Metagenomic, Genomic and Functional Genomic Approaches for the Characterisation of Antibiotic Resistance

by

Ross Stuart McInnes

A thesis submitted to the University of Birmingham for the degree of Doctor of Philosophy

Institute of Microbiology and Infection
College of Medical and Dental Sciences
University of Birmingham
April 2022

Abstract

Antibiotic resistance is an emerging threat to global public health. To tackle antibiotic resistance, we must take a One Health approach in which we study antibiotic resistance in combined human, animal and environmental contexts. Here I studied antibiotic resistance at the environmental, clinical and molecular level using different genomics-based approaches. Bacterial culturing and metagenomics were used to uncover that a major reservoir of antibiotic resistance in Bangladesh was the urban surface water environment and that the discharge of untreated human waste into these water bodies was the possible source of antibiotic-resistant bacteria. Short and long-read whole genome shotgun sequencing were used to reveal the structure of a hospital vancomycin resistant *Enterococcus faecium* (VRE) outbreak and to examine the genetic background of vancomycin-variable *E. faecium* (VVE) isolates found in this environment. Long-read sequencing and RT-qPCR were also used to uncover that the VVE isolates could rapidly revert to a vancomycin-resistant phenotype by co-transcribing the *vanHAX* vancomycin resistance genes with one of the rRNA operons. Transcriptomics and transposon insertion sequencing were leveraged to identify genes important for the fitness of VRE strains in the presence of vancomycin at sub-inhibitory concentrations. Exposure to vancomycin led to wide-ranging alterations in the transcriptional program of *E. faecium*, particularly the downregulation of phosphotransferase system genes. A currently uncharacterised gene *avrB,* was identified by both RNA-seq and Tn-seq and shown to confer a growth advantage to *E. faecium* E745 in the presence of vancomycin at ½ minimum inhibitory concentration compared to a deletion mutant. This study demonstrated that important insights can be gained by taking a One Health approach when studying antibiotic resistance. These data suggest that increased sanitation in Bangladesh could prevent the high levels of antibiotic resistance in the urban

environment. Increased sequencing of *E. faecium* in the clinical setting may be required in the future to prevent treatment failures caused by vancomycin-variable *E. faecium* strains. Novel targets were identified for the development of drugs that can be used in conjunction with vancomycin for the treatment of *E. faecium* infections.

For June and Bert

Acknowledgments

I would firstly like to thank the Wellcome Trust for providing the funding for this PhD, without which none of this work would have been possible. Thank you to my supervisors Professor Willem van Schaik and Professor Alan McNally for their support and guidance. I would also like to thank all of the members of NG01/NG20 (the McSchaiks) for their support and their wealth of knowledge, particularly to Dr Steven Dunn and Dr Robert Moran. Thank you to all the members of ARM past and present who have provided support when the going got tough. I'd like to say a particularly big thank you to Dr Lisa Lamberte, who is the only member of ARM that has been with me since my first day in the lab. Thank you to my parents, Stuart and Lynn, who have always encouraged and supported me in my academic endeavours. A special thank you goes to my partner Lizzy Darby, who has supported me both personally and academically throughout my PhD.

"A Man's a Man for a' that"

- Robert Burns

Publications and conference presentations arising from this thesis

Publications

McInnes, R.S. *et al.* (2020) 'Horizontal transfer of antibiotic resistance genes in the human gut microbiome', *Current Opinion in Microbiology*, 53, pp. 35–43. doi:10.1016/j.mib.2020.02.002.

McInnes, R.S. *et al.* (2021) 'Metagenome-Wide Analysis of Rural and Urban Surface Waters and Sediments in Bangladesh Identifies Human Waste as a Driver of Antibiotic Resistance', *mSystems*, 6(4), pp. e00137-21. doi:10.1128/mSystems.00137-21.

Conference presentations

Ross S. McInnes and Willem van Schaik. Construction of novel reporter systems to identify horizontal gene transfer of antimicrobial resistance genes in complex bacterial communities. Copenhagen Bioscience Conferences, Averting the post -antibiotic era – challenges and developments. Copenhagen, Denmark. 31st October – 3rd November 2018. Poster #69.

Ross S. McInnes, M. Hassan Zaman, Imam Taskin Alam, Siu Fung Stanley Ho, Niyaz Ahmed, M. Sirajul Islam and Willem van Schaik. Metagenomes and resistomes of rural and urban surface water in Bangladesh. British Society for Antimicrobial Chemotherapy, Antibiotic resistance and mechanisms. Birmingham, England. 28-29th November 2019. Poster #38.

Ross S. McInnes, M. Hassan Zaman, Imam Taskin Alam, Siu Fung Stanley Ho, Niyaz Ahmed, M. Sirajul Islam and Willem van Schaik. Metagenomes and resistomes of rural and urban surface water in Bangladesh. Microbiology Society Annual Conference. Edinburgh, Scotland. 31st March-3rd April 2020. Poster. (Cancelled due to COVID-19).

Ross S. McInnes, Vincent de Maat and Willem van Schaik. Identification of auxiliary vancomycin resistance genes in *Enterococcus faecium.* Microbiology Society Annual Conference. Online. 26-30th April 2021. Oral presentation.

Ross S. McInnes, Vincent de Maat and Willem van Schaik. Identification of auxiliary vancomycin resistance genes in *Enterococcus faecium.* World Enterococcus Day. Online. 15th July 2021. Oral presentation.

Declaration of authorship

Ross Stuart McInnes was the first author and main contributor of the publication:
McInnes, R.S. *et al.* (2021) 'Metagenome-Wide Analysis of Rural and Urban Surface Waters and Sediments in Bangladesh Identifies Human Waste as a Driver of Antibiotic Resistance', *mSystems*, 6(4), pp. e00137-21. doi:10.1128/mSystems.00137-21.

Results from this publication are presented in Chapter 3 of this thesis.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ABC	ATP-binding Cassette
Abs	Absorbance
AMR	Antimicrobial Resistance
ANI	Average Nucleotide Identity
ANOVA	Analysis of Variance
ARB	Antibiotic Resistant Bacteria
ARG	Antibiotic Resistance Gene
ATP	Adenosine Triphosphate
BHI	Brain Heart Infusion
bp	Base pair(s)
CFU	Colony Forming Units
CLIMB	Cloud Infrastructure for Microbial Bioinformatics
COG	Cluster of Orthologous Groups
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DGE	Differential Gene Expression
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
ESBL	Extended Spectrum beta-lactamase
ESKAPE	*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter spp*.
EtOH	Ethanol
EUCAST	European Committee on Antimicrobial Susceptibility Testing
FC	Fold Change
FDR	False Discovery Rate
g	Gram
GDP	Gross Domestic Product
HGT	Horizontal Gene Transfer
HITS	High-throughput Insertion Tracking by deep Sequencing
HMW	High Molecular Weight
h	Hour(s)
ICE	Integrative and Conjugative Elements
IDT	Integrated DNA Technologies
IPCR	Inverse Polymerase Chain Reaction
IQR	Interquartile Range
IS	Insertion Sequence
KO	Knockout
KPC	*Klebsiella pneumoniae* Carbapenemase
LB	Lysogeny Broth
LMIC	Low- and Middle-Income Countries
LRE	Linezolid Resistant Enterococci
Mbp	Mega-base pair(s)
MDR	Multidrug-Resistant
mg	Milligram
MGE	Mobile Genetic Element

