Tam, P. T., van der Hoek, L., Simmonds, P., Rambaut, A., Toan, T. K., Van Vinh Chau, N., Hien, T. T., Wolfe, N., Farrar, J. J., Thwaites, G., Kellam, P., Woolhouse, M. E. & Baker, S. The Vietnam Initiative on Zoonotic Infections (VIZIONS): A strategic approach to studying emerging zoonotic infectious diseases. *Ecohealth* **12**, 726-735, (2015).

84 Thompson, C. N., Phan, M. V., Hoang, N. V., Minh, P. V., Vinh, N. T., Thuy, C. T., Nga, T. T., Rabaa, M. A., Duy, P. T., Dung, T. T., Phat, V. V., Nga, T. V., Tu le, T. P., Tuyen, H. T., Yoshihara, K., Jenkins, C., Duong, V. T., Phuc, H. L., Tuyet, P. T., Ngoc, N. M., Vinh, H., Chinh, N. T., Thuong, T. C., Tuan, H. M., Hien, T. T., Campbell, J. I., Chau, N. V., Thwaites, G. & Baker, S. A prospective multi-center observational study of children hospitalized with diarrhea in Ho Chi Minh City, Vietnam. *Am J Trop Med Hyg* **92**, 1045-1052, (2015).

85 Antibiotic resistance threats in the United States. (Centers for Disease Control and Prevention (CDC), Atlanta, 2013).

86 Klevens, R. M., Edwards, J. R., Richards, C. L., Horan, T. C., Gaynes, R. P., Pollock, D. A. & Cardo, D. M. Estimating health care-associated infections and deaths in U.S. hospitals, 2002. *Public Health Rep* **122**, 160-166, (2007).

87 Santajit, S. & Indrawattana, N. Mechanisms of antimicrobial resistance in ESKAPE pathogens. *Biomed Res Int* **2016**, 2475067, (2016).

88 Taylor, L. H., Latham, S. M. & Woolhouse, M. E. Risk factors for human disease emergence. *Philos Trans R Soc Lond B Biol Sci* **356**, 983-989, (2001).

89 *Antibiotic-resistant priority pathogens list* (World Health Organisation, Geneva, Switzerland, 2017).

90 Vital signs: carbapenem-resistant Enterobacteriaceae. *MMWR. Morbidity and mortality weekly report* **62**, 165-170, (2013).

91 Berglund, F., Marathe, N. P., Osterlund, T., Bengtsson-Palme, J., Kotsakis, S., Flach, C. F., Larsson, D. G. J. & Kristiansson, E. Identification of 76 novel B1 metallo-beta-lactamases through large-scale screening of genomic and metagenomic data. *Microbiome* **5**, 134, (2017).

92 Argaw-Denboba, A., Abejew, A. A. & Mekonnen, A. G. Antibiotic-resistant bacteria are major threats of otitis media in Wollo area, northeastern Ethiopia: A ten-year retrospective analysis. *Int J Microbiol* **2016**, 8724671, (2016).

93 Jacoby, G. A. AmpC beta-lactamases. *Clin Microbiol Rev* **22**, 161-182, Table of Contents, (2009).

94 Canton, R., Gonzalez-Alba, J. M. & Galan, J. C. CTX-M enzymes: Origin and diffusion. *Front Microbiol* **3**, 110, (2012).

95 Evans, B. A. & Amyes, S. G. OXA beta-lactamases. *Clin Microbiol Rev* **27**, 241-263, (2014).

96 Akova, M., Daikos, G. L., Tzouvelekis, L. & Carmeli, Y. Interventional strategies and current clinical experience with carbapenemase-producing Gram-negative bacteria. *Clin Microbiol Infect* **18**, 439-448, (2012).

97 Walsh, F. The multiple roles of antibiotics and antibiotic resistance in nature. *Front Microbiol* **4**, 255, (2013).

98 Kempf, I., Fleury, M. A., Drider, D., Bruneau, M., Sanders, P., Chauvin, C., Madec, J. Y. & Jouy, E. What do we know about resistance to colistin in Enterobacteriaceae in avian and pig production in Europe? *Int J Antimicrob Agents* **42**, 379-383, (2013).

99 Liu, Y.-Y., Wang, Y., Walsh, T. R., Yi, L.-X., Zhang, R., Spencer, J., Doi, Y., Tian, G., Dong, B., Huang, X., Yu, L.-F., Gu, D., Ren, H., Chen, X., Lv, L., He, D., Zhou, H., Liang, Z., Liu, J.-H. & Shen, J. Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. *Lancet Infec Dis* **16**, 161-168, (2016).

100 Wang, X., Wang, Y., Zhou, Y., Wang, Z., Wang, Y., Zhang, S. & Shen, Z. Emergence of colistin resistance gene mcr-8 and its variant in *Raoultella ornithinolytica*. *Front Microbiol* **10**2019).

101 Zhang, H., Hou, M., Xu, Y., Srinivas, S., Huang, M., Liu, L. & Feng, Y. Action and mechanism of the colistin resistance enzyme MCR-4. *Commun Biol* **2**, 36, (2019).

102 Cantas, L., Shah, S. Q., Cavaco, L. M., Manaia, C. M., Walsh, F., Popowska, M., Garelick, H., Burgmann, H. & Sorum, H. A brief multi-disciplinary review on antimicrobial resistance in medicine and its linkage to the global environmental microbiota. *Front Microbiol* **4**, 96, (2013).

103 Fleming, A., Chain, E. & Florey, H. in *URL http://www.nobelprize.org/nobel_prizes/medicine/laureates/1945/fleming-lecture.html* (1945).

104 Costelloe, C., Metcalfe, C., Lovering, A., Mant, D. & Hay, A. D. Effect of antibiotic prescribing in primary care on antimicrobial resistance in individual patients: systematic review and meta-analysis. *BMJ* **340**, c2096, (2010).

105 Uranga, A., España, P. P., Bilbao, A., Quintana, J. M., Arriaga, I., Intxausti, M., Lobo, J. L., Tomás, L., Camino, J., Nuñez, J. & Capelastegui, A. Duration of antibiotic treatment in community-acquired pneumonia: A multicenter randomized clinical trial. *JAMA Intern Med* **176**, 1257-1265, (2016).

106 Abeles, S. R., Jones, M. B., Santiago-Rodriguez, T. M., Ly, M., Klitgord, N., Yooseph, S., Nelson, K. E. & Pride, D. T. Microbial diversity in individuals and their household contacts following typical antibiotic courses. *Microbiome* **4**, 39, (2016).

107 Ventola, C. L. The Antibiotic Resistance Crisis. Part 1: Causes and Threats. *P T* **40**, 277-283, (2015).

108 Hutchinson, J. M., Patrick, D. M., Marra, F., Ng, H., Bowie, W. R., Heule, L., Muscat, M. & Monnet, D. L. Measurement of antibiotic consumption: A practical guide to the use of the Anatomical Therapeutic

Chemical classification and Defined Daily Dose system methodology in Canada. *Can J Infect Dis* **15**, 29-35, (2004).

109 *ATC/DDD Index*, <https://www.whocc.no/atc_ddd_index/?code=J01CR02&showdescription=yes> (2018).

110 WHO report on surveillance of antibiotic consumption: 2016-2018 early implementation. Report No. 9241514884, (World Health Organisation, Geneva, Switzerland, 2018).

111 Karabay, O. Attitudes and behaviors of family physicians regarding use of antibiotics. *JMID* **01**, 53-57, (2011).

112 *Turkey takes strong action to reduce antibiotic consumption and resistance*, <http://www.euro.who.int/en/countries/turkey/news/news/2017/11/turkey-takes-strong-action-to-reduce-antibiotic-consumption-and-resistance> (2017).

113 Robinson, T. P., Bu, D. P., Carrique-Mas, J., Fevre, E. M., Gilbert, M., Grace, D., Hay, S. I., Jiwakanon, J., Kakkar, M., Kariuki, S., Laxminarayan, R., Lubroth, J., Magnusson, U., Thi Ngoc, P., Van Boeckel, T. P. & Woolhouse, M. E. Antibiotic resistance is the quintessential One Health issue. *Trans R Soc Trop Med Hyg* **110**, 377-380, (2016).

114 Association, A. V. M. One Health Initiative Task Force - Final Report. 2008).

115 Manyi-Loh, C., Mamphweli, S., Meyer, E. & Okoh, A. Antibiotic use in agriculture and its consequential resistance in environmental sources: Potential public health implications. *Molecules* **23**, 795, (2018).

116 Dibner, J. J. & Richards, J. D. Antibiotic growth promoters in agriculture: history and mode of action. *Poultry Science* **84**, 634-643, (2005).

117 Meek, R. W., Vyas, H. & Piddock, L. J. Nonmedical uses of antibiotics: Time to restrict their use? *PLoS Biol* **13**, e1002266, (2015).

118 Swann, M. M. Report of The Joint Committee on the Use of Antibiotics in Animal Husbandry Veterinary Medicine. (HM Stationery Office, London, UK, 1969).

119 Founou, L. L., Founou, R. C. & Essack, S. Y. Antibiotic resistance in the food chain: A developing country-perspective. *Front Microbiol* **7**2016).

120 Gouliouris, T., Raven, K. E., Moradigaravand, D., Ludden, C., Coll, F., Blane, B., Naydenova, P., Horner, C., Brown, N. M., Corander, J., Limmathurotsakul, D., Parkhill, J. & Peacock, S. J. Detection of vancomycin-resistant *Enterococcus faecium* hospital-adapted lineages in municipal wastewater treatment plants indicates widespread distribution and release into the environment. *Genome Res* **29**, 626-634, (2019).

121 Raven, K. E., Ludden, C., Gouliouris, T., Blane, B., Naydenova, P., Brown, N. M., Parkhill, J. & Peacock, S. J. Genomic surveillance of *Escherichia coli* in municipal wastewater treatment plants as an indicator of clinically relevant pathogens and their resistance genes. *Microb Genom* **5, (**2019).

122 Leonard, A. F. C., Zhang, L., Balfour, A. J., Garside, R., Hawkey, P. M., Murray, A. K., Ukoumunne, O. C. & Gaze, W. H. Exposure to and colonisation by antibiotic-resistant *E. coli* in UK coastal water users: Environmental surveillance, exposure assessment, and epidemiological study (Beach Bum Survey). *Environ Int* **114**, 326-333, (2018).

123 Salyers, A. A., Gupta, A. & Wang, Y. Human intestinal bacteria as reservoirs for antibiotic resistance genes. *Trends Microbiol* **12**, 412-416, (2004).

124 Bergstrom, A., Skov, T. H., Bahl, M. I., Roager, H. M., Christensen, L. B., Ejlerskov, K. T., Molgaard, C., Michaelsen, K. F. & Licht, T. R. Establishment of intestinal microbiota during early life: a longitudinal, explorative study of a large cohort of Danish infants. *Appl Environ Microbiol* **80**, 2889-2900, (2014).

125 Rajilic-Stojanovic, M. & de Vos, W. M. The first 1000 cultured species of the human gastrointestinal microbiota. *FEMS Microbiol Rev* **38**, 996-1047, (2014).

126 Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207-214, (2012).

127 Barka, E. A., Vatsa, P., Sanchez, L., Gaveau-Vaillant, N., Jacquard, C., Meier-Kolthoff, J. P., Klenk, H. P., Clement, C., Ouhdouch, Y. & van Wezel, G. P. Taxonomy, physiology, and natural products of Actinobacteria. *Microbiol Mol Biol Rev* **80**, 1-43, (2016).

128 Hahnke, R. L., Meier-Kolthoff, J. P., Garcia-Lopez, M., Mukherjee, S., Huntemann, M., Ivanova, N. N., Woyke, T., Kyrpides, N. C., Klenk, H. P. & Goker, M. Genome-based taxonomic classification of Bacteroidetes. *Front Microbiol* **7**, 2003, (2016).

129 Shin, N. R., Whon, T. W. & Bae, J. W. Proteobacteria: microbial signature of dysbiosis in gut microbiota. *Trends Biotechnol* **33**, 496-503, (2015).

130 Eckburg, P. B., Bik, E. M., Bernstein, C. N., Purdom, E., Dethlefsen, L., Sargent, M., Gill, S. R., Nelson, K. E. & Relman, D. A. Diversity of the human intestinal microbial flora. *Science* **308**, 1635, (2005).

131 Gensollen, T., Iyer, S. S., Kasper, D. L. & Blumberg, R. S. How colonization by microbiota in early life shapes the immune system. *Science* **352**, 539-544, (2016).

132 Lloyd-Price, J., Abu-Ali, G. & Huttenhower, C. The healthy human microbiome. *Genome Med* **8**, 51, (2016).

133 Lawley, T. D., Clare, S., Walker, A. W., Stares, M. D., Connor, T. R., Raisen, C., Goulding, D., Rad, R., Schreiber, F., Brandt, C., Deakin, L. J., Pickard, D. J., Duncan, S. H., Flint, H. J., Clark, T. G., Parkhill, J. & Dougan, G. Targeted restoration of the intestinal microbiota with a simple, defined bacteriotherapy resolves relapsing *Clostridium difficile* disease in mice. *PLoS Pathog* **8**, e1002995, (2012).

134 Korpela, K., Salonen, A., Virta, L. J., Kekkonen, R. A., Forslund, K., Bork, P. & de Vos, W. M. Intestinal microbiome is related to lifetime antibiotic use in Finnish pre-school children. *Nat Commun* **7**, 10410, (2016).