MIC	Minimum Inhibitory Concentration
min(s)	Minute(s)
ml	Millilitre
MLST	Multi-Locus Sequence Type
MV	Membrane Vesicles
NDM	New-Delhi Metallo beta-lactamase
NEB	New England Biolabs
ng	Nanogram
NGS	Next Generation Sequencing
NMDS	Non-metric Multidimensional Scaling
NP	Not Present
NS	Not Significant
OD	Optical Density
ONT	Oxford Nanopore Technologies
ORF	Open Reading Frame
PAM	Protospacer Adjacent Motif
PBP	Penicillin Binding Protein
PBS	Phosphate Buffered Saline
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
PEG	Polyethylene Glycol
PERMANOVA	Permutational Multivariate Analysis of Variance
PTS	Phosphotransferase System
PRP	Pheromone-Responsive Plasmid(s)
QC	Quality Control
RACE	Rapid Amplification of cDNA Ends
RCR	Rolling Circle Replication
RNA	Ribonucleic Acid
RPKM	Relative abundance in reads Per Kilobase of reference sequence per Million sample reads
RPM	Revolutions Per Minute
RT	Room Temperature
RT-PCR	Reverse Transcription Polymerase Chain Reaction
s	Second(s)
SBS	Sequencing by Synthesis
SMRT	Single Molecule Real-time
SNP	Single Nucleotide Polymorphism
SPI	Signal Peptidase I
SPRI	Solid Phase Reversible Immobilization
ST	Sequence Type
TBE	Tris-Borate-EDTA
TIS	Transposon Insertion Sequencing
Tm	Melting temperature
TSS	Transformation and Storage Solution
UK	United Kingdom
URT	Universal Repair Template
USA	United States of America
UTI	Urinary Tract Infection

VRE	Vancomycin Resistant Enterococci
VVE	Vancomycin Variable Enterococci
WGS	Whole Genome Shotgun sequencing
WT	Wildtype
ya	Years ago

Chapter 1
General introduction

1.1 Antibiotics and Antibiotic Resistance

1.1.1 History of antibiotics

Antibiotics are drugs which can either kill or inhibit the growth of bacteria (Baquero and Levin, 2021). In 1909, the first modern antibiotic drug, the organoarsenic compound Salvarsan, was developed for the treatment of syphilis (Sepkowitz, 2011) (Figure 1.1). Although Salvarsan was widely used, it had major drawbacks in terms of its stability and hepatotoxicity (Zaffiri, Gardner and Toledo-Pereyra, 2012). The Salvarsan-based drugs were followed in 1935 by the sulphonamide drug Prontosil. This prodrug had activity against Gram-positive organisms and was used worldwide for the treatment of *Streptococcus pyogenes* infections (Millett, 1937). The first beta-lactam drug penicillin was discovered by Sir Alexander Fleming in 1928 but its utility as a therapeutic was initially overlooked due to issues with its production and scalability (Fleming, 1929; Gaynes, 2017). It was not until 1942 when Howard Florey and Norman Heatley made significant breakthroughs in the mass production of penicillin that it was recognised as a viable drug for the treatment of Gram-positive bacterial infections (Chain *et al.*, 1940). Its mass production allowed it to be used extensively in the second half of World War II and is still in use today (Quinn, 2013). The development of antibiotics falls into two main periods, the "Golden Age" of antibiotic development between 1940 and 1960 and the "lean years" from 1960 until the present day. During the Golden Age many of the clinically relevant antibiotics which we use today such as the beta-lactams, aminoglycosides and glycopeptides were developed (Gould, 2016). However, after 1960 few new classes of antibiotics were in production and it was not until the early 2000s that the oxazolidinones and lipopeptide antibiotics were approved for clinical use (Nicolaou and Rigol, 2018).

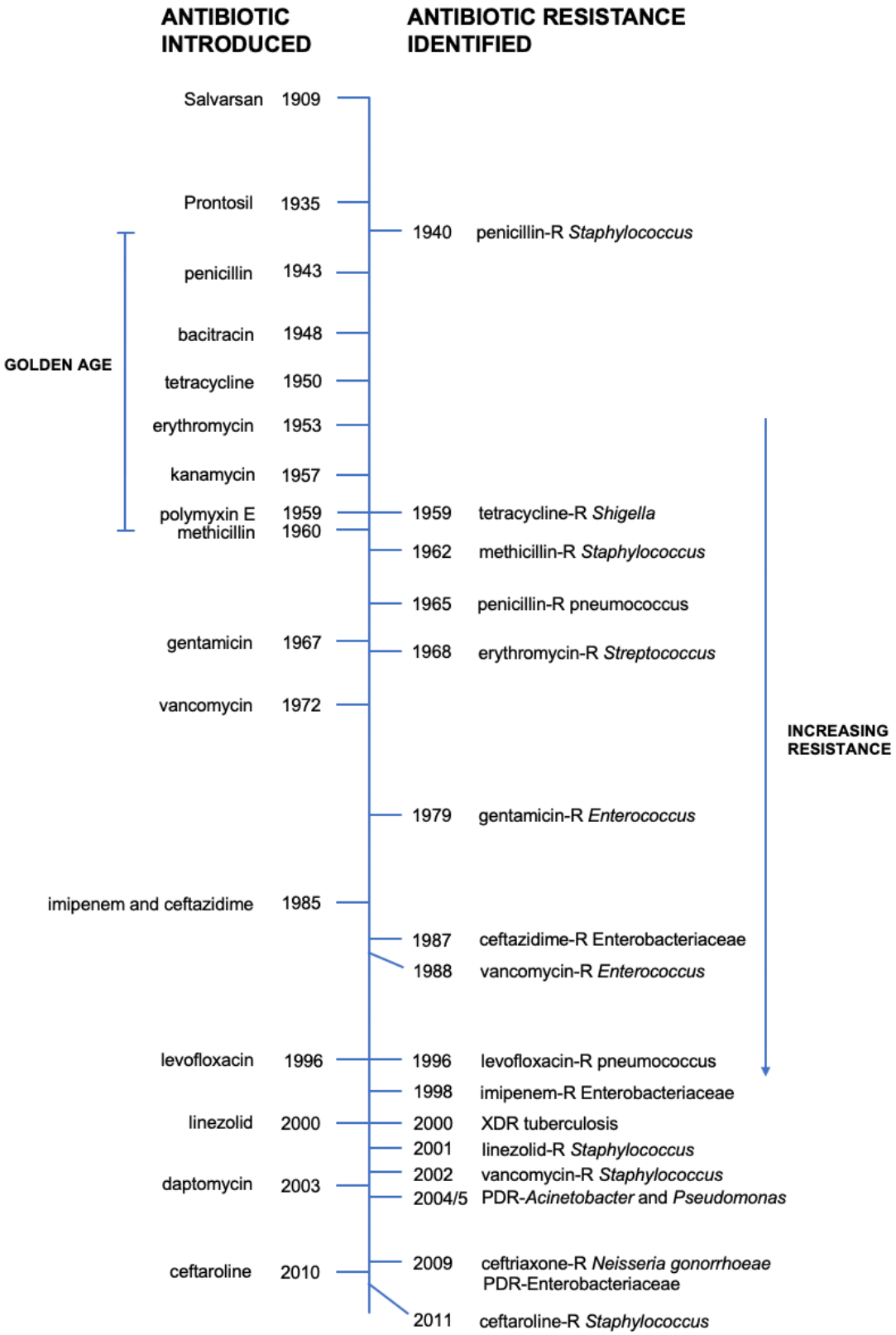

Figure 1.1: A timeline of notable antibiotics and the subsequent rise of resistance to these antibiotics. Figure adapted from (Centers for Disease Control & Prevention, 2013).

1.1.2 Antibiotic resistance

Antimicrobial resistance (AMR) is the resistance of fungi, parasites, viruses and bacteria to the therapeutic drugs used to treat them (Prestinaci, Pezzotti and Pantosti, 2015). Resistance to antimicrobials, and in particular to antibiotics, is an increasing public health concern (Ferri *et al.*, 2017). It has been reported that by 2050 10 million deaths per year could be attributed to antimicrobial resistance (O'Neill, 2014). This is a worst-case scenario and likely overestimates deaths due to confounding factors in estimation. However it does illustrate the growing threat of AMR to public health (de Kraker, Stewardson and Harbarth, 2016). A recent study estimated that in 2019, 4.95 million deaths were associated with antibiotic resistant bacteria (ARB) and 1.27 million deaths were directly attributable to ARB (Murray *et al.*, 2022). Antimicrobial resistance not only affects public health but also has a negative impact on the global economy. The World Bank produced a report titled "Drug resistant infections a threat to our economic future" which modelled the impact of antimicrobial resistance on global GDP. It demonstrated in its high-impact scenario that global GDP could drop by 3.8% and that low-income countries would be worst affected (Jonas *et al.*, 2017).

Multiple human factors contribute to the increase in AMR, including over-prescription of drugs, failure to finish courses of drugs and the use of clinically relevant drugs in agriculture (Michael, Dominey-Howes and Labbate, 2014). It should also be noted that even the appropriate use of antibiotics will select for drug-resistant bacteria. The issues linked to AMR are likely to increase as few new classes of antimicrobial drugs which target priority pathogens are currently in development (Theuretzbacher *et al.*, 2020).

Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa, and *Enterobacter* species are a group of bacteria

collectively known as the ESKAPE pathogens (Santajit and Indrawattana, 2016). These opportunistic pathogens are often associated with the hospital environment and are in many cases multidrug resistant (De Oliveira *et al.*, 2020). Drug resistant ESKAPE pathogens, particularly *S. aureus* and *K. pneumoniae* are a major cause of death globally (Murray *et al.*, 2022). These drug resistant pathogens use a variety of different mechanism to survive exposure to antibiotic drugs.

1.1.3 Mechanisms of antibiotic resistance

Bacteria have evolved many different strategies to resist antibiotic drugs. The mechanisms of resistance can be grouped into three major categories namely; prevention of the antibiotic from reaching toxic levels inside the cell, modification of the antibiotic target and modification of the antibiotic itself (Blair *et al.*, 2015) (Figure 1.2).

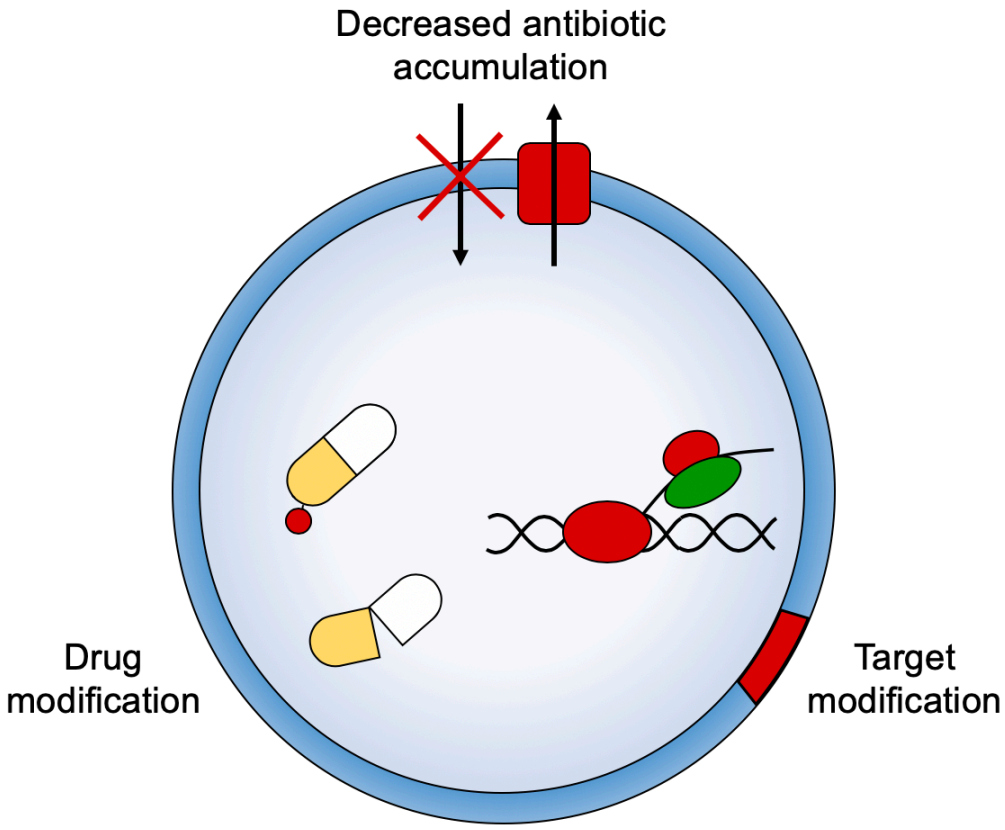

Figure 1.2: Mechanisms of antibiotic resistance. Drug modification: Antibiotics can be degraded or modified by bacterial enzymes. Decreased antibiotic accumulation: Increased expression of efflux pumps and decreased expression of porins (in Gram-negative bacteria) prevent accumulation of the antibiotic inside the bacterial cell. Target modification: Mutations in DNA encoding the antibiotic target result in the antibiotic no longer being able to bind.

Antibiotics may enter the bacterial cell through either the membrane or through porin proteins in the outer membrane of Gram-negative bacteria. Bacteria can resist antibiotics entering the cell by downregulating the expression of the genes encoding porin proteins or

by having mutated porin proteins that obstruct the passage of antibiotics into the cell. Bacteria also actively pump antibiotics out of the cell using a variety of different efflux pumps which can also be upregulated during antibiotic treatment (Delcour, 2009). Modification of the antibiotic target is a strategy that is employed by many bacteria and is often a self-resistance mechanism of cells that produce antibiotics. Two examples of this strategy provide resistance to the cell wall targeting antibiotics penicillin and vancomycin. Mutations in the penicillin binding protein (PBP) genes of bacteria can produce PBPs that have a much lower binding affinity for penicillin than the wildtype proteins. Resistance to vancomycin is conferred by an alteration of D-Ala-D-Ala residues to D-Ala-D-Lac or D-Ala-D-Ser in peptidoglycan precursor molecules which greatly reduces the binding affinity of vancomycin (Peterson and Kaur, 2018). The final strategy employed by bacteria to resist antibiotics is the modification or destruction of the antibiotic itself. Resistance to aminoglycosides is conferred by acetyltransferase enzymes which acetylate aminoglycoside drugs leading to steric hinderance at the target site, thus lowering the antibiotic's binding affinity to its target. Other enzymes such as beta-lactamases cleave bonds within the drug which in turn stops them binding to their target (Munita and Arias, 2016). Antibiotic resistance genes are found on both the chromosome and on mobile genetic elements (MGEs) such as plasmids and phage. These MGEs are capable of exchanging genetic material, including antibiotic resistance genes, by different mechanisms of horizontal gene transfer. The most important of which are, transformation (the uptake of naked DNA from the environment), conjugation (the transfer of plasmids and integrative and conjugative elements between strains) and transduction (the transfer of DNA between strains via a bacteriophage) (McInnes *et al.*, 2020).