135 Johanesen, P. A., Mackin, K. E., Hutton, M. L., Awad, M. M., Larcombe, S., Amy, J. M. & Lyras, D. Disruption of the gut microbiome: *Clostridium difficile* infection and the threat of antibiotic resistance. *Genes (Basel)* **6**, 1347-1360, (2015).

136 Baron, S. A., Diene, S. M. & Rolain, J.-M. Human microbiomes and antibiotic resistance. *Human Microbiome Journal* **10**, 43-52, (2018).

137 Forslund, K., Sunagawa, S., Kultima, J. R., Mende, D. R., Arumugam, M., Typas, A. & Bork, P. Country-specific antibiotic use practices impact the human gut resistome. *Genome Res* **23**, 1163-1169, (2013).

138 Blake, D. P., Hillman, K., Fenlon, D. R. & Low, J. C. Transfer of antibiotic resistance between commensal and pathogenic members of the Enterobacteriaceae under ileal conditions. *J Appl Microbiol* **95**, 428-436, (2003).

139 Sommer, M. O., Dantas, G. & Church, G. M. Functional characterisation of the antibiotic resistance reservoir in the human microflora. *Science* **325**, 1128-1131, (2009).

140 Hu, Y., Yang, X., Qin, J., Lu, N., Cheng, G., Wu, N., Pan, Y., Li, J., Zhu, L., Wang, X., Meng, Z., Zhao, F., Liu, D., Ma, J., Qin, N., Xiang, C., Xiao, Y., Li, L., Yang, H., Wang, J., Yang, R., Gao, G. F., Wang, J. & Zhu, B. Metagenome-wide analysis of antibiotic resistance genes in a large cohort of human gut microbiota. *Nat Commun* **4**, 2151, (2013).

141 Moore, A. M., Patel, S., Forsberg, K. J., Wang, B., Bentley, G., Razia, Y., Qin, X., Tarr, P. I. & Dantas, G. Pediatric fecal microbiota harbor diverse and novel antibiotic resistance genes. *PLoS One* **8**, e78822, (2013).

142 Bengtsson-Palme, J., Angelin, M., Huss, M., Kjellqvist, S., Kristiansson, E., Palmgren, H., Larsson, D. G. & Johansson, A. The human gut microbiome as a transporter of antibiotic resistance genes between continents. *Antimicrob Agents Chemother* **59**, 6551-6560, (2015).

143 Pehrsson, E. C., Tsukayama, P., Patel, S., Mejia-Bautista, M., Sosa-Soto, G., Navarrete, K. M., Calderon, M., Cabrera, L., Hoyos-Arango, W., Bertoli, M. T., Berg, D. E., Gilman, R. H. & Dantas, G. Interconnected microbiomes and resistomes in low-income human habitats. *Nature* **533**, 212-216, (2016).

144 Sommer, M. O., Church, G. M. & Dantas, G. The human microbiome harbors a diverse reservoir of antibiotic resistance genes. *Virulence* **1**, 299-303, (2010).

145 Zhang, L., Calvo-Bado, L., Murray, A. K., Amos, G. C. A., Hawkey, P. M., Wellington, E. M. & Gaze, W. H. Novel clinically relevant antibiotic resistance genes associated with sewage sludge and industrial waste streams revealed by functional metagenomic screening. *Environ Int* **132**, 105120, (2019).

146 Browne, H. P., Forster, S. C., Anonye, B. O., Kumar, N., Neville, B. A., Stares, M. D., Goulding, D. & Lawley, T. D. Culturing of 'unculturable' human microbiota reveals novel taxa and extensive sporulation. *Nature* **533**, 543-546, (2016).

147 Leclercq, R., Canton, R., Brown, D. F., Giske, C. G., Heisig, P., MacGowan, A. P., Mouton, J. W., Nordmann, P., Rodloff, A. C., Rossolini, G. M., Soussy, C. J., Steinbakk, M., Winstanley, T. G. & Kahlmeter, G. EUCAST expert rules in antimicrobial susceptibility testing. *Clin Microbiol Infect* **19**, 141-160, (2013).

148 Wexler, H. M. *Bacteroides*: the good, the bad, and the nitty-gritty. *Clin Microbiol Rev* **20**, 593-621, (2007).

149 Rettedal, E. A., Gumpert, H. & Sommer, M. O. Cultivation-based multiplex phenotyping of human gut microbiota allows targeted recovery of previously uncultured bacteria. *Nat Commun* **5**, 4714, (2014).

150 Lagier, J. C., Khelaifia, S., Alou, M. T., Ndongo, S., Dione, N., Hugon, P., Caputo, A., Cadoret, F., Traore, S. I., Seck, E. H., Dubourg, G., Durand, G., Mourembou, G., Guilhot, E., Togo, A., Bellali, S., Bachar, D., Cassir, N., Bittar, F., Delerce, J., Mailhe, M., Ricaboni, D., Bilen, M., Dangui Nieko, N. P., Dia Badiane, N. M., Valles, C., Mouelhi, D., Diop, K., Million, M., Musso, D., Abrahao, J., Azhar, E. I., Bibi, F., Yasir, M., Diallo, A., Sokhna, C., Djossou, F., Vitton, V., Robert, C., Rolain, J. M., La Scola, B., Fournier, P. E., Levasseur, A. & Raoult, D. Culture of previously uncultured members of the human gut microbiota by culturomics. *Nat Microbiol* **1**, 16203, (2016).

151 Forster, S. C., Kumar, N., Anonye, B. O., Almeida, A., Viciani, E., Stares, M. D., Dunn, M., Mkandawire, T. T., Zhu, A., Shao, Y., Pike, L. J., Louie, T., Browne, H. P., Mitchell, A. L., Neville, B. A., Finn, R. D. & Lawley, T. D. A human gut bacterial genome and culture collection for improved metagenomic analyses. *Nat Biotechnol* **37**, 186-192, (2019).

152 Hu, Y., Yang, X., Li, J., Lv, N., Liu, F., Wu, J., Lin, I. Y., Wu, N., Weimer, B. C., Gao, G. F., Liu, Y. & Zhu, B. The transfer network of bacterial mobile resistome connecting animal and human microbiome. *Appl Environ Microbiol* **82**, 6672-6681, (2016).

153 Gladstone, R. A., Lo, S. W., Lees, J. A., Croucher, N. J., van Tonder, A. J., Corander, J., Page, A. J., Marttinen, P., Bentley, L. J., Ochoa, T. J., Ho, P. L., du Plessis, M., Cornick, J. E., Kwambana-Adams, B., Benisty, R., Nzenze, S. A., Madhi, S. A., Hawkins, P. A., Everett, D. B., Antonio, M., Dagan, R., Klugman, K. P., von Gottberg, A., McGee, L., Breiman, R. F., Bentley, S. D. & Global Pneumococcal Sequencing, C. International genomic definition of pneumococcal lineages, to contextualise disease, antibiotic resistance and vaccine impact. *EBioMedicine* **43**, 338-346, (2019).

154 Duranti, S., Lugli, G. A., Mancabelli, L., Turroni, F., Milani, C., Mangifesta, M., Ferarrio, C., Anzalone, R., Viappiani, A., van Sinderen, D. & Ventura, M. Prevalence of antibiotic resistance genes among human gut-derived *Bifidobacteria*. *Appl Environ Microbiol* **83**, e02894-02816, (2017).

155 Wright, G. D. The antibiotic resistome: the nexus of chemical and genetic diversity. *Nat Rev Microbiol* **5**, 175-186, (2007).

156 Pal, C., Bengtsson-Palme, J., Kristiansson, E. & Larsson, D. G. The structure and diversity of human, animal and environmental resistomes. *Microbiome* **4**, 54, (2016).

157 Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., Mende, D. R., Li, J., Xu, J., Li, S., Li, D., Cao, J., Wang, B., Liang, H., Zheng, H., Xie, Y., Tap, J., Lepage, P., Bertalan, M., Batto, J.-M., Hansen, T., Le Paslier, D., Linneberg, A., Nielsen, H. B., Pelletier, E., Renault, P., Sicheritz-Ponten, T., Turner, K., Zhu, H., Yu, C., Li, S., Jian, M., Zhou, Y., Li, Y., Zhang, X., Li, S., Qin, N., Yang, H., Wang, J., Brunak, S., Doré, J., Guarner, F., Kristiansen, K., Pedersen, O., Parkhill, J., Weissenbach, J., Meta, H. I. T. C., Bork, P., Ehrlich, S. D. & Wang, J. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59-65, (2010).

158 van Schaik, W. The human gut resistome. *Philos Trans R Soc Lond B Biol Sci* **370**, 20140087, (2015).

159 Liu, B. & Pop, M. ARDB--Antibiotic Resistance Genes Database. *Nucleic Acids Res* **37**, D443-447, (2009).

160 Hunt, M., Mather, A. E., Sánchez-Busó, L., Page, A. J., Parkhill, J., Keane, J. A. & Harris, S. R. ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads. *Microb Genom* **3**, e000131, (2017).

161 Gupta, S. K., Padmanabhan, B. R., Diene, S. M., Lopez-Rojas, R., Kempf, M., Landraud, L. & Rolain, J. M. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrob Agents Chemother* **58**, 212-220, (2014).

162 Bush, K. & Jacoby, G. A. Updated functional classification of beta-lactamases. *Antimicrob Agents Chemother* **54**, 969-976, (2010).

163 McArthur, A. G., Waglechner, N., Nizam, F., Yan, A., Azad, M. A., Baylay, A. J., Bhullar, K., Canova, M. J., De Pascale, G., Ejim, L., Kalan, L., King, A. M., Koteva, K., Morar, M., Mulvey, M. R., O'Brien, J. S., Pawlowski, A. C., Piddock, L. J., Spanogiannopoulos, P., Sutherland, A. D., Tang, I., Taylor, P. L., Thaker, M., Wang, W., Yan, M., Yu, T. & Wright, G. D. The comprehensive antibiotic resistance database. *Antimicrob Agents Chemother* **57**, 3348-3357, (2013).

164 Jia, B., Raphenya, A. R., Alcock, B., Waglechner, N., Guo, P., Tsang, K. K., Lago, B. A., Dave, B. M., Pereira, S., Sharma, A. N., Doshi, S., Courtot, M., Lo, R., Williams, L. E., Frye, J. G., Elsayegh, T., Sardar, D., Westman, E. L., Pawlowski, A. C., Johnson, T. A., Brinkman, F. S., Wright, G. D. & McArthur, A. G. CARD

2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res* **45**, D566-D573, (2017).

165 Arango-Argoty, G., Garner, E., Pruden, A., Heath, L. S., Vikesland, P. & Zhang, L. DeepARG: a deep learning approach for predicting antibiotic resistance genes from metagenomic data. *Microbiome* **6**, 23, (2018).

166 Thai, Q. K., Bos, F. & Pleiss, J. The Lactamase Engineering Database: a critical survey of TEM sequences in public databases. *BMC Genomics* **10**, 390, (2009).

167 Lakin, S. M., Dean, C., Noyes, N. R., Dettenwanger, A., Ross, A. S., Doster, E., Rovira, P., Abdo, Z., Jones, K. L., Ruiz, J., Belk, K. E., Morley, P. S. & Boucher, C. MEGARes: an antimicrobial resistance database for high throughput sequencing. *Nucleic Acids Res* **45**, D574-D580, (2017).

168 Zankari, E., Allesøe, R., Joensen, K. G., Cavaco, L. M., Lund, O. & Aarestrup, F. M. PointFinder: a novel web tool for WGS-based detection of antimicrobial resistance associated with chromosomal point mutations in bacterial pathogens. *J Antimicrob Chemother*2017).

169 Gibson, M. K., Forsberg, K. J. & Dantas, G. Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. *ISME J* **9**, 207-216, (2015).

170 Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., Aarestrup, F. M. & Larsen, M. V. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother* **67**, 2640-2644, (2012).

171 Inouye, M., Dashnow, H., Raven, L.-A., Schultz, M. B., Pope, B. J., Tomita, T., Zobel, J. & Holt, K. E. SRST2: Rapid genomic surveillance for public health and hospital microbiology labs. *Genome Med* **6**2014).

172 Yin, X., Jiang, X. T., Chai, B., Li, L., Yang, Y., Cole, J. R., Tiedje, J. M. & Zhang, T. ARGs-OAP v2.0 with an expanded SARG database and Hidden Markov Models for enhancement characterization and quantification of antibiotic resistance genes in environmental metagenomes. *Bioinformatics* **34**, 2263-2270, (2018).

173 Zhou, W., Wang, Y. & Lin, J. Functional cloning and characterization of antibiotic resistance genes from the chicken gut microbiome. *Appl Environ Microbiol* **78**, 3028-3032, (2012).

174 Clemente, J. C., Pehrsson, E. C., Blaser, M. J., Sandhu, K., Gao, Z., Wang, B., Magris, M., Hidalgo, G., Contreras, M., Noya-Alarcon, O., Lander, O., McDonald, J., Cox, M., Walter, J., Oh, P. L., Ruiz, J. F., Rodriguez, S., Shen, N., Song, S. J., Metcalf, J., Knight, R., Dantas, G. & Dominguez-Bello, M. G. The microbiome of uncontacted Amerindians. *Science advances* **1**, e1500183, (2015).