1.1.4 Global spread of antibiotic resistance genes

There have been many examples of novel resistance genes arising and subsequently being spread globally. An example of this process is provided by the spread of resistance to the antibiotic colistin. Colistin was generally avoided for use in humans due to its toxicity. However, as organisms have become increasingly resistant to other drugs, colistin was brought in for use in humans as a drug of last resort (Zavascki *et al.*, 2007). Resistance to colistin had previously been observed due to mutations in chromosomally encoded genes including *mgrB* and *pmrB* (Cannatelli *et al.*, 2014; Berglund *et al.*, 2018). In 2015, *mcr-1* a mobilizable colistin resistance gene was detected and by 2017 the gene had been detected on five continents (Liu *et al.*, 2016; Wang *et al.*, 2018). A similar situation was seen with the beta-lactamase gene bla_{NDM-1} which was first identified in 2008 and within a few years had spread globally (Yong *et al.*, 2009; Khan, Maryam and Zarrilli, 2017). Mutations in an antibiotics target site can also lead to the global emergence of antibiotic resistance. Ampicillin resistance in *Enterococcus faecium* is conferred by the expression of a mutated penicillin-binding protein PBP5 which has a reduced affinity for β-lactam antibiotics. Approximately 90% of all clinical *E. faecium* isolates are resistant to ampicillin because they possess this low affinity protein whereas in *Enterococcus faecalis*, another important hospital associated pathogen, resistance rates are around 2% (Gagetti *et al.*, 2019). The spread of resistance genes can be mediated either by the spread of the bacterial host and/or the spread of the resistance gene itself on mobile genetic elements (McInnes *et al.*, 2020). Resistance genes that are chromosomally encoded are usually spread vertically and are inherited by daughter cells. However, transduction mechanisms can mobilise chromosomal genes via bacteriophage, and in some cases insertion sequences can mobilise chromosomally encoded genes on to plasmids allowing them to spread horizontally (Ebmeyer, Kristiansson

and Larsson, 2021). Antibiotic resistance genes which are found on mobile genetic elements such as plasmids, integrative and conjugative elements (ICE) and bacteriophages can be shared horizontally with other bacteria (Partridge *et al.*, 2018). The horizontal transfer of antibiotic resistance genes occurs often in microbial ecosystems due to the close spatial relationship of bacteria and the large numbers of potential donors and recipients (van Elsas and Bailey, 2002).

1.2 Metagenomics

Bacteria live in complex ecosystems that can be studied using modern DNA-sequencing technologies. These technologies can be leveraged to perform metagenomics, which is the study of the total genetic content of an ecosystem, or the metagenome (Thomas, Gilbert and Meyer, 2012). There are many different techniques that can be used to identify organisms that contribute to a metagenome, and these are typically grouped into culture-dependent and culture-independent techniques (Pandya *et al.*, 2017). The simplest way to identify the organisms that contribute to a microbial ecosystem is to culture them on selective media and sequence their genomes. However, many microorganisms that are found in natural environments cannot be cultured in the laboratory so would be missed using routine culturing approaches (Bodor *et al.*, 2020). To overcome this problem, a group of techniques termed culturomics were developed which leveraged high-throughput methods to culture organisms in hundreds or thousands of different conditions increasing the likelihood of their isolation (Lagier *et al.*, 2012). Although culturomics vastly increased the number of organisms that could be isolated, it still failed to identify organisms which had unknown growth requirements (Lagier *et al.*, 2018).

Culture-independent techniques have now been developed that do not require isolation of the microorganisms thus overcoming many of the drawbacks of culture-dependent approaches (Forbes *et al.*, 2017). There are two major culture-independent approaches to identify microorganisms in the environment, both of which use high-throughput sequencing to sequence the microbial DNA (Semenov, 2021). The first technique is meta-barcoding, this technique uses the amplification and sequencing of marker genes to identify and quantify the organisms present in a microbiome (Huggins *et al.*, 2019). A common marker gene used for the identification of bacteria in a sample is the 16S ribosomal RNA gene which

contains nine variable regions with the 250 bp V4 region being most commonly used (Bukin *et al.*, 2019). Metabarcoding is an extremely useful technique to identify organisms in an environment as it is relatively low-cost and extremely high-throughput (Elbrecht *et al.*, 2017). Metabarcoding does, however, have a number of drawbacks. Due to the short length of the variable regions that are typically sequenced in a metabarcoding study, it is not always possible to identify organisms to the species level. (Shahi *et al.*, 2019). Due to their low abundance these short regions of DNA must be amplified by PCR, which can lead to sequencing errors and can preclude the identification of certain organisms. PCR can also introduce biases through primer-mismatch and preferential target amplification skewing community composition estimates (Sze and Schloss, 2019). Arguably, the greatest drawback of metabarcoding is that it only identifies the organisms, missing vital genomic information that can give insights into the biology of these organisms, including any virulence or antibiotic resistance genes that they may carry (Bharti and Grimm, 2021).

Shotgun metagenomic sequencing is a technique that is used to sequence all of the DNA in a sample rather than just the marker genes (Quince *et al.*, 2017). By sequencing all of the DNA in a sample it is not only possible to identify and quantify the organisms present, but also to identify all other genes of interest in the sample (Chiu and Miller, 2019). Sequencing of all DNA also makes it possible to identify organisms from different kingdoms whereas in metabarcoding the marker genes are often kingdom-specific such as the 16S rRNA gene for bacteria or the ITS gene for fungi (Edgar, 2018). Metagenomes that have been sequenced using shotgun metagenomic sequencing to a high read depth, can also be assembled to construct metagenome assembled genomes (MAGs) which can be near-complete genome sequences of organisms in a microbial ecosystem (Meziti *et al.*, 2021). The advantage of recovering MAGs from a sample is that it reveals the functional potential of specific

organisms in an environment (Liu *et al.*, 2020). Metagenomics has been leveraged to profile the metagenomes of both human-associated "built environments", such as homes and office spaces, and natural environments including soil and water.

1.2.1 Metagenomes of the built environment

Many studies have looked at the microbiome of the home as a model for a human associated microbiome (Jeon, Chun and Kim, 2013; Adams *et al.*, 2014; Lax *et al.*, 2014; Casas *et al.*, 2019). These studies identified that the microbiome of the home was mainly associated with the bacteria carried by their human occupants. It was also found that specific areas of the home could be linked to specific occupants and that the transmission of these bacteria could be traced throughout the household (Lax *et al.*, 2014). It has also been shown that domestic animals such as cats and dogs can have a significant impact on the microbiome of the home (Fujimura *et al.*, 2010). Highly controlled built environments have also been associated with human and environmental microbiomes. The cleanroom used to construct the Mars 2020 Perseverance Rover was found to be contaminated with human-associated bacteria but as the distance increased from the entrance of the cleanroom, the bacteria were more associated with the soil environment (Hendrickson *et al.*, 2021). Operating theatres also have human-associated microbiomes, bacteria from the walls of the operating theatres were most associated with human skin but this varied spatially and the lower walls were more associated with human faeces (Derilus *et al.*, 2020).

1.2.2 Metagenomes of the natural environment

As well as the built environment, the natural environment has been studied in great depth, particularly the microbiomes of sediment and water. Although these environments directly interact with each other, they contain very different bacterial communities (Huang *et al.*,

2018). In freshwater, sediment samples are often dominated by bacteria belonging to the phyla Proteobacteria while the corresponding water samples are dominated by bacteria belonging to the phyla Cyanobacteria (Huang *et al.*, 2017; Marmen *et al.*, 2021). It has also been shown that the microbial composition of these environments can be altered dramatically through natural processes like the change of seasons or by human processes such as runoff from agriculture (Wilhelm *et al.*, 2014; Beattie *et al.*, 2020). The discharge of human waste into urban rivers is associated with high environmental levels of human gut associated bacteria. These human associated bacteria often carry antibiotic resistance genes that confer resistance to clinically relevant antibiotics (Hiruy *et al.*, 2022).

1.2.3 Resistomes

The resistome is the total genetic material in a sample that confers resistance to antibiotic drugs (van Schaik, 2015). DNA sequences assembled from metagenomic sequencing can be queried against antibiotic resistance gene databases such as ResFinder (Zankari *et al.*, 2012) or CARD (Alcock *et al.*, 2020) which contain known antibiotic resistance genes. This allows the identification of the antibiotic resistance genes (ARGs) present in a sample and to identify whether ARGs are enriched for certain antibiotic classes (Gupta, Tiwari and Cytryn, 2020).

The resistomes of the built environment have been studied extensively. Antibiotic resistance genes are ubiquitous in the built environment and in some cases levels of antibiotic resistance genes in household dust can exceed that of soil contaminated with sewage (Ding, Zhou and Zhu, 2020). High levels of cleaning in clinical environments has also been linked to a shift in the bacterial communities to a less diverse state with higher levels of antibiotic resistance (Mahnert *et al.*, 2019). “One Health” is a multidisciplinary approach to improve health by targeting human, animal and environmental health particularly at its intersection (McEwen

and Collignon, 2018)(Figure 1.3). Using this approach, many studies have shown that antibiotic resistance genes are present at high levels in lakes and rivers used for drinking water and recreational activities (Lee *et al.*, 2020; Han *et al.*, 2022). The use of antibiotics in aquaculture and the release of untreated human waste into environmental water bodies provides selection for antibiotic resistant bacteria and also introduces antibiotic resistant bacteria directly into the environment (Karkman, Pärnänen and Larsson, 2019; Stanton *et al.*, 2020). This is particularly relevant for developing and low- and middle-income (LMIC) countries where policies surrounding antibiotic use and sanitation infrastructure are generally not as developed as those in high-income countries (Tattevin *et al.*, 2020).

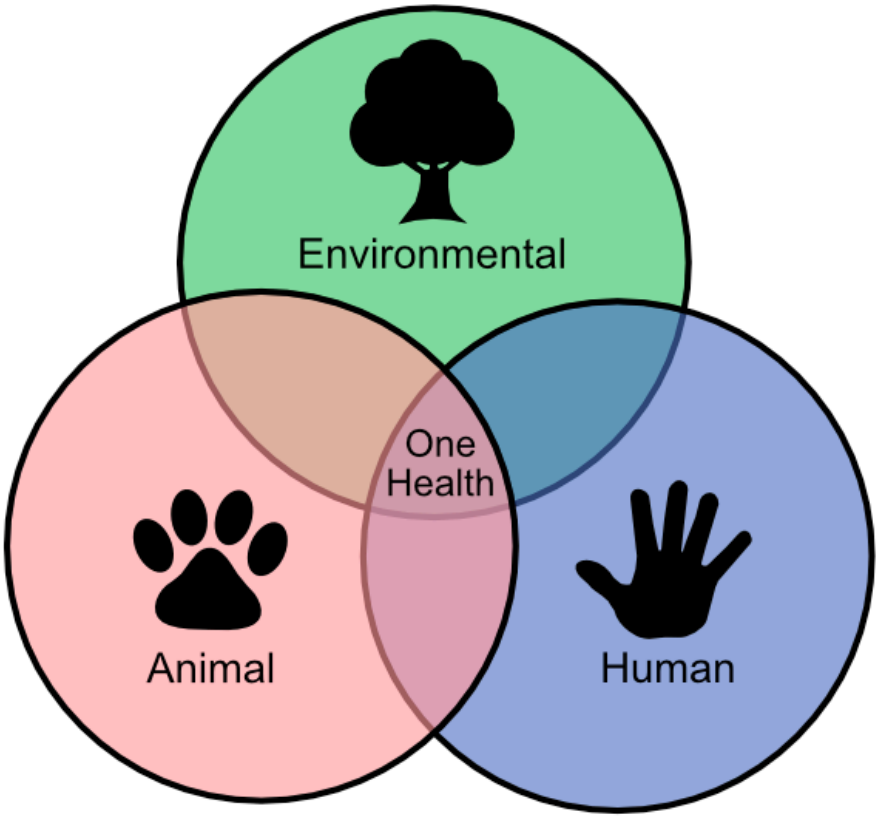

Figure 1.3: One Health: One Health is the approach of looking at health not only from the human perspective but also animals and the environment.

1.3 Bacterial Genomics

Bacterial genomics is the study of the bacterial genome, including chromosomal and extrachromosomal elements. It can be used to study different features of the genome such as genes and larger elements such as transposons and plasmids (Raskin *et al.*, 2006). Bacterial genomics can also be used to compare the genomes of closely related strains to understand the evolution of a species or to track an outbreak of bacterial infections (Van Goethem *et al.*, 2019; Koonin, Makarova and Wolf, 2021). The advent of next generation sequencing (NGS) has made bacterial genomics accessible to most laboratories which has enabled the sequencing of hundreds of thousands of bacterial genomes (Blackwell *et al.*, 2021). NGS is a group of techniques that can be used to sequence DNA and can be broadly grouped into short-read sequencing and the more recently introduced long-read sequencing technologies (Goodwin, McPherson and McCombie, 2016).

1.3.1 Short-read sequencing

Short-read sequencing is the highly parallel sequencing of short fragments of DNA around 250 bp long. The most widely used short-read sequencing technology is from Illumina which employs a method called “Sequencing by Synthesis” (SBS) (McCombie, McPherson and Mardis, 2019). DNA which has been extracted from an organism of interest is sheared into short fragments that are first hybridised to a flow cell. Fluorescently labelled nucleotides are passed over the flow cell and the fluorescent signal is recorded as they are incorporated into the growing chain. The fluorescent signals can then be translated into nucleotide bases to obtain the DNA sequence of the fragment. These short fragments of DNA sequence are then assembled to give the entire genome sequence of the organism (Modi *et al.*, 2021). This technology can also be used to study the functional capacities of the genome which will be discussed further in sections 1.4 (Transposon Insertion Sequencing) and 1.5

(Transcriptomics). Short-read sequencing with Illumina technology is high throughput, relatively inexpensive and highly accurate, producing less than 1 error per 1000 bases (Tan *et al.*, 2019). A drawback of short-read sequencing is a function of its short read length. Bacterial genomes, particularly those of the low-GC Gram-positive organisms contain many repetitive elements such as insertion sequences (IS). Short-reads cannot distinguish between different copies of the same element as they have no genomic context, which can lead to the fragmentation of short-read assemblies Important information about the genome can be lost (Makałowski *et al.*, 2019). To overcome these challenges, new technologies have been developed which generate longer read lengths.