175 Hatosy, S. M. & Martiny, A. C. The ocean as a global reservoir of antibiotic resistance genes. *Appl Environ Microbiol* **81**, 7593-7599, (2015).

176 Allen, H. K., Moe, L. A., Rodbumrer, J., Gaarder, A. & Handelsman, J. Functional metagenomics reveals diverse beta-lactamases in a remote Alaskan soil. *ISME J* **3**, 243-251, (2009).

177 Thompson, L. R., Sanders, J. G., McDonald, D., Amir, A., Ladau, J., Locey, K. J., Prill, R. J., Tripathi, A., Gibbons, S. M., Ackermann, G., Navas-Molina, J. A., Janssen, S., Kopylova, E., Vazquez-Baeza, Y., Gonzalez, A., Morton, J. T., Mirarab, S., Zech Xu, Z., Jiang, L., Haroon, M. F., Kanbar, J., Zhu, Q., Jin Song, S., Kosciolek, T., Bokulich, N. A., Lefler, J., Brislawn, C. J., Humphrey, G., Owens, S. M., Hampton-Marcell, J., Berg-Lyons, D., McKenzie, V., Fierer, N., Fuhrman, J. A., Clauset, A., Stevens, R. L., Shade, A., Pollard,

K. S., Goodwin, K. D., Jansson, J. K., Gilbert, J. A., Knight, R. & Earth Microbiome Project, C. A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* **551**, 457-463, (2017).

178 Dethlefsen, L., Huse, S., Sogin, M. L. & Relman, D. A. The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS Biol* **6**, e280, (2008).

179 Antonopoulos, D. A., Huse, S. M., Morrison, H. G., Schmidt, T. M., Sogin, M. L. & Young, V. B. Reproducible community dynamics of the gastrointestinal microbiota following antibiotic perturbation. *Infect Immun* **77**, 2367-2375, (2009).

180 Raymond, F., Ouameur, A. A., Deraspe, M., Iqbal, N., Gingras, H., Dridi, B., Leprohon, P., Plante, P. L., Giroux, R., Berube, E., Frenette, J., Boudreau, D. K., Simard, J. L., Chabot, I., Domingo, M. C., Trottier, S., Boissinot, M., Huletsky, A., Roy, P. H., Ouellette, M., Bergeron, M. G. & Corbeil, J. The initial state of the human gut microbiome determines its reshaping by antibiotics. *ISME J* **10**, 707-720, (2016).

181 Iizumi, T., Battaglia, T., Ruiz, V. & Perez Perez, G. I. Gut Microbiome and Antibiotics. *Arch Med Res* **48**, 727-734, (2017).

182 Yarza, P., Yilmaz, P., Pruesse, E., Glockner, F. O., Ludwig, W., Schleifer, K. H., Whitman, W. B., Euzeby, J., Amann, R. & Rossello-Mora, R. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol* **12**, 635-645, (2014).

183 Page, A. J., de Silva, N., Hunt, M., Quail, M. A., Parkhill, J., Harris, S. R., Otto, T. D. & Keane, J. A. Robust high throughout prokaryote de novo assembly and improvement pipeline for Illumina data. *bioRxiv* (2016).

184 Sunagawa, S., Mende, D. R., Zeller, G., Izquierdo-Carrasco, F., Berger, S. A., Kultima, J. R., Coelho, L. P., Arumugam, M., Tap, J., Nielsen, H. B., Rasmussen, S., Brunak, S., Pedersen, O., Guarner, F., de Vos, W. M., Wang, J., Li, J., Dore, J., Ehrlich, S. D., Stamatakis, A. & Bork, P. Metagenomic species profiling using universal phylogenetic marker genes. *Nat Methods* **10**, 1196-1199, (2013).

185 Katoh, K., Misawa, K., Kuma, K. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **30**, 3059-3066, (2002).

186 Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30**, 772-780, (2013).

187 Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* **26**, 1641-1650, (2009).

188 Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2 – approximately maximum-likelihood trees for large alignments. *PloS One* **5**, e9490, (2009).

189 Ondov, B. D., Bergman, N. H. & Phillippy, A. M. Interactive metagenomic visualization in a Web browser. *BMC Bioinformatics* **12**, 385, (2011).

190 Tukey, J. W. *Exploratory Data Analysis: Limited Preliminary Ed*. Vol. 2 (Addison-Wesley Publishing Company, 1970).

191 Letunic, I. & Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res* **44**, W242-245, (2016).

192 Fisher, R. A. *Statistical methods for research workers*. 16 edn, (Genesis Publishing Pvt Ltd, 2006).

193 Benjamini, Y., Krieger, A. M. & Yekutieli, D. Adaptive linear step-up procedures that control the false discovery rate. *Biometrika* **93**, 491-507, (2006).

194 Wattam, A. R., Davis, J. J., Assaf, R., Boisvert, S., Brettin, T., Bun, C., Conrad, N., Dietrich, E. M., Disz, T., Gabbard, J. L., Gerdes, S., Henry, C. S., Kenyon, R. W., Machi, D., Mao, C., Nordberg, E. K., Olsen, G. J., Murphy-Olson, D. E., Olson, R., Overbeek, R., Parrello, B., Pusch, G. D., Shukla, M., Vonstein, V., Warren, A., Xia, F., Yoo, H. & Stevens, R. L. Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center. *Nucleic Acids Res* **45**, D535-D542, (2017).

195 Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* **25**, 1043-1055, (2015).

196 Duncan, S. H. Growth requirements and fermentation products of Fusobacterium prausnitzii, and a proposal to reclassify it as *Faecalibacterium prausnitzii* gen. nov., comb. nov. *Int J Syst Evol Microbiol* **52**, 2141-2146, (2002).

197 Rice, P., Longden, I. & Bleasby, A. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* **16**, 276-277, (2000).

198 Team, R. C. (R Foundation for Statistical Computing, Vienna, Austria, 2013).

199 Wickham, H. *ggplot2: elegant graphics for data analysis*. 2 edn, (Springer, 2016).

200 Jain, C., Rodriguez, R. L., Phillippy, A. M., Konstantinidis, K. T. & Aluru, S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* **9**, 5114, (2018).

201 Page, A. J., Cummins, C. A., Hunt, M., Wong, V. K., Reuter, S., Holden, M. T., Fookes, M., Falush, D., Keane, J. A. & Parkhill, J. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **31**, 3691-3693, (2015).

202 Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792-1797, (2004).

203 Edgar, R. C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**, 113, (2004).

204 Gouy, M., Guindon, S. & Gascuel, O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* **27**, 221-224, (2009).

205 Chang, A. C. & Cohen, S. N. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J Bacteriol* **134**, 1141-1156, (1978).

206 Embree, J. High dose amoxicillin: rationale for use in otitis media treatment failture. *Peaediatr Child Health* **4**, 321-323, (1999).

207 Nawrocki, E. *Structural RNA homology search and alignment using covariance models* PhD thesis, Washington University, (2009).

208 Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., Sahl, J. W., Stres, B., Thallinger, G. G., Van Horn, D. J. & Weber, C. F. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75**, 7537-7541, (2009).

209 Kaminski, J., Gibson, M. K., Franzosa, E. A., Segata, N., Dantas, G. & Huttenhower, C. High-specificity targeted functional profiling in microbial communities with ShortBRED. *PLoS Comput Biol* **11**, e1004557, (2015).

210 Carver, T., Harris, S. R., Berriman, M., Parkhill, J. & McQuillan, J. A. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics* **28**, 464-469, (2012).

211 Wood, D. E. & Salzberg, S. L. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* **15**, R46, **(**2014).

212 Mann, H. B. & Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Statist.* **18**, 50-60, (1947).

213 Ondov, B. D., Treangen, T. J., Melsted, P., Mallonee, A. B., Bergman, N. H., Koren, S. & Phillippy, A. M. Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol* **17**, 132, (2016).

214 Penders, J., Stobberingh, E. E., Savelkoul, P. H. & Wolffs, P. F. The human microbiome as a reservoir of antimicrobial resistance. *Front Microbiol* **4**, 87, (2013).

215 Feng, J., Li, B., Jiang, X., Yang, Y., Wells, G. F., Zhang, T. & Li, X. Antibiotic resistome in a large-scale healthy human gut microbiota deciphered by metagenomic and network analyses. *Environ Microbiol* **20**, 355-368, (2018).

216 Almeida, A., Mitchell, A. L., Boland, M., Forster, S. C., Gloor, G. B., Tarkowska, A., Lawley, T. D. & Finn, R. D. A new genomic blueprint of the human gut microbiota. *Nature* **568**, 499-504, (2019).

217 Yassour, M., Vatanen, T., Siljander, H., Hamalainen, A.-M., Harkonen, T., Ryhanen, S. J., Franzosa, E. A., Vlamakis, H., Huttenhower, C., Gevers, D., Lander, E. S., Knip, M. & Xavier, R. J. Natural history of the infant gut microbiome and impact of antibiotic treatment on bacterial strain diversity and stability. *Science* **8**, 343-381, (2016).

218 Schmidt, K., Mwaigwisya, S., Crossman, L. C., Doumith, M., Munroe, D., Pires, C., Khan, A. M., Woodford, N., Saunders, N. J., Wain, J., O'Grady, J. & Livermore, D. M. Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore-based metagenomic sequencing. *J Antimicrob Chemother* **72**, 104-114, (2017).

219 Malbruny, B., Werno, A. M., Murdoch, D. R., Leclercq, R. & Cattoir, V. Cross-resistance to lincosamides, streptogramins A, and pleuromutilins due to the lsa(C) gene in *Streptococcus agalactiae* UCN70. *Antimicrob Agents Chemother* **55**, 1470-1474, (2011).

220 Davin-Regli, A. & Pages, J. M. Enterobacter aerogenes and Enterobacter cloacae; versatile bacterial pathogens confronting antibiotic treatment. *Front Microbiol* **6**, 392, (2015).

221 Singh, L., Cariappa, M. P. & Kaur, M. *Klebsiella oxytoca*: An emerging pathogen? *Med J Armed Forces India* **72**, S59-S61, (2016).

222 Guzman Prieto, A. M., van Schaik, W., Rogers, M. R., Coque, T. M., Baquero, F., Corander, J. & Willems, R. J. Global emergence and dissemination of Enterococci as nosocomial pathogens: Attack of the clones? *Front Microbiol* **7**, 788, (2016).

223 Telenti, A., Imboden, P., Marchesi, F., Lowrie, D., Cole, S., Colston, M. J., Matter, L., Schopfer, K. & Bodmer, T. Detection of rifampicin-resistance mutations in *Mycobacterium tuberculosis*. *Lancet* **341**, 647-650, (1993).

224 Sterling, T. R., Villarino, M. E., Borisov, A. S., Shang, N., Gordin, F., Bliven-Sizemore, E., Hackman, J., Hamilton, C. D., Menzies, D., Kerrigan, A., Weis, S. E., Weiner, M., Wing, D., Conde, M. B., Bozeman, L.,

Horsburgh, C. R. & Chaisson, R. E. Three months of rifapentine and isoniazid for latent tuberculosis infection. *New England Journal of Medicine* **365**, 2155-2166, (2011).

225 Malik, M., Li, L., Zhao, X., Kerns, R. J., Berger, J. M. & Drlica, K. Lethal synergy involving bicyclomycin: an approach for reviving old antibiotics. *J Antimicrob Chemother* **69**, 3227-3235, (2014).

226 Kwon, D. H. & Lu, C. D. Polyamine effects on antibiotic susceptibility in bacteria. *Antimicrob Agents Chemother* **51**, 2070-2077, (2007).

227 Fleeman, R. M., Debevec, G., Antonen, K., Adams, J. L., Santos, R. G., Welmaker, G. S., Houghten, R. A., Giulianotti, M. A. & Shaw, L. N. Identification of a novel polyamine scaffold with potent efflux pump Inhibition activity toward nulti-drug resistant Bacterial pathogens. *Front Microbiol* **9**, 1301, (2018).

228 Yoon, B. K., Jackman, J. A., Valle-Gonzalez, E. R. & Cho, N. J. Antibacterial free fatty acids and monoglycerides: Biological activities, experimental testing, and therapeutic applications. *Int J Mol Sci* **19**, E1114, (2018).

229 Björkman, J. & Andersson, D. I. The cost of antibiotic resistance from a bacterial perspective. *Drug Resistance Updates* **3**, 237-245, (2000).

230 Andreani, N. A., Hesse, E. & Vos, M. Prokaryote genome fluidity is dependent on effective population size. *ISME J* **11**, 1719-1721, (2017).

231 Coque, T. E., Singh, K. V., Weinstock, G. M. & Murray, B. E. Characterisation of dihydrofolate reductase genes from trimethoprim-susceptible and trimethoprim-resistant strains of *Enteroccus faecalis*. *Antimicrob Agents Chemother* **43**, 141-147, (1999).

232 Nakano, V., Silva, A. d. N. e., Merino, V. R. C., Wexler, H. M. & Avila-Campos, M. J. Antimicrobial resistance and prevalence of resistance genes in intestinal Bacteroidales strains. *Clinics (Sao Paulo)* **66**, 543-547, (2011).

233 Shoemaker, N. B., Vlamakis, H., Hayes, K. & Salyers, A. A. Evidence for extensive resistance gene transfer among *Bacteroides* spp. and among *Bacteroides* and other genera in the human colon. *Appl Environ Microbiol* **67**, 561-568, (2001).

234 Mazel, D. Integrons and the origin of antibiotic resistance gene cassettes. *ASM News-American Society for Microbiology* **70**, 520-525, (2004).

235 Fair, R. J. & Tor, Y. Antibiotics and bacterial resistance in the 21st century. *Perspect Medicin Chem* **6**, 25-64, (2014).

236 Annavajhala, M. K., Gomez-Simmonds, A. & Uhlemann, A. C. Multidrug-resistant Enterobacter cloacae complex emerging as a global, diversifying threat. *Front Microbiol* **10**, 44, (2019).

237 Partridge, S. R. Analysis of antibiotic resistance regions in Gram-negative bacteria. *FEMS Microbiol Rev* **35**, 820-855, (2011).

238 Foster, T. J. Antibiotic resistance in *Staphylococcus aureus*. Current status and future prospects. *FEMS Microbiol Rev* **41**, 430-449, (2017).

239 Schumacher, A., Vranken, T., Malhotra, A., Arts, J. J. C. & Habibovic, P. *In vitro* antimicrobial susceptibility testing methods: agar dilution to 3D tissue-engineered models. *Eur J Clin Microbiol Infect Dis* **37**, 187-208, (2018).

240 Su, M., Satola, S. W. & Read, T. D. Genome-based prediction of bacterial antibiotic resistance. *J Clin Microbiol* **57**, e01405-01418, (2019).

241 Kos, V. N., Deraspe, M., McLaughlin, R. E., Whiteaker, J. D., Roy, P. H., Alm, R. A., Corbeil, J. & Gardner, H. The resistome of *Pseudomonas aeruginosa* in relationship to phenotypic susceptibility. *Antimicrob Agents Chemother* **59**, 427-436, (2015).

242 Drouin, A., Letarte, G., Raymond, F., Marchand, M., Corbeil, J. & Laviolette, F. Interpretable genotype-to-phenotype classifiers with performance guarantees. *Sci Rep* **9**, 4071, (2019).

243 Wainwright, M. Acridine—a neglected antibacterial chromophore. *J Antimicrob Chemother* **47**, 1-13, (2001).

244 Zhang, Y., Bao, Q., Gagnon, L. A., Huletsky, A., Oliver, A., Jin, S. & Langaee, T. ampG gene of *Pseudomonas aeruginosa* and its role in beta-lactamase expression. *Antimicrob Agents Chemother* **54**, 4772-4779, (2010).

245 Beam, T. R., Jr. Ceftriaxone: a beta-lactamase-stable, broad-spectrum cephalosporin with an extended half-life. *Pharmacotherapy* **5**, 237-253, (1985).

246 Bhalodi, A. A., van Engelen, T. S. R., Virk, H. S. & Wiersinga, W. J. Impact of antimicrobial therapy on the gut microbiome. *J Antimicrob Chemother* **74**, i6-i15, (2019).

247 Sagheddu, V., Patrone, V., Miragoli, F., Puglisi, E. & Morelli, L. Infant early gut colonization by Lachnospiraceae: High frequency of *Ruminococcus gnavus*. *Front Pediatr* **4**, 57, (2016).

248 Pesesky, M. W., Hussain, T., Wallace, M., Patel, S., Andleeb, S., Burnham, C. D. & Dantas, G. Evaluation of machine learning and rules-based approaches for predicting antimicrobial Resistance profiles in Gram-negative Bacilli from whole genome sequence data. *Front Microbiol* **7**, 1887, (2016).

249 Higdon, R., Louie, B. & Kolker, E. Modeling sequence and function similarity between proteins for protein functional annotation. *Proc Int Symp High Perform Distrib Comput* **2010**, 499-502, (2010).

250 Brolund, A., Lagerqvist, N., Byfors, S., Struelens, M. J., Monnet, D. L., Albiger, B., Kohlenberg, A. & European Antimicrobial Resistance Genes Surveillance Network, E.-N. C. S. G. Worsening epidemiological situation of carbapenemase-producing Enterobacteriaceae in Europe, assessment by national experts from 37 countries, July 2018. *Euro Surveill* **24**2019).

251 Hong, H.-J., Hutchings, M. I. & Buttner, M. J. in *Bacterial Signal Transduction: Networks and Drug Targets* (ed Ryutaro Utsumi) 200-213 (Springer New York, 2008).

252 Moran, R. A., Anantham, S., Holt, K. E. & Hall, R. M. Prediction of antibiotic resistance from antibiotic resistance genes detected in antibiotic-resistant commensal *Escherichia coli* using PCR or WGS. *J Antimicrob Chemother* **72**, 700-704, (2016).

253 Li, L., Zhou, Q., Voss, T. C., Quick, K. L. & LaBarbera, D. V. High-throughput imaging: Focusing in on drug discovery in 3D. *Methods* **96**, 97-102, (2016).

254 Baltekin, Ö., Boucharin, A., Tano, E., Andersson, D. I. & Elf, J. Antibiotic susceptibility testing in less than 30 min using direct single-cell imaging. *PNAS*, 201708558, (2017).

255 de Vries, L. E., Valles, Y., Agerso, Y., Vaishampayan, P. A., Garcia-Montaner, A., Kuehl, J. V., Christensen, H., Barlow, M. & Francino, M. P. The gut as reservoir of antibiotic resistance: microbial diversity of tetracycline resistance in mother and infant. *PLoS One* **6**, e21644, (2011).

256 Cheng, G., Hu, Y., Yin, Y., Yang, X., Xiang, C., Wang, B., Chen, Y., Yang, F., Lei, F., Wu, N., Lu, N., Li, J., Chen, Q., Li, L. & Zhu, B. Functional screening of antibiotic resistance genes from human gut microbiota reveals a novel gene fusion. *FEMS Microbiol Lett* **336**, 11-16, (2012).

257 Balis, E., Vatopoulos, A. C., Kanelopoulou, M., Mainas, E., Haztoudis, G., Kontogianni, V., Malamou-Lada, H., Kitsou-Kiriakopoulou, S. & Kalopothaki, V. Indications of *in vivo* transfer of an epidemic R plasmid from *Samonella enteritidis* to *Escherichia coli* of the normal human gut flora. *J Clin Microbiol* **34**, 977-979, (1996).

258 Gruzza, M., Fons, M. D.-I., Y. & Duculuzeau, R. Study of gene transfer in vitro and in the digestive tract of gnotobiotic mice from *Lactococcus lactis* strains to various strains belonging to human intestinal flora. *Microb Releases* **2**, 1830189, (1994).

259 Forsberg, K. J., Reyes, A., Wang, B., Selleck, E. M., Sommer, M. O. & Dantas, G. The shared antibiotic resistome of soil bacteria and human pathogens. *Science* **337**, 1107-1111, (2012).

260 Croswell, A., Amir, E., Teggatz, P., Barman, M. & Salzman, N. H. Prolonged impact of antibiotics on intestinal microbial ecology and susceptibility to enteric *Salmonella* infection. *Infect Immun* **77**, 2741, (2009).

261 Rodrigues, R. R., Greer, R. L., Dong, X., Dsouza, K. N., Gurung, M., Wu, J. Y., Morgun, A. & Shulzhenko, N. Antibiotic-induced alterations in gut microbiota are associated with changes in glucose metabolism in healthy mice. *Front Microbiol* **8**, 2306-2306, (2017).

262 Oldenburg, C. E., Sié, A., Coulibaly, B., Ouermi, L., Dah, C., Tapsoba, C., Bärnighausen, T., Ray, K. J., Zhong, L., Cummings, S., Lebas, E., Lietman, T. M., Keenan, J. D. & Doan, T. Effect of commonly used pediatric antibiotics on gut microbial diversity in preschool children in Burkina Faso: A randomized clinical trial. *Open Forum Infect Dis* **5**, ofy289-ofy289, (2018).

263 Jovel, J., Patterson, J., Wang, W., Hotte, N., O'Keefe, S., Mitchel, T., Perry, T., Kao, D., Mason, A. L., Madsen, K. L. & Wong, G. K. Characterization of the gut microbiome using 16S or shotgun metagenomics. *Front Microbiol* **7**, 459, (2016).

264 Hildebrand, F., Moitinho-Silva, L., Blasche, S., Jahn, M. T., Gossmann, T. I., Heuerta-Cepas, J., Hercog, R., Luetge, M., Bahram, M., Pryszlak, A., Alves, R. J., Waszak, S. M., Zhu, A., Ye, L., Costea, P. I., Aalvink, S., Belzer, C., Forslund, S. K., Sunagawa, S., Hentschel, U., Merten, C., Patil, K. R., Benes, V. & Bork, P. Antibiotics-induced monodominance of a novel gut bacterial order. *Gut*2019).

265 Walker, A. W., Duncan, S. H., Louis, P. & Flint, H. J. Phylogeny, culturing, and metagenomics of the human gut microbiota. *Trends Microbiol* **22**, 267-274, (2014).

266 Forbes, J. D., Knox, N. C., Ronholm, J., Pagotto, F. & Reimer, A. Metagenomics: The next culture-independent game changer. *Front Microbiol* **8**, 1069, (2017).

267 Truong, D. T., Tett, A., Pasolli, E., Huttenhower, C. & Segata, N. Microbial strain-level population structure and genetic diversity from metagenomes. *Genome Res* **27**, 626-638, (2017).

268 Jacobsen, L., Wilcks, A., Hammer, K., Huys, G., Gevers, D. & Andersen, S. R. Horizontal transfer of tet(M) and erm(B) resistance plasmids from food strains of *Lactobacillus plantarum* to *Enterococcus faecalis* JH2-2 in the gastrointestinal tract of gnotobiotic rats. *FEMS Microbiol Ecol* **59**, 158-166, (2007).

269 Schjorring, S., Struve, C. & Krogfelt, K. A. Transfer of antimicrobial resistance plasmids from *Klebsiella pneumoniae* to *Escherichia coli* in the mouse intestine. *J Antimicrob Chemother* **62**, 1086-1093, (2008).

270 Schjorring, S. & Krogfelt, K. A. Assessment of bacterial antibiotic resistance transfer in the gut. *Int J Microbiol* **2011**, 312956, (2011).

271 Lagkouvardos, I., Pukall, R., Abt, B., Foesel, B. U., Meier-Kolthoff, J. P., Kumar, N., Bresciani, A., Martínez, I., Just, S., Ziegler, C., Brugiroux, S., Garzetti, D., Wenning, M., Bui, T. P. N., Wang, J., Hugenholtz, F., Plugge, C. M., Peterson, D. A., Hornef, M. W., Baines, J. F., Smidt, H., Walter, J., Kristiansen, K., Nielsen, H. B., Haller, D., Overmann, J., Stecher, B. & Clavel, T. The Mouse Intestinal Bacterial Collection (miBC) provides host-specific insight into cultured diversity and functional potential of the gut microbiota. *Nat Microbiol* **1**, 16131, (2016).

272 Stephens, M. A. EDF statistics for goodness of fit and some comparisons. *J. Am. Stat. Assoc* **69**, 730-737, (1974).

273 D'Agostino, R. & Pearson, E. S. Tests for departure from normality. Empirical results for the distributions of b 2 andV b. *Biometrika* **60**, 613-622, (1973).

274 Shapiro, S. S. & Wilk, M. B. An analysis of variance test for normality (complete samples)†. *Biometrika* **52**, 591-611, (1965).

275 Massey Jr, F. J. The Kolmogorov-Smirnov test for goodness of fit. *J AM STAT ASSOC* **46**, 68-78, (1951).

276 Welch, B. L. The generalisation of student's problems when several different population variances are involved. *Biometrika* **34**, 28-35, (1947).

277 Goker, M., Gronow, S., Zeytun, A., Nolan, M., Lucas, S., Lapidus, A., Hammon, N., Deshpande, S., Cheng, J. F., Pitluck, S., Liolios, K., Pagani, I., Ivanova, N., Mavromatis, K., Ovchinikova, G., Pati, A., Tapia, R., Han, C., Goodwin, L., Chen, A., Palaniappan, K., Land, M., Hauser, L., Jeffries, C. D., Brambilla, E. M., Rohde, M., Detter, J. C., Woyke, T., Bristow, J., Markowitz, V., Hugenholtz, P., Eisen, J. A., Kyrpides, N. C. & Klenk, H. P. Complete genome sequence of *Odoribacter splanchnicus* type strain (1651/6). *Stand Genomic Sci* **4**, 200-209, (2011).

278 MacPherson, C. W., Mathieu, O., Tremblay, J., Champagne, J., Nantel, A., Girard, S. A. & Tompkins, T. A. Gut bacterial microbiota and its resistome rapidly recover to basal state levels after short-term amoxicillin-clavulanic acid treatment in healthy adults. *Sci Rep* **8**, 11192, (2018).

279 Drusano, G. L., Okusanya, O. O., Okusanya, A. O., van Scoy, B., Brown, D. L., Fregeau, C., Kulawy, R., Kinzig, M., Sorgel, F., Heine, H. S. & Louie, A. Impact of spore biology on the rate of kill and suppression of resistance in *Bacillus anthracis*. *Antimicrob Agents Chemother* **53**, 4718-4725, (2009).

280 Vandeputte, D., Kathagen, G., D'hoe, K., Vieira-Silva, S., Valles-Colomer, M., Sabino, J., Wang, J., Tito, R. Y., De Commer, L., Darzi, Y., Vermeire, S., Falony, G. & Raes, J. Quantitative microbiome profiling links gut community variation to microbial load. *Nature* **551**, 507-511, (2017).