1.3.2 Long-read sequencing

Long-read sequencing is an approach that allows the sequencing of considerably longer fragments of DNA than short-read sequencing can achieve. Unlike short-read sequencing where the Illumina platform has a monopoly on the market, there are two leading long-read sequencing technologies. Pacific Biosciences' Single Molecule Real-time (SMRT) sequencing technology uses an approach similar to Illumina in that it is a sequencing-by-synthesis approach and fluorescence is measured as each base is incorporated in the growing DNA strand. SMRT sequencing can achieve DNA reads lengths of around 10 kbp (Logsdon, Vollger and Eichler, 2020). Oxford Nanopore Technologies use a radically different approach to sequencing, instead of sequencing-by-synthesis, the DNA is sequenced directly. The DNA is unwound by a helicase enzyme and then a single strand of DNA is fed through a pore in a conductive membrane. As each base is passed through the membrane the change in current is measured. These changes in current can be translated into nucleotide bases giving the sequence of the DNA strand (Mantere, Kersten and Hoischen, 2019). A major advantage of long-read sequencing is that the read length can be longer than repetitive

elements, meaning that the genome can be assembled in its entirety, and information is not lost (Amarasinghe *et al.*, 2020). Theoretically, DNA fragments of any length can be sequenced using nanopore technology but, in reality, this is limited by the integrity of the DNA. It is now commonplace to achieve read lengths of over 1 Mb which equates to around 1/3 of the entire genome of some bacteria (Payne *et al.*, 2019). A current disadvantage of nanopore sequencing is that it is less accurate than Illumina sequencing or PacBio SMRT sequencing with a read accuracy of 1 error per 50-100 bp whereas these other platforms regularly achieve error rates of 1 error per 1000 bp (Hu *et al.*, 2021).

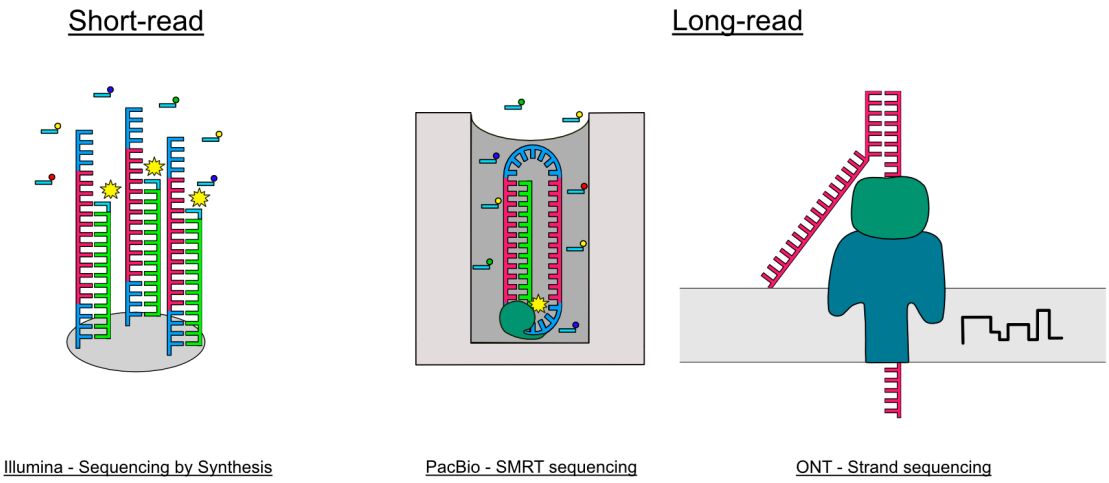

Figure 1.4: Next generation sequencing technologies. Illumina short-read sequencing using cluster generation and sequencing by synthesis. Pacific Biosciences (PacBio) SMRT sequencing also using a sequencing by synthesis approach. Oxford Nanopore Technologies (ONT) sequencing uses changes in current as DNA passes through a nanopore to sequence DNA.

1.3.3 Hybrid assembly – The best of both worlds

Short and long-read technologies can be used in a complementary manner to overcome both of their shortcomings. The more inaccurate long reads can be used to create a highly contiguous genome assembly and the higher quality short reads can then be used to remove any errors in the assembly (Chen, Erickson and Meng, 2020). Hybrid assembly can produce a complete and highly accurate genome with one replicon per contig, which is especially useful for studying plasmid sequences (Wick *et al.*, 2017).

1.4 Transposon insertion sequencing

Transposon insertion sequencing (TIS) is a functional genomics method that can be used to query gene essentially and function in a bacterium (van Opijnen and Camilli, 2013). Several different implementations of TIS exist, but they all involve the creation of a mutant library in which a transposon with a selectable marker is distributed throughout the genome such that there are many insertions within every gene in the genome, except those which are essential for survival (Larivière *et al.*, 2021) (Figure 1.5). The mutant pool is then exposed to a test condition such as the presence of an antibiotic and a control condition in which selection for mutants is minimal. DNA is then extracted and a library is prepared such that the end of the transposon and a small region of the genome adjacent to the transposon is sequenced (Cain *et al.*, 2020). Reads can then be mapped to the parent strain's genome and the transposon insertion site can be identified. The reads/insertions can then be summed for each gene in the genome and different methods used to compare genes in the test condition compared to the control (Chao *et al.*, 2016).

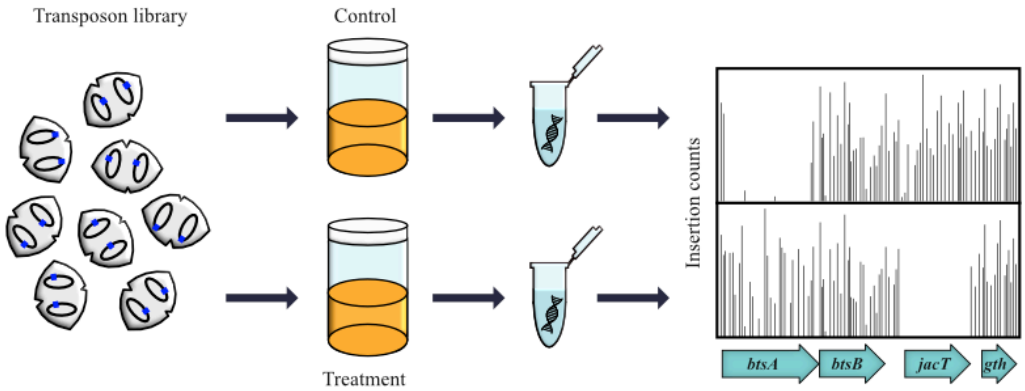

Figure 1.5: Transposon insertion sequencing method. The transposon library is grown in a treatment condition such as an antibiotic and a control condition. The DNA is then extracted, and the sequencing library is prepared according to the specific TIS method. The libraries are then sequenced, and the reads are aligned to the parent genome. Genes that contribute to survival in the treatment condition have fewer mapped reads than the control, in this example the hypothetical gene *jacT* is essential for survival in the treatment condition. Genes that contribute to survival in the control condition have fewer mapped reads than the treatment, in this example the hypothetical gene *btsA* is essential for survival in the control condition.

1.4.2 Different types of transposon sequencing

Transposon insertion sequencing (TIS) has four main incarnations; Tn-seq, TraDIS, INSeq and HITS (Cain *et al.*, 2020). Tn-seq and INSeq are highly similar techniques which both use *mariner* transposons that non-specifically target thymidine-adenine (TA) dinucleotide motifs. Both techniques use a MmeI restriction site in the inverted repeat region of the transposon to capture a 16 bp (INSeq) or 20 bp (Tn-seq) region of genomic DNA adjacent to the transposon (Goodman *et al.*, 2009; van Opijnen, Bodi and Camilli, 2009). TraDIS does not use a mariner transposon and instead uses a Tn5 transposon derivative which inserts into the genome non-specifically. TraDIS does not utilise restriction enzymes but instead uses sonication to shear the DNA into varying lengths (Langridge *et al.*, 2009). Random shearing of the DNA is useful as it creates longer reads which are easier to map back to the genome. However, as the reads are different lengths this can introduce a PCR bias into the analysis, which is not the case for the uniform read lengths produced by Tn-seq and INSeq (van Opijnen and Camilli, 2013). HITS has features of both Tn-seq/INSeq and TraDIS, as it uses a *mariner* transposon to create insertions but uses random shearing of the DNA prior to adapter ligation rather than restriction enzymes. Unlike the other TIS techniques, HITS employs a streptavidin based selection step to recover only fragments of DNA which contain a transposon genome junction (Gawronski *et al.*, 2009).