281 Dubin, K. & Pamer, E. G. Enterococci and their interactions with the intestinal microbiome. *Microbiol Spectr* **5**, (2014).

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

283 Nord, C. E. & Edlund, C. Impact of antimicrobial agents on human intestinal microflora. *Journal of chemotherapy (Florence, Italy)* **2**, 218-237, (1990).

284 Sullivan, Å., Edlund, C. & Nord, C. E. Effect of antimicrobial agents on the ecological balance of human microflora. *Lancet Infect Dis* **1**, 101-114, (2001).

285 Baughn, A. D. & Malamy, M. H. The strict anaerobe *Bacteroides fragilis* grows in and benefits from nanomolar concentrations of oxygen. *Nature* **427**, 441-444, (2004).

286 Mishra, S. & Imlay, J. A. An anaerobic bacterium, *Bacteroides thetaiotaomicron*, uses a consortium of enzymes to scavenge hydrogen peroxide. *Mol Microbiol* **90**, 1356-1371, (2013).

287 Oliver, A., Kay, M. & Cooper, K. K. Comparative genomics of cocci-shaped *Sporosarcina* strains with diverse spatial isolation. *BMC Genomics* **19**, 310, (2018).

288 Wolfgang, W. J., Coorevits, A., Cole, J. A., De Vos, P., Dickinson, M. C., Hannett, G. E., Jose, R., Nazarian, E. J., Schumann, P., Van Landschoot, A., Wirth, S. E. & Musser, K. A. *Sporosarcina newyorkensis* sp. nov. from clinical specimens and raw cow's milk. *Int J Syst Evol Microbiol* **62**, 322-329, (2012).

289 Hildebrand, F., Pallen, M. J. & Bork, P. Towards standardisation of naming novel prokaryotic taxa in the age of high-throughput microbiology. *Gut*2019).

290 Parker, A. C. & Smith, C. J. Genetic and biochemical analysis of a novel Ambler class A beta-lactamase responsible for cefoxitin resistance in *Bacteroides* species. *Antimicrob Agents Chemother* **37**, 1028-1036, (1993).

291 García, N., Gutiérrez, G., Lorenzo, M., García, J. E., Píriz, S. & Quesada, A. Genetic determinants for cfxA expression in *Bacteroides* strains isolated from human infections. *J Antimicrob Chemother* **62**, 942-947, (2008).

292 Veloo, A. C. M., Baas, W. H., Haan, F. J., Coco, J. & Rossen, J. W. Prevalence of antimicrobial resistance genes in *Bacteroides* spp. and *Prevotella* spp. Dutch clinical isolates. *Clin Microbiol Infect* **25**, 1156 e1159-1156 e1113, (2019).

293 Connelly, S., Subramanian, P., Hasan, N. A., Colwell, R. R. & Kaleko, M. Distinct consequences of amoxicillin and ertapenem exposure in the porcine gut microbiome. *Anaerobe* **53**, 82-93, (2018).

294 Stiefel, U., Nerandzic, M. M., Pultz, M. J. & Donskey, C. J. Gastrointestinal colonization with a cephalosporinase-producing *Bacteroides* species preserves colonization resistance against vancomycin-resistant *Enterococcus* and *Clostridium difficile* in cephalosporin-treated mice. *Antimicrob Agents Chemother* **58**, 4535-4542, (2014).

295 Feldgarden, M., Brover, V., Haft, D. H., Prasad, A. B., Slotta, D. J., Tolstoy, I., Tyson, G. H., Zhao, S., Hsu, C.-H., McDermott, P. F., Tadesse, D. A., Morales, C., Simmons, M., Tillman, G., Wasilenko, J., Folster, J. P. & Klimke, W. Using the NCBI AMRFinder tool to determine antimicrobial resistance genotype-phenotype correlations within a collection of NARMS isolates. *bioRxiv*, 550707, (2019).

296 Ayling, M., Clark, M. D. & Leggett, R. M. New approaches for metagenome assembly with short reads. *Brief Bioinform* **bbz020**, (2019).

297 Lagier, J. C., Armougom, F., Million, M., Hugon, P., Pagnier, I., Robert, C., Bittar, F., Fournous, G., Gimenez, G., Maraninchi, M., Trape, J. F., Koonin, E. V., La Scola, B. & Raoult, D. Microbial culturomics: paradigm shift in the human gut microbiome study. *Clin Microbiol Infect* **18**, 1185-1193, (2012).

298 Doyle, R. M., O'Sullivan, D. M., Aller, S. D., Bruchmann, S., Clark, T., Pelegrin, A. C., Cormican, M., Benavente, E. D., Ellington, M. J., McGrath, E., Motro, Y., Thuy Nguyen, T. P., Phelan, J., Shaw, L. P., Stabler, R. A., van Belkum, A., van Dorp, L., Woodford, N., Moran-Gilad, J., Huggett, J. F. & Harris, K. A. Discordant bioinformatic predictions of antimicrobial resistance from whole-genome sequencing data of bacterial isolates: An inter-laboratory study. *bioRxiv*, 793885, (2019).

299 McDermott, P. F., Tyson, G. H., Kabera, C., Chen, Y., Li, C., Folster, J. P., Ayers, S. L., Lam, C., Tate, H. P. & Zhao, S. Whole-genome sequencing for detecting antimicrobial resistance in nontyphoidal *Salmonella*. *Antimicrob Agents Chemother* **60**, 5515-5520, (2016).

300 Ravcheev, D. A. & Thiele, I. Genomic analysis of the human gut microbiome suggests novel enzymes involved in quinone biosynthesis. *Front Microbiol* **7**, 128, (2016).

Appendix 1. CARD determinant groupings

Table A1.1. Genetic determinants of antibiotic resistance as described in CARD were grouped by the antibiotic to which they are described as conferring resistance. If more than one antibiotic class was included for a single determinant, these determinants were grouped as "non-specific". The exception is for resistances to Macrolide, Lincosamide, Pleuromutilin and Streptogramin (MLPS) antibiotics, as resistance determinants against these antibiotics can have cross-resistance to each other and are grouped together in a single, separate category.

Number of determinants	Custom group	CARD antibiotic class
1	Acridine dye	Acridine dye
9	Aminocoumarin	Aminocoumarin antibiotic
170	Aminoglycoside	Aminoglycoside antibiotic
2	Antibacterial free fatty acids	Antibacterial free fatty acids
51	Beta-lactam	Carbapenem
3	Beta-lactam	Carbapenem;cephalosporin
3	Beta-lactam	Carbapenem;cephalosporin;cephamycin
48	Beta-lactam	Carbapenem;cephalosporin;cephamycin;penam
85	Beta-lactam	Carbapenem;cephalosporin;cephamycin;penam;penem
194	Beta-lactam	Carbapenem;cephalosporin;penam
2	Beta-lactam	Carbapenem;penam
184	Beta-lactam	Cephalosporin
29	Beta-lactam	Cephalosporin;cephamycin
10	Beta-lactam	Cephalosporin;cephamycin;penam

Number of determinants	Custom group	CARD antibiotic class
347	Beta-lactam	Cephalosporin;penam
114	Beta-lactam	Cephamycin
31	Beta-lactam	Monobactam;carbapenem;cephalosporin
11	Beta-lactam	Monobactam;carbapenem;cephalosporin;cephamycin;penam
12	Beta-lactam	Monobactam;carbapenem;cephalosporin;cephamycin;penam;penem
18	Beta-lactam	Monobactam;carbapenem;cephalosporin;penam
7	Beta-lactam	Monobactam;carbapenem;cephalosporin;penam;penem
1	Beta-lactam	Monobactam;carbapenem;penam
11	Beta-lactam	Monobactam;cephalosporin
16	Beta-lactam	Monobactam;cephalosporin;cephamycin
8	Beta-lactam	Monobactam;cephalosporin;cephamycin;penam;penem
33	Beta-lactam	Monobactam;cephalosporin;penam
167	Beta-lactam	Monobactam;cephalosporin;penam;penem
29	Beta-lactam	Penam
24	Beta-lactam	Penam;penem
4	Beta-lactam	N/A
1	Bicyclomycin	Bicyclomycin
47	Chloramphenicol	Phenicol antibiotic
8	Elfamycin	Elfamycin antibiotic
31	Diaminopyrimidine	Diaminopyrimidine antibiotic
125	Fluoroquinolone	Fluoroquinolone antibiotic
28	Fosfomycin	Fosfomycin
6	Fusidic acid	Fusidic acid
77	Glycopeptide	Glycopeptide antibiotic
4	Isoniazid	Isoniazid
4	Mupirocin	Mupirocin

Number of determinants	Custom group	CARD antibiotic class
13	MLPS	Lincosamide antibiotic
1	MLPS	Lincosamide antibiotic;pleuromutilin antibiotic
5	MLPS	Lincosamide antibiotic;streptogramin antibiotic;pleuromutilin antibiotic
3	MLPS	Macrolide antibiotic
4	MLPS	Macrolide antibiotic;lincosamide antibiotic
33	MLPS	Macrolide antibiotic;lincosamide antibiotic;streptogramin antibiotic
5	MLPS	Macrolide antibiotic;streptogramin antibiotic
1	MLPS	Pleuromutilin antibiotic
11	MLPS	Streptogramin antibiotic
6	MLPS	Streptogramin antibiotic;pleuromutilin antibiotic
2	Nitrofuran	Nitrofuran antibiotic
1	Nitroimidazole	Nitroimidazole antibiotic
3	Non-specific	Aminoglycoside antibiotic;aminocoumarin antibiotic
5	Non-specific	Aminoglycoside antibiotic;cephalosporin;cephamycin;penam
2	Non-specific	Aminoglycoside antibiotic;tetracycline antibiotic;phenicol antibiotic
1	Non-specific	Cephalosporin;penam;peptide antibiotic
6	Non-specific	Fluoroquinolone antibiotic;acridine dye
1	Non-specific	Fluoroquinolone antibiotic;acridine dye;triclosan
5	Non-specific	Fluoroquinolone antibiotic;aminocoumarin antibiotic
5	Non-specific	Fluoroquinolone antibiotic;aminoglycoside antibiotic
2	Non-specific	Fluoroquinolone antibiotic;cephalosporin;cephamycin;penam
1	Non-specific	Fluoroquinolone antibiotic;cephalosporin;glycylcycline;cephamycin;penam;tetracycline antibiotic;rifamycin antibiotic;phenicol antibiotic;triclosan
1	Non-specific	Fluoroquinolone antibiotic;cephalosporin;glycylcycline;penam;tetracycline antibiotic;acridine dye;rifamycin antibiotic;phenicol antibiotic;triclosan
15	Non-specific	Fluoroquinolone antibiotic;cephalosporin;glycylcycline;penam;tetracycline antibiotic;rifamycin antibiotic;phenicol antibiotic;triclosan

Number of determinants	Custom group	CARD antibiotic class
1	Non-specific	Fluoroquinolone antibiotic;cephalosporin;penam;tetracycline antibiotic;peptide antibiotic;acridine dye
6	Non-specific	Fluoroquinolone antibiotic;diaminopyrimidine antibiotic;phenicol antibiotic
2	Non-specific	Fluoroquinolone antibiotic;glycylcycline;tetracycline antibiotic;diaminopyrimidine antibiotic;nitrofuran antibiotic
1	Non-specific	Fluoroquinolone antibiotic;lincosamide antibiotic;nucleoside antibiotic;acridine dye;phenicol antibiotic
1	Non-specific	Fluoroquinolone antibiotic;monobactam;carbapenem;cephalosporin;cephamycin;penam;tetracycline antibiotic;phenicol antibiotic;penem
4	Non-specific	Fluoroquinolone antibiotic;monobactam;carbapenem;cephalosporin;glycylcycline;cephamycin;penam;tetracycline antibiotic;rifamycin antibiotic;phenicol antibiotic;triclosan;penem
3	Non-specific	Fluoroquinolone antibiotic;monobactam;cephalosporin
1	Non-specific	Fluoroquinolone antibiotic;nucleoside antibiotic;acridine dye;phenicol antibiotic
16	Non-specific	Fluoroquinolone antibiotic;nybomycin
5	Non-specific	Fluoroquinolone antibiotic;tetracycline antibiotic
4	Non-specific	Fluoroquinolone antibiotic;tetracycline antibiotic;acridine dye
1	Non-specific	Fluoroquinolone antibiotic;tetracycline antibiotic;nitroimidazole antibiotic
10	Non-specific	Glycylcycline;tetracycline antibiotic
1	Non-specific	Isoniazid;ethionamide
3	Non-specific	Isoniazid;triclosan
1	Non-specific	Macrolide antibiotic;aminocoumarin antibiotic
3	Non-specific	Macrolide antibiotic;carbapenem;tetracycline antibiotic;acridine dye;diaminopyrimidine antibiotic;phenicol antibiotic
1	Non-specific	Macrolide antibiotic;fluoroquinolone antibiotic
3	Non-specific	Macrolide antibiotic;fluoroquinolone antibiotic;aminoglycoside antibiotic;carbapenem;cephalosporin;cephamycin;penam;tetracycline antibiotic;acridine dye;phenicol antibiotic
1	Non-specific	Macrolide antibiotic;fluoroquinolone antibiotic;aminoglycoside antibiotic;carbapenem;tetracycline antibiotic
4	Non-specific	Macrolide antibiotic;fluoroquinolone antibiotic;aminoglycoside antibiotic;cephalosporin;penam;tetracycline antibiotic;aminocoumarin antibiotic;diaminopyrimidine antibiotic;phenicol antibiotic
4	Non-specific	Macrolide antibiotic;fluoroquinolone antibiotic;aminoglycoside antibiotic;cephalosporin;tetracycline antibiotic
1	Non-specific	Macrolide antibiotic;fluoroquinolone antibiotic;cephalosporin;cephamycin;penam;tetracycline antibiotic