When the techniques Tn-seq and TraDIS were first described they were used for the identification of essential genes in *Streptococcus pneumoniae* and *Salmonella* Typhi respectively (Langridge *et al.*, 2009; van Opijnen, Bodi and Camilli, 2009). INSeq leveraged transposon insertion sequencing to identify the genes that were essential for colonisation of the gut environment by *Bacteroides thetaiotaomicron* both in the presence and absence of a resident microbiome (Goodman *et al.*, 2009). HITS was originally used to identify genes in

Haemophilus influenzae which contributed to its delayed clearance in a murine pulmonary model (Gawronski *et al.*, 2009).

Since the original publication of the TIS techniques in 2009, they have been used to identify genes which contribute to increased fitness in many different species and conditions. A major use of TIS has been the identification of virulence genes. TIS techniques were able to identify that a type III secretion system was important for *Burkholderia pseudomallei* to cause respiratory melioidosis (Gutierrez, Yoder-Himes and Warawa, 2015). Similarly, type III secretion systems have been linked to survival of enterohaemorrhagic *E. coli* in an animal model (Warr *et al.*, 2019). Possibly the most widespread use of TIS is the identification of genes which contribute to antibiotic resistance against both old and new antibiotics in many of the ESKAPE pathogens (Cain *et al.*, 2020). TraDIS identified that efflux genes *adeIJK* and lipooligosaccharide synthesis genes *lpxC* and *lpsO* were required for survival of *Acinetobacter baumannii* when exposed to the polymyxin drug colistin (Boinett *et al.*, 2019). TIS was also able to help characterise the mechanism of action of the lysocin drugs, identifying that they work through binding to lipid II (Santiago *et al.*, 2018). Transposon sequencing has also been used to identify genes which contribute to antibiotic resistance in *Enterococcus faecalis*. A transposon insertion library of multidrug resistant strain, *E. faecalis* MMH594, was exposed to 10 different antibiotics at $1/8^{th}$ MIC. Genes were identified which were important for survival in the presence of ceftriaxone, polymyxin and daptomycin but not for vancomycin. These data suggests that in *E. faecalis* vancomycin resistance is conferred solely by acquired genes (Gilmore *et al.*, 2020).

1.5 Transcriptomics

Transcriptomics is the study of all of the RNA transcripts in a cell usually excluding the ribosomal RNA genes (Lowe *et al.*, 2017). Transcriptomics was originally performed using Sanger sequencing technology, but this was expensive, laborious and allowed only a fraction of the transcriptome to be studied at one time (Adams *et al.*, 1991). The first advance in transcriptomics was the development of microarrays which queried thousands of transcripts at one time, this dramatically reduced the price and labour (Baldi and Hatfield, 2002). However, microarrays also had drawbacks, as commercial microarrays were only available for model organisms and the dynamic range of gene expression that could be measured was narrow (Jaluria *et al.*, 2007; Zhao *et al.*, 2014). The advent of next generation sequencing opened the door for studies mapping the total transcriptional activity of cells. It is now possible to query the transcriptional profile of every gene in the genome relatively easily and cheaply using RNA sequencing (RNA-seq) (Wang, Gerstein and Snyder, 2009). Total RNA is extracted from the sample and DNA contamination is removed by treating the RNA with a DNA degrading enzyme. Ribosomal RNA makes up the majority of the RNA in a bacterial cell. It is removed using magnetic beads coated in oligonucleotides complimentary to the ribosomal RNA, leaving behind mRNA transcripts from all other transcribed genes. The RNA is then reverse transcribed into cDNA using random hexamer primers and the cDNA library can then be prepared for sequencing on next generation sequencing platforms using adapters specific for the sequencing platform (Kukurba and Montgomery, 2015). A recent advance in transcriptomics is the ability to directly sequence RNA without first reverse transcribing it to cDNA, with Oxford Nanopore Technologies developing a kit which allows RNA to be directly sequenced on their long-read sequencing platforms (Workman *et al.*, 2019).

Transcriptomics has been used widely in microbiology to investigate a wide array of processes in the bacterial cell. RNA-seq has been used extensively to probe the biology of bacteria during infection. It has identified genes and sRNAs in *Salmonella enterica* serovar Typhimurium that are important for the infection of murine macrophages (Srikumar *et al.*, 2015) and uncovered the mechanisms by which *Yersinia* spp. evade the host immune system when infecting the Peyer's patches. RNA-seq has also been used to investigate gene expression in biofilms of clinically important pathogens such as *Staphylococcus aureus, Campylobacter jejuni* and *Pseudomonas aeruginosa* (Tram *et al.*, 2020; D'Arpa *et al.*, 2021; Tomlinson, Malof and Shaw, 2021). As with biofilm formation, antibiotic resistance is another important phenotype that has been studied extensively in human associated pathogens using RNA-seq. Genes have been discovered for amikacin resistance in *A. baumanii* and tigecycline resistance in *Mycobacterium abscessus* (Qin *et al.*, 2018; Schildkraut *et al.*, 2021), and the two-component regulatory system *croRS* was shown to provide tolerance to teixobactin in *E. faecalis* (Darnell *et al.*, 2019).

1.6 *Enterococcus faecium*

The genus *Enterococcus* consists of a group of low GC% Gram-positive organisms belonging to the phylum Bacillota. The genus is thought to have emerged approximately 500 million years ago (Lebreton *et al.*, 2017). To date, 78 species in the genus *Enterococcus* have been described. They are found ubiquitously in nature and are often associated with the gastrointestinal tract of animals and insects (Hammerum, 2012; Parte *et al.*, 2020). *Enterococcus faecalis* and *Enterococcus faecium* are the species most frequently isolated from human infections (Zhou et al., 2020). Phylogenetically, the species *E. faecium* has been split into two major clades, clade A and clade B (Figure 1.6). Clade A is also split into two subclades; clade A1 which is associated with the hospital environment and clade A2 which is associated with animals. Clade B organisms are found in healthy individuals and are considered to be commensals of the human gastrointestinal tract (Lebreton *et al.*, 2013). Recently it has been suggested that the clade B isolates should be reclassified as *Enterococcus lactis* as their average nucleotide identity clusters more closely with *E. lactis* isolates than non-clade B *E. faecium* isolates (Belloso Daza *et al.*, 2021). A multilocus sequence typing (MLST) scheme for *Enterococcus faecium* was developed in 2002 and consists of seven housekeeping genes (*gdh*, *purK*, *pstS atpA*, *gyd*, *adk* and *ddl*) (Homan *et al.*, 2002). Mutations in these genes have, so far, given rise to 2,268 sequence types (ST) of *E. faecium* (Jolley, Bray and Maiden, 2018). These sequence types can be grouped into clonal complexes (CC), the most notable of which is CC-17. CC-17 is a hospital-associated lineage that contains isolates which often carry resistance to antibiotics including vancomycin, fluoroquinolones and aminoglycosides. As well as antibiotic resistance genes, CC-17 isolates often contain virulence factors particularly the enterococcal surface protein (Esp) which is thought to contribute to biofilm formation (Willems *et al.*, 2005; Top,