Number of determinants	Custom group	CARD antibiotic class
4	Non-specific	Macrolide antibiotic;fluoroquinolone antibiotic;cephalosporin;fusidic acid
1	Non-specific	Macrolide antibiotic;fluoroquinolone antibiotic;cephalosporin;glycylcycline;cephamycin;penam;tetracycline antibiotic;aminocoumarin antibiotic;rifamycin antibiotic;phenicol antibiotic;triclosan
1	Non-specific	Macrolide antibiotic;fluoroquinolone antibiotic;cephalosporin;penam;tetracycline antibiotic;aminocoumarin antibiotic;diaminopyrimidine antibiotic;phenicol antibiotic
4	Non-specific	Macrolide antibiotic;fluoroquinolone antibiotic;lincosamide antibiotic;carbapenem;cephalosporin;tetracycline antibiotic;rifamycin antibiotic;diaminopyrimidine antibiotic;phenicol antibiotic;penem
1	Non-specific	Macrolide antibiotic;fluoroquinolone antibiotic;monobactam;aminoglycoside antibiotic;carbapenem;cephalosporin;cephamycin;penam;tetracycline antibiotic;peptide antibiotic;acridine dye;aminocoumarin antibiotic;diaminopyrimidine antibiotic;sulfonamide antibiotic;phenicol antibiotic;penem
2	Non-specific	Macrolide antibiotic;fluoroquinolone antibiotic;monobactam;aminoglycoside antibiotic;carbapenem;cephalosporin;cephamycin;penam;tetracycline antibiotic;peptide antibiotic;aminocoumarin antibiotic;diaminopyrimidine antibiotic;sulfonamide antibiotic;phenicol antibiotic;penem
6	Non-specific	Macrolide antibiotic;fluoroquinolone antibiotic;monobactam;carbapenem;cephalosporin;cephamycin;penam;tetracycline antibiotic;peptide antibiotic;aminocoumarin antibiotic;diaminopyrimidine antibiotic;sulfonamide antibiotic;phenicol antibiotic;penem
5	Non-specific	Macrolide antibiotic;fluoroquinolone antibiotic;penam
2	Non-specific	Macrolide antibiotic;fluoroquinolone antibiotic;penam;tetracycline antibiotic
2	Non-specific	Macrolide antibiotic;fluoroquinolone antibiotic;rifamycin antibiotic
2	Non-specific	Macrolide antibiotic;fluoroquinolone antibiotic;tetracycline antibiotic;acridine dye;phenicol antibiotic
3	Non-specific	Macrolide antibiotic;fluoroquinolone antibiotic;tetracycline antibiotic;phenicol antibiotic
6	Non-specific	Macrolide antibiotic;lincosamide antibiotic;streptogramin antibiotic;oxazolidinone antibiotic;phenicol antibiotic
1	Non-specific	Macrolide antibiotic;lincosamide antibiotic;streptogramin antibiotic;oxazolidinone antibiotic;phenicol antibiotic;pleuromutilin antibiotic
1	Non-specific	Macrolide antibiotic;lincosamide antibiotic;streptogramin antibiotic;tetracycline antibiotic
4	Non-specific	Macrolide antibiotic;monobactam;tetracycline antibiotic;aminocoumarin antibiotic
4	Non-specific	Macrolide antibiotic;penam
1	Non-specific	Macrolide antibiotic;penam;antibacterial free fatty acids
4	Non-specific	Macrolide antibiotic;peptide antibiotic
3	Non-specific	Macrolide antibiotic;tetracycline antibiotic;triclosan

Number of determinants	Custom group	CARD antibiotic class
4	Non-specific	Monobactam;carbapenem;cephalosporin;cephamycin;penam;phenicol antibiotic;penem
1	Non-specific	Monobactam;carbapenem;cephalosporin;cephamycin;penam;tetracycline antibiotic;penem
3	Non-specific	Nucleoside antibiotic;acridine dye
1	Non-specific	Peptide antibiotic;polyamine antibiotic
6	Non-specific	Peptide antibiotic;rifamycin antibiotic
1	Non-specific	Rifamycin antibiotic;isoniazid
1	Non-specific	Rifamycin antibiotic;polyamine antibiotic
1	Non-specific	Tetracycline antibiotic;benzalkonium chloride;rhodamine
7	Non-specific	N/A
4	Nucleoside	Nucleoside antibiotic
1	Oxazolidinone	Oxazolidinone antibiotic
3	Para-aminosalicylic acid	Para-aminosalicylic acid
56	Peptide	Peptide antibiotic
8	Polyamine	Polyamine antibiotic
1	Pyrazinamide	Pyrazinamide
13	Rifamycin	Rifamycin antibiotic
1	Sulfonamide	Sulfonamide antibiotic
6	Sulfonamide	Sulfonamide antibiotic;sulfone antibiotic
52	Tetracycline	Tetracycline antibiotic
7	Triclosan	Triclosan

Appendix 2. HBC antibiotic resistance determinant groupings

Table A2.1. The genetic determinants of antibiotic resistance predicted in the HBC genomes, as described in CARD, were grouped by the antibiotic to which they are described as conferring resistance. If more than one antibiotic class was included for a single determinant, these determinants were grouped as "non-specific". The exception is for resistances to Macrolide, Lincosamide, Pleuromutilin and Streptogramin (MLPS) antibiotics, as resistance determinants against these antibiotics can have cross-resistance to each other.

Custom Group	CARD antibiotic classes	Antibiotic resistance gene/determinant	Mechanism
Acridine dye	Acridine dye	*emeA*	Antibiotic efflux
Aminocoumarin	Aminocoumarin antibiotic	*mdtA*	Antibiotic efflux
Aminocoumarin	Aminocoumarin antibiotic	*mdtB*	Antibiotic efflux
Aminocoumarin	Aminocoumarin antibiotic	*mdtC*	Antibiotic efflux
Aminoglycoside	Aminoglycoside	*aac(6')-Ie-aph(2'')-Ia*	Antibiotic inactivation
Aminoglycoside	Aminoglycoside	*aac(6')-Ii*	Antibiotic inactivation
Aminoglycoside	Aminoglycoside	*aac(6')-Iih*	Antibiotic inactivation
Aminoglycoside	Aminoglycoside	*aadA*	Antibiotic inactivation
Aminoglycoside	Aminoglycoside	*aadK*	Antibiotic inactivation
Aminoglycoside	Aminoglycoside	*ant(6)-Ia*	Antibiotic inactivation
Aminoglycoside	Aminoglycoside	*Ant(6)-Ib*	Antibiotic inactivation
Aminoglycoside	Aminoglycoside	*aph(2'')*	Antibiotic inactivation
Aminoglycoside	Aminoglycoside	*aph(2'')-If*	Antibiotic inactivation
Aminoglycoside	Aminoglycoside	*aph(2'')-IIa*	Antibiotic inactivation

Custom Group	CARD antibiotic classes	Antibiotic resistance gene/determinant	Mechanism
Aminoglycoside	Aminoglycoside	*aph(3'')-Ib*	Antibiotic inactivation
Aminoglycoside	Aminoglycoside	*aph(3')-Ia*	Antibiotic inactivation
Aminoglycoside	Aminoglycoside	*aph(3')-IIIa*	Antibiotic inactivation
Aminoglycoside	Aminoglycoside	*aph(6)-Id*	Antibiotic inactivation
Aminoglycoside	Aminoglycoside	*kdpE*	Antibiotic efflux
Aminoglycoside	Aminoglycoside antibiotic	*acrD*	Antibiotic efflux
Beta-lactam	Carbapenem, cephalosporin, penam	*shv*	Antibiotic inactivation
Beta-lactam	Cephalosporin	*cblA_1*	Antibiotic inactivation
Beta-lactam	Cephalosporin	*cepA*	Antibiotic inactivation
Beta-lactam	Cephalosporin, monobactam, penam	*oxy_2*	Antibiotic inactivation
Beta-lactam	Cephalosporin, penam	*bcl*	Antibiotic inactivation
Beta-lactam	Cephalosporin, penam	*bclI*	Antibiotic inactivation
Beta-lactam	Cephalosporin, penam	*oxa_114a*	Antibiotic inactivation
Beta-lactam	Cephamycin	*cfxA*	Antibiotic inactivation
Beta-lactam	Cephamycin	*cmy_101*	Antibiotic inactivation
Beta-lactam	Monobactam, cephalosporin, penam, penem	*tem*	Antibiotic inactivation
Beta-lactam	Monobactam, cephamycin, cephalosporin, penam, carbapenem, penem	*K. pneumoniae ompK35*	Resistance by absence or reduced permeability to antibiotic
Beta-lactam	Penam	*blaZ*	Antibiotic inactivation
Beta-lactam	Penam, carbapenem, cephamycin, cephalosporin	*act*	Antibiotic inactivation
Beta-lactam	Penam carbapenem, cephamycin, cephalosporin	*act_1*	Antibiotic inactivation
Beta-lactam	Penam, carbapenem, cephamycin, cephalosporin, penem, monobactam	*K. pneumoniae ompK36*	Resistance by absence or reduced permeability to antibiotic
Beta-lactam	Penam carbapenem, cephamycin, cephalosporin, penem, monobactam	*ompC*	Resistance by absence or reduced permeability to antibiotic
Beta-lactam	Penam, cephalosporin	*ampC_2*	Antibiotic inactivation

Custom Group	CARD antibiotic classes	Antibiotic resistance gene/determinant	Mechanism
Beta-lactam	Penam, cephalosporin	*oxa_347*	Antibiotic inactivation
Beta-lactam	Penam ,cephalosporin	*oxa_4*	Antibiotic inactivation
Beta-lactam	Penam ,cephalosporin, monobactam	*acc*	Antibiotic inactivation
Beta-lactam	Penam, cephalosporin, monobactam	*oxy*	Antibiotic inactivation
Beta-lactam	Penam, monobactam, penam, carbapenem, cephamycin, cephalosporin	*K. pneumoniae ompK37*	Resistance by absence or reduced permeability to antibiotic
Beta-lactam	Penem, penam	*len_1*	Antibiotic inactivation
Chloramphenicol	Phenicol antibiotic	*cat*	Antibiotic inactivation
Chloramphenicol	Phenicol antibiotic	*E. faecalis chloramphenicol+*	Antibiotic inactivation
Diaminopyrimidine	Diaminopyrimidine antibiotic	*dfrA14*	Antibiotic target replacement
Diaminopyrimidine	Diaminopyrimidine antibiotic	*dfrC*	Antibiotic target replacement
Diaminopyrimidine	Diaminopyrimidine antibiotic	*dfrE*	Antibiotic target replacement
Diaminopyrimidine	Diaminopyrimidine antibiotic	*dfrF*	Antibiotic target replacement
Diaminopyrimidine	Diaminopyrimidine antibiotic	*dfrG*	Antibiotic target replacement
Elfamycin	Elfamycin	*E. coli ef-tu*	Antibiotic target alteration
Fluoroquinolone	Fluoroquinolone antibiotic	*emrA*	Antibiotic efflux
Fluoroquinolone	Fluoroquinolone antibiotic	*emrB*	Antibiotic efflux
Fluoroquinolone	Fluoroquinolone antibiotic	*emrR*	Antibiotic efflux
Fluoroquinolone	Fluoroquinolone antibiotic	*mdtH*	Antibiotic efflux
Fluoroquinolone	Fluoroquinolone antibiotic	*patA*	Antibiotic efflux
Fosfomycin	Fosfomycin	*E. coli cyaA*	Antibiotic target alteration
Fosfomycin	Fosfomycin	*E. coli glpT*	Antibiotic target alteration
Fosfomycin	Fosfomycin	*E. coli uhpA*	Antibiotic target alteration
Fosfomycin	Fosfomycin	*E. coli uhpT*	Antibiotic target alteration
Fosfomycin	Fosfomycin	*fosA2*	Antibiotic inactivation

Custom Group	CARD antibiotic classes	Antibiotic resistance gene/determinant	Mechanism
Fosfomycin	Fosfomycin	*fosA5*	Antibiotic inactivation
Fosfomycin	Fosfomycin	*fosA7*	Antibiotic inactivation
Fosfomycin	Fosfomycin	*fosB*	Antibiotic inactivation
Fosfomycin	Fosfomycin	*mdtG*	Antibiotic efflux
Glycopeptide	Glycopeptide antibiotic	*vanA*	Antibiotic target alteration
Glycopeptide	Glycopeptide antibiotic	*vanC*	Antibiotic target alteration
Glycopeptide	Glycopeptide antibiotic	*vanD*	Antibiotic target alteration
Glycopeptide	Glycopeptide antibiotic	*vanHA*	Antibiotic target alteration
Glycopeptide	Glycopeptide antibiotic	*vanHD*	Antibiotic target alteration
Glycopeptide	Glycopeptide antibiotic	*vanRA*	Antibiotic target alteration
Glycopeptide	Glycopeptide antibiotic	*vanRC*	Antibiotic target alteration
Glycopeptide	Glycopeptide antibiotic	*vanRD*	Antibiotic target alteration
Glycopeptide	Glycopeptide antibiotic	*vanSA*	Antibiotic target alteration
Glycopeptide	Glycopeptide antibiotic	*vanSC*	Antibiotic target alteration
Glycopeptide	Glycopeptide antibiotic	*vanTC*	Antibiotic target alteration
Glycopeptide	Glycopeptide antibiotic	*vanXA*	Antibiotic target alteration
Glycopeptide	Glycopeptide antibiotic	*vanXD*	Antibiotic target alteration
Glycopeptide	Glycopeptide antibiotic	*vanXYC*	Antibiotic target alteration
MLPS	Lincosamide antibiotic	*lmrB*	Antibiotic efflux
MLPS	Lincosamide antibiotic	*lnuC*	Antibiotic inactivation
MLPS	Lincosamide antibiotic	*lnuG*	Antibiotic inactivation
MLPS	Lincosamide antibiotic macrolide antibiotic, streptogramin antibiotic	*ermB*	Antibiotic target alteration
MLPS	Lincosamide antibiotic, streptogramin antibiotic, macrolide antibiotic	*ermG*	Antibiotic target alteration
MLPS	Macrolide antibiotic	*mefA*	antibiotic efflux
MLPS	Macrolide antibiotic, lincosamide antibiotic, streptogramin antibiotic	*ermF*	Antibiotic target alteration

Custom Group	CARD antibiotic classes	Antibiotic resistance gene/determinant	Mechanism
MLPS	Pleuromutilin antibiotic, lincosamide antibiotic	*lsaE*	Antibiotic efflux
MLPS	Pleuromutilin antibiotic, lincosamide antibiotic, streptogramin antibiotic	*eatAv*	Antibiotic efflux
MLPS	Streptogramin antibiotic, lincosamide antibiotic, macrolide antibiotic	*ermQ*	Antibiotic target alteration
MLPS	Streptogramin antibiotic, lincosamide antibiotic, pleuromutilin antibiotic	*lsaA*	Antibiotic efflux
MLPS	Streptogramin antibiotic, macrolide antibiotic	*mel*	Antibiotic efflux
MLPS	Streptogramin antibiotic, macrolide antibiotic	*msrC*	Antibiotic efflux
MLPS	Streptogramin antibiotic, macrolide antibiotic, lincosamide antibiotic	*ermD*	Antibiotic efflux
MLPS	Streptogramin antibiotic, pleuromutilin antibiotic	*vgaC*	Antibiotic efflux
Mupirocin	Mupirocin	*ileS*	antibiotic target alteration
Nitrofuran	Nitrofuran antibiotic	*nfsA*	antibiotic target alteration
Nitroimidazole	Nitroimidazole antibiotic	*msbA*	Antibiotic efflux
Non-specific	Acridine dye, fluoroquinolone antibiotic	*blt*	Antibiotic efflux
Non-specific	Acridine dye, fluoroquinolone antibiotic	*cdeA*	Antibiotic efflux
Non-specific	Aminocoumarin antibiotic, aminoglycoside antibiotic	*baeR*	Antibiotic efflux
Non-specific	Aminocoumarin antibiotic, aminoglycoside antibiotic	*baeS*	Antibiotic efflux
Non-specific	Aminocoumarin antibiotic, aminoglycoside antibiotic	*cpxA*	Antibiotic efflux
Non-specific	Aminoglycoside antibiotic, fluoroquinolone antibiotic	*mipA*	Resistance by absence or reduced permeability to antibiotic
Non-specific	Aminoglycoside antibiotic, tetracycline antibiotic, phenicol antibiotic	*ykkC*	Antibiotic efflux
Non-specific	Cephalosporin, macrolide antibiotic, aminoglycoside antibiotic, fluoroquinolone antibiotic	*axyX*	Antibiotic efflux
Non-specific	Cephalosporin, macrolide antibiotic, aminoglycoside antibiotic, fluoroquinolone antibiotic	*axyZ*	Antibiotic efflux
Non-specific	Fluoroquinolone antibiotic, acridine dye	*norA*	Antibiotic efflux
Non-specific	Fluoroquinolone antibiotic, macrolide antibiotic, penam, cephamycin, cephalosporin, tetracycline antibiotic	*h_ns*	Antibiotic efflux

Custom Group	CARD antibiotic classes	Antibiotic resistance gene/determinant	Mechanism
Non-specific	Fluoroquinolone antibiotic, penam, cephamycin, cephalosporin	*acrF*	Antibiotic efflux
Non-specific	Fluoroquinolone antibiotic, penam, macrolide antibiotic	*crp*	Antibiotic efflux
Non-specific	Fluoroquinolone antibiotic, penam, macrolide antibiotic	*mdtE*	Antibiotic efflux
Non-specific	Fluoroquinolone antibiotic, penam, tetracycline antibiotic, macrolide antibiotic	*evgA*	Antibiotic efflux
Non-specific	Glycopeptide antibiotic, streptogramin antibiotic, lincosamide antibiotic, macrolide antibiotic, phenicol antibiotic, pleuromutilin antibiotic	*Propionibacterium 23S rRNA*	Antibiotic target alteration
Non-specific	Glycopeptide antibiotic, streptogramin antibiotic, lincosamide antibiotic, macrolide antibiotic, phenicol antibiotic, pleuromutilin antibiotic	*S. pneumoniae 23S rRNA*	Antibiotic target alteration
Non-specific	Glycopeptide antibiotic, streptogramin antibiotic, macrolide antibiotic, phenicol antibiotic, lincosamide antibiotic, pleuromutilin antibiotic	*E. coli 23S rRNA*	Antibiotic target alteration
Non-specific	Glycopeptide antibiotic, streptogramin antibiotic, macrolide antibiotic, phenicol antibiotic, lincosamide antibiotic, pleuromutilin antibiotic, oxazolidinone antibiotic	*S. aureus 23S rRNA*	Antibiotic target alteration
Non-specific	Glycopeptide antibiotic, tetracycline antibiotic, glycylcycline, nucleoside antibiotic, aminoglycoside antibiotic, peptide antibiotic	*rrsB*	Antibiotic target alteration
Non-specific	Glycopeptide antibiotic, tetracycline antibiotic, peptide antibiotic, aminoglycoside antibiotic, nucleoside antibiotic, glycylcycline	*M. abscessus 16S rRNA*	Antibiotic target alteration
Non-specific	Macrolide antibiotic, cephalosporin, aminoglycoside antibiotic, fluoroquinolone antibiotic	*axyY*	Antibiotic efflux
Non-specific	Macrolide antibiotic, cephalosporin, aminoglycoside antibiotic, fluoroquinolone antibiotic	*oprZ*	Antibiotic efflux
Non-specific	Macrolide antibiotic, fluoroquinolone antibiotic	*efmA*	Antibiotic efflux
Non-specific	Macrolide antibiotic, penam, fluoroquinolone antibiotic	*gadW*	Antibiotic efflux
Non-specific	Macrolide antibiotic, penam, fluoroquinolone antibiotic	*gadX*	Antibiotic efflux
Non-specific	Nucleoside antibiotic, acridine dye	*mdtN*	Antibiotic efflux
Non-specific	Nucleoside antibiotic, acridine dye	*mdtO*	Antibiotic efflux
Non-specific	Nucleoside antibiotic, acridine dye	*mdtP*	Antibiotic efflux
Non-specific	Nucleoside antibiotic, phenicol antibiotic, acridine dye, fluoroquinolone antibiotic	*bmr*	Antibiotic efflux

Custom Group	CARD antibiotic classes	Antibiotic resistance gene/determinant	Mechanism
Non-specific	Nucleoside antibiotic, phenicol antibiotic, lincosamide antibiotic, fluoroquinolone antibiotic, acridine dye	*mdtM*	Antibiotic efflux
Non-specific	Penam, cephamycin, cephalosporin, fluoroquinolone antibiotic	*acrE*	Antibiotic efflux
Non-specific	Penam, macrolide antibiotic, fluoroquinolone antibiotic	*mdtF*	Antibiotic efflux
Non-specific	Peptide antibiotic, macrolide antibiotic	*mgrB*	Resistance by absence, antibiotic efflux and antibiotic target alteration
Non-specific	Tetracycline antibiotic, benzalkonium chloride, rhodamine	*mdfA*	Antibiotic efflux
Non-specific	Tetracycline antibiotic, cephalosporin, rifamycin antibiotic, phenicol antibiotic, glycylcycline, penam, fluoroquinolone antibiotic, triclosan	*acrA1*	Antibiotic efflux
Non-specific	Tetracycline antibiotic, diaminopyrimidine antibiotic, glycylcycline, nitrofuran antibiotic, fluoroquinolone antibiotic	*oqxA*	Antibiotic efflux
Non-specific	Tetracycline antibiotic, fluoroquinolone antibiotic	*E. coli lamB*	Resistance by absence or reduced permeability to antibiotic
Non-specific	Tetracycline antibiotic, glycylcycline, rifamycin antibiotic, phenicol antibiotic, fluoroquinolone antibiotic, penam, cephamycin, cephalosporin, triclosan	*acrS*	Antibiotic efflux
Non-specific	Tetracycline antibiotic, macrolide antibiotic, fluoroquinolone antibiotic, penam	*evgS*	Antibiotic efflux
Non-specific	Tetracycline antibiotic, nitrofuran antibiotic, fluoroquinolone antibiotic, glycylcycline, diaminopyrimidine antibiotic	*oqxB*	Antibiotic efflux
Non-specific	Tetracycline antibiotic, penam, cephalosporin, glycylcycline, rifamycin antibiotic, phenicol antibiotic, triclosan, fluoroquinolone antibiotic	*acrB*	Antibiotic efflux
Non-specific	Tetracycline antibiotic, penam, cephalosporin, rifamycin antibiotic, phenicol antibiotic, glycylcycline, fluoroquinolone antibiotic, triclosan	*acrA*	Antibiotic efflux
Non-specific	Tetracycline antibiotic, penam, cephalosporin, rifamycin antibiotic, phenicol antibiotic, glycylcycline, fluoroquinolone antibiotic, triclosan	*marR*	Antibiotic efflux and antibiotic target alteration
Non-specific	Tetracycline antibiotic, penem, penam, carbapenem, cephamycin, cephalosporin, rifamycin antibiotic, phenicol antibiotic, monobactam, glycylcycline, fluoroquinolone antibiotic, triclosan	*marA*	Antibiotic efflux and reduced permeability to beta-lactams

Custom Group	CARD antibiotic classes	Antibiotic resistance gene/determinant	Mechanism
Non-specific	Tetracycline antibiotic, phenicol antibiotic, aminoglycoside antibiotic	*ykkD*	Antibiotic efflux
Non-specific	Tetracycline antibiotic, phenicol antibiotic, glycylcycline, penam, cephalosporin, rifamycin antibiotic, fluoroquinolone antibiotic, triclosan	*acrA2*	Antibiotic efflux
Non-specific	Tetracycline antibiotic, rifamycin antibiotic, phenicol antibiotic, aminocoumarin antibiotic, fluoroquinolone antibiotic, penam, cephamycin, cephalosporin, glycylcycline, macrolide antibiotic, triclosan	*tolC*	Antibiotic efflux
Non-specific	Tetracycline antibiotic, rifamycin antibiotic, phenicol antibiotic, glycylcycline, penam, cephalosporin, triclosan, fluoroquinolone antibiotic	*ramR*	Antibiotic efflux
Non-specific		*emrD*	Antibiotic efflux
Non-specific		*emrE*	Antibiotic efflux
Nucleoside	Nucleoside antibiotic	*sat2*	Antibiotic inactivation
Nucleoside	Nucleoside antibiotic	*sat4*	Antibiotic inactivation
Nucleoside	Nucleoside antibiotic	*tmrB*	Reduced permeability to antibiotic
Peptide	Peptide antibiotic	*bacA*	Antibiotic target alteration
Peptide	Peptide antibiotic	*bcrA*	Antibiotic efflux
Peptide	Peptide antibiotic	*bcrB*	Antibiotic efflux
Peptide	Peptide antibiotic	*bcrC*	Antibiotic efflux
Peptide	Peptide antibiotic	*eptA*	Antibiotic target alteration
Peptide	Peptide antibiotic	*mprF*	Antibiotic target alteration
Peptide	Peptide antibiotic	*mprF3*	Antibiotic target alteration
Peptide	Peptide antibiotic	*mprF4*	Antibiotic target alteration
Peptide	Peptide antibiotic	*pmrF*	Antibiotic target alteration
Peptide	Peptide antibiotic	*ugd*	Antibiotic target alteration
Peptide	Peptide antibiotic	*yojI*	Antibiotic efflux

Custom Group	CARD antibiotic classes	Antibiotic resistance gene/determinant	Mechanism
Sulphonamide	Sulfonamide antibiotic, sulfone antibiotic	*sul2*	Antibiotic target replacement
Tetracycline	Tetracycline antibiotic	*emrK*	Antibiotic efflux
Tetracycline	Tetracycline antibiotic	*emrY*	Antibiotic efflux
Tetracycline	Tetracycline antibiotic	*P. acnes 16S rRNA*	Antibiotic target alteration
Tetracycline	Tetracycline antibiotic	*tet_40*	Antibiotic efflux
Tetracycline	Tetracycline antibiotic	*tet_J*	Antibiotic efflux
Tetracycline	Tetracycline antibiotic	*tet_K*	Antibiotic efflux
Tetracycline	Tetracycline antibiotic	*tet44*	Antibiotic target protection
Tetracycline	Tetracycline antibiotic	*tetA_46*	Antibiotic efflux
Tetracycline	Tetracycline antibiotic	*tetA_P*	Antibiotic efflux
Tetracycline	Tetracycline antibiotic	*tetB_P*	Antibiotic target protection
Tetracycline	Tetracycline antibiotic	*tetM*	Antibiotic target protection
Tetracycline	Tetracycline antibiotic	*tetO*	Antibiotic target protection
Tetracycline	Tetracycline antibiotic	*tetQ*	Antibiotic target protection
Tetracycline	Tetracycline antibiotic	*tetS*	Antibiotic target protection
Tetracycline	Tetracycline antibiotic, glycylcycline	*tetA*	Antibiotic efflux

Table A2.2. A complete list of HBC isolates and the antibiotic resistance genes and mutations identified in their genomes. The genetic determinants of antibiotic resistance predicted in the HBC genomes, as described in CARD, were grouped by the antibiotic to which they are described as conferring resistance. If more than one antibiotic class was included for a single determinant, these determinants were grouped as "non-specific". The exception is for resistances to Macrolide, Lincosamide, Pleuromutilin and Streptogramin (MLPS) antibiotics, as resistance determinants against these antibiotics can have cross-resistance to each other. The table can be viewed here:

https://docs.google.com/spreadsheets/d/1zwmhUicOW3JVW_9y6P6LssbavW47EFiRq4_wwnS9CMg/edit?usp=sharing

This table can also be found on the CD included with the hard copy of this thesis.

Appendix 3: Gut microbiota community composition in mice with human-derived gut microbiota

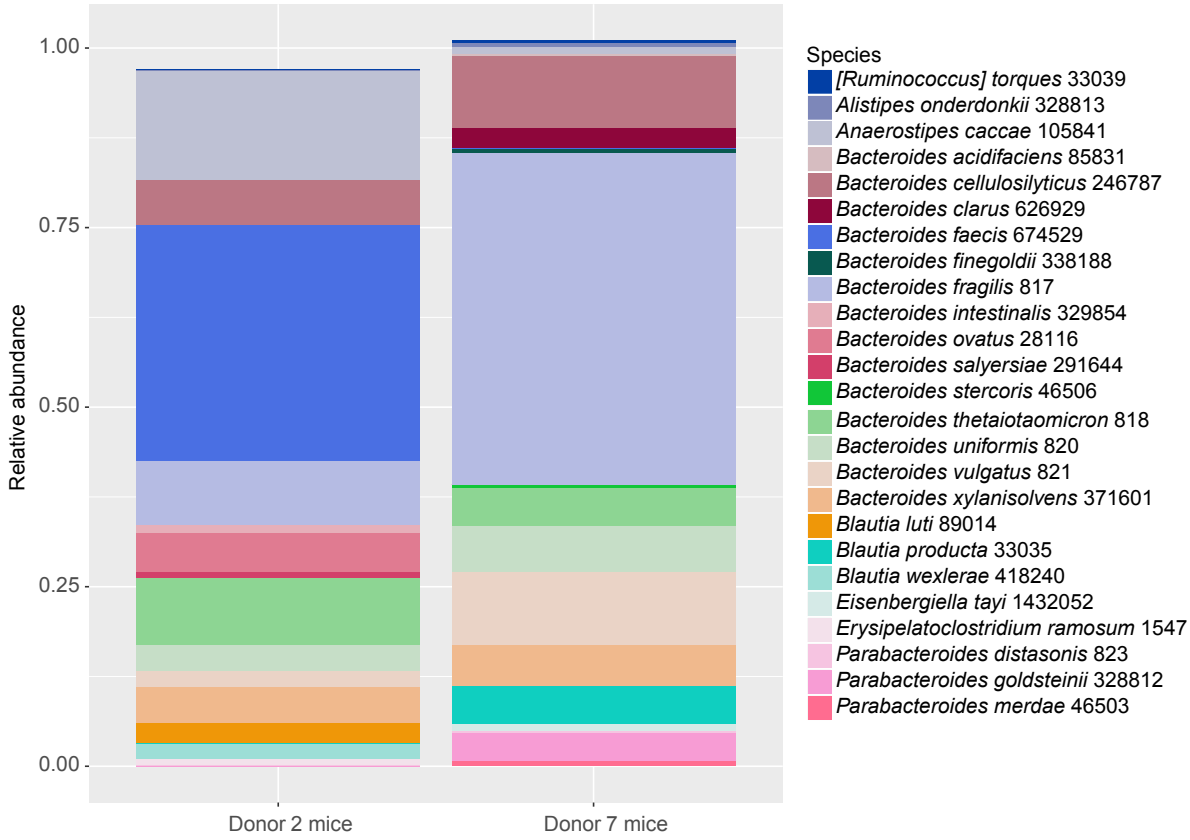


Figure A3.1. The relative abundance of species cultured from stool of Donor 2- and Donor 7-derived mice under anaerobic vegetative conditions. One faecal pellet was collected from each of five mice housed together in a single cage, of both Donor 2 and Donor 7 mouse lines. The stools were weighed and homogenised 100mg/ml, then pooled per mouse cage in equal volumes. The pooled samples were diluted and plated on YCFA agar and cultured at 37°C for 48 hours. The total growth was collected from each plate and total DNA extracted for metagenomic sequencing. Sequences reads were taxonomically classified using Kraken and a database of all publicly available gut bacterial genomes (as of December 2017) plus the HBC reference genomes. Classified read counts per species were normalised against the total number of classified reads per sample. The relative abundance of the top 25 most abundant species is shown.

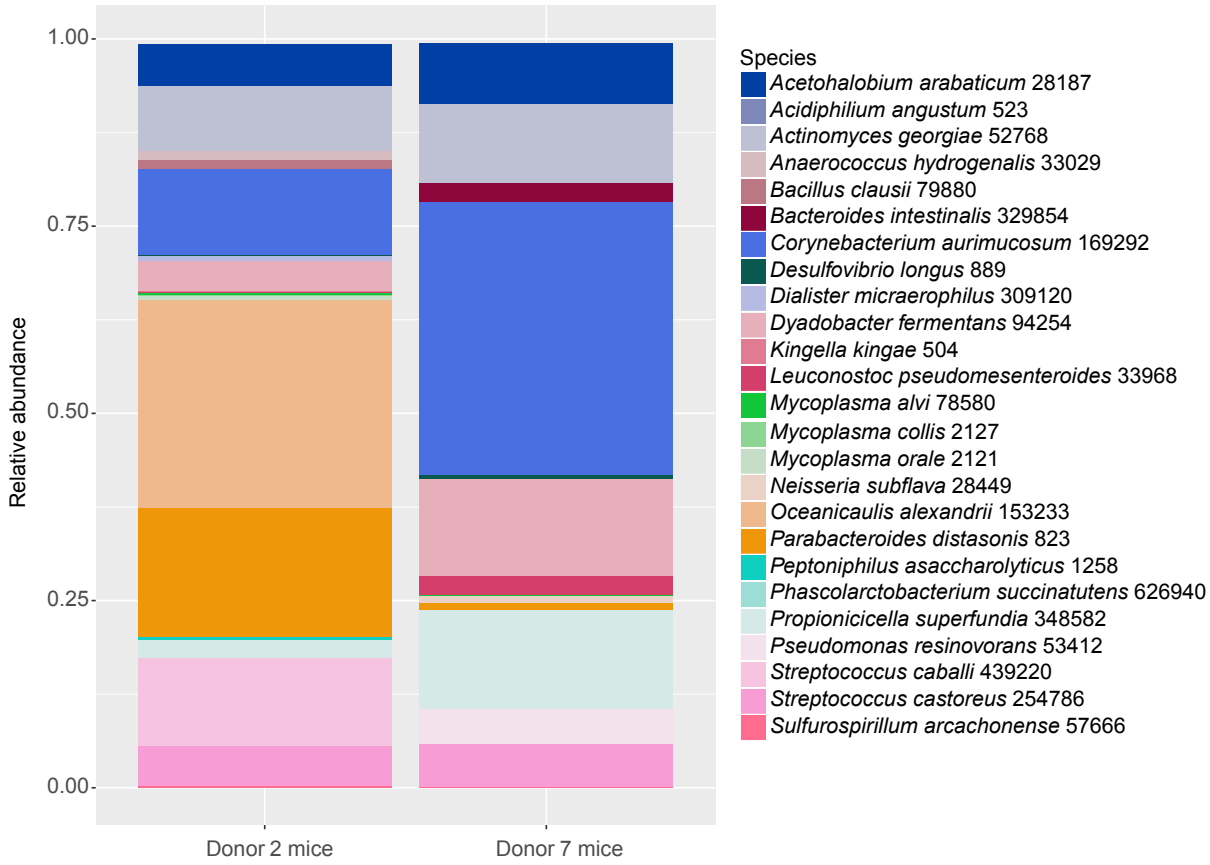


Figure A3.2. The relative abundance of amoxicillin-resistant species cultured from stool of Donor 2- and Donor 7-derived mice under anaerobic vegetative conditions. One stool was collected from each of five mice housed together in a single cage, of both Donor 2 and Donor 7 mouse lines. The stools were weighed and homogenised 100mg/ml, then pooled per cage in equal volumes. The pooled samples were diluted and plated on YCFA agar containing 8 mg/L amoxicillin (a level considered 'clinically resistant' by CLSI and EUCAST) and cultured at 37C for 48 hours. The complete growth was collected from each plate and total DNA extracted for metagenomic sequencing. Sequences reads were taxonomically classified using Kraken and a database of all publicly available gut bacterial genomes plus the HBC reference genomes. Classified read counts per species were normalised against the total number of classified reads per sample. The relative abundance of the top 25 most abundant species is shown.

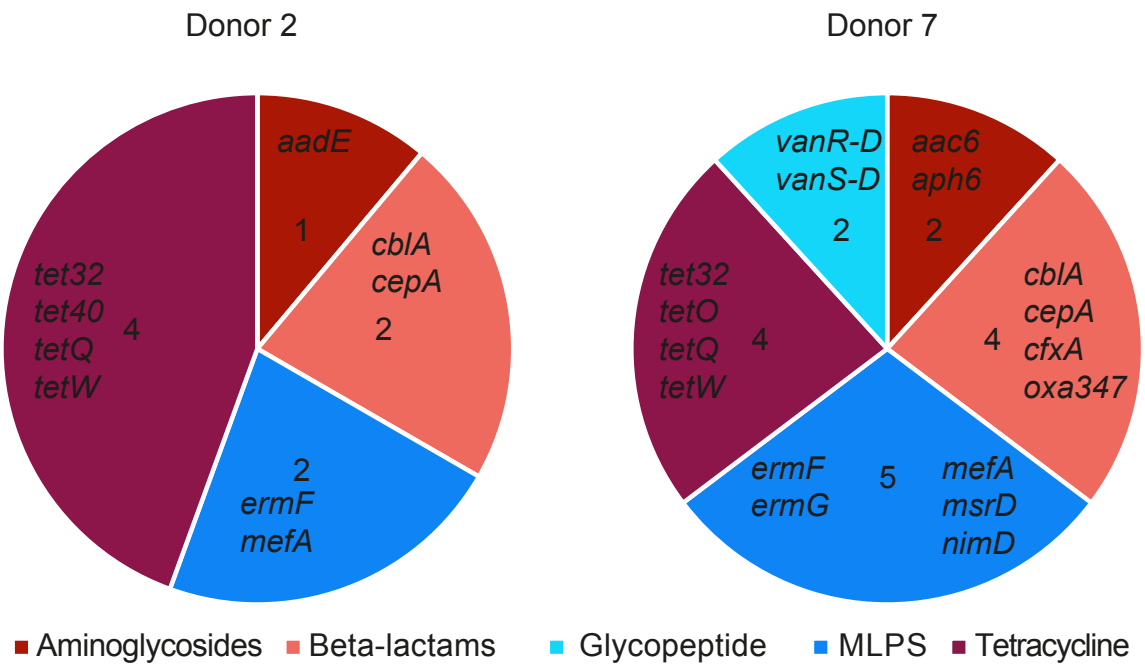


Figure A3.3. Antibiotic resistance genes (ARGs) identified in Donor 2- and Donor 7-derived mice. One stool was collected from each of five mice housed together in a single cage, of both Donor 2 and Donor 7 mouse lines. The stools were weighed and homogenised 100mg/ml, then pooled per cage in equal volumes. The pooled samples were diluted and plated on YCFA agar and cultured at 37C for 48 hours. The complete growth was collected from each plate and total DNA extracted for metagenomic sequencing. The presence of ARGs in the CARD, ResFinder and MegaRes databases were predicted from the metagenomic sequence reads using ARIBA. Results using each database were combined to count the different ARGs predicted to be present in the gut community of each mouse line. 19 different ARGs were predicted in total, with Donor 7 mice predicted to harbour more antibiotic resistance genes than Donor 2 (17 vs. 9); this includes the presence of a Class D beta-lactamase (Oxa-347) in the Donor 7 mice.