Willems and Bonten, 2008). Many different sequence types can be found in the hospital environment, however there are a few CC-17 sequence types that regularly cause outbreaks in the hospital environment. ST80 is currently the dominant sequence type in Stockholm, Sweden and has caused outbreaks in a number of hospitals (Fang *et al.*, 2021). This sequence type has also caused outbreaks of VRE in Australia and Ireland with some isolates additionally possessing linezolid resistance (Pratama *et al.*, 2021; Egan *et al.*, 2022). ST17 is another widespread hospital-associated sequence type that has caused several outbreaks in Spain, and also in Slovakia, Australia, Columbia and others (Valdezate *et al.*, 2009, 2012; Hughes *et al.*, 2019; Rios *et al.*, 2020; Jozefíková *et al.*, 2022). Like ST80 and ST17, ST78 is another CC-17 sequence type that is associated with hospital outbreaks in several countries including those in South East Asia and South America. Outbreaks of ST78, have involved both *vanA* and *vanB*-type vancomycin resistance (Hsieh *et al.*, 2010; Khan *et al.*, 2010; Yang *et al.*, 2016). *E. faecium* is a lactic acid producing bacterium that has been used widely in food production to flavour fermented products such as cheese and cured meats (Hanchi *et al.*, 2018). Despite their use in food manufacturing, enterococci are opportunistic human pathogens that can cause a variety of diseases (Gao, Howden and Stinear, 2018).

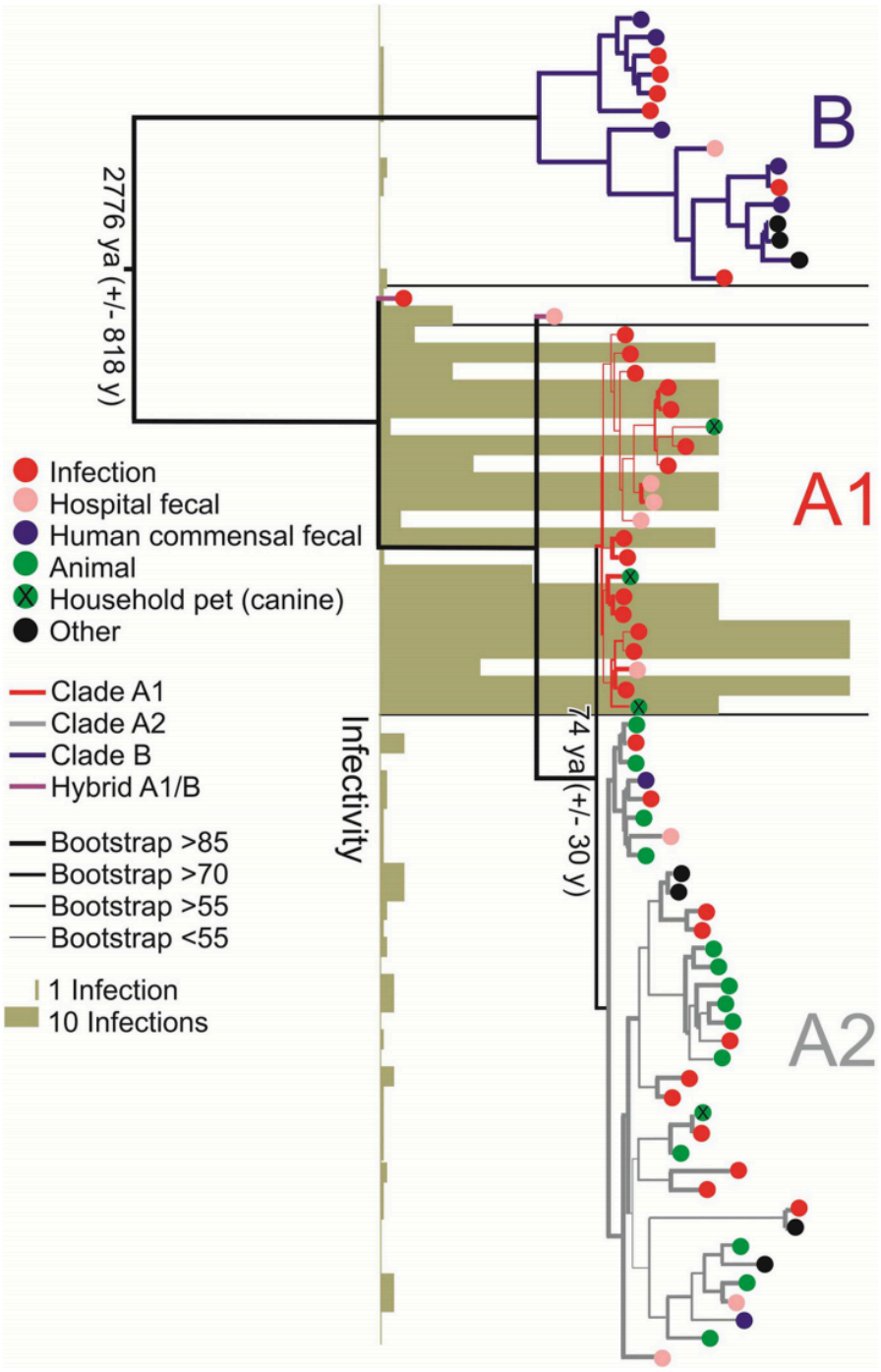

Figure 1.6: Phylogenetic tree based on SNP difference in 1,344 single-copy core genes across 73 *E. faecium* genomes. Dates are estimated by a BEAST analysis and are indicated by (ya, years ago). The infectivity score is the number of strains of a particular sequence type that have been isolated from an infection, in the MLST database. Clades are coloured: clade B in purple, clade A1 in red, and clade A2 in grey. Figure reproduced from (Lebreton *et al.*, 2013).

1.6.1 Enterococcal diseases

Enterococcus species can cause a number of different diseases in immune compromised patients (Figure 1.7). Urinary tract infection (UTI) is the most common disease caused by enterococci and accounts for 5% of all uncomplicated UTIs and 11% of complicated UTIs (Flores-Mireles *et al.*, 2015). UTIs, particularly within the upper urinary tract can lead to severe disease as they have been shown to contribute to the development of bacteraemia (Giannakopoulos *et al.*, 2019). Enterococci are also often isolated from cases of polymicrobial intra-abdominal infections. As with UTIs these infections can contribute to the development of bacteraemia (Montravers *et al.*, 1997). Enterococci are a leading cause of bacteraemia, accounting for around 8% of blood stream infections annually (Diekema *et al.*, 2019). These enterococcal blood stream infections have high mortality rates which vary between 50.5% in cases of vancomycin-resistant *E. faecium* and 24.4% in cases of vancomycin-sensitive *E. faecalis* (Kramer *et al.*, 2018). Bacteraemia, urinary-tract infections and the gastrointestinal tract can act as sources for the development of endocarditis. Enterococci are the third leading cause of bacterial endocarditis with a mortality rate of 23.8% (Holland *et al.*, 2016; Pericàs *et al.*, 2020). Around 10% of skin and soft tissue infections in North America are caused by *Enterococcus* species (Moet *et al.*, 2007). In rare occasions, *Enterococcus* species can also cause meningitis following traumatic brain injury or surgery (Zhang, Jiang and Zhou, 2021).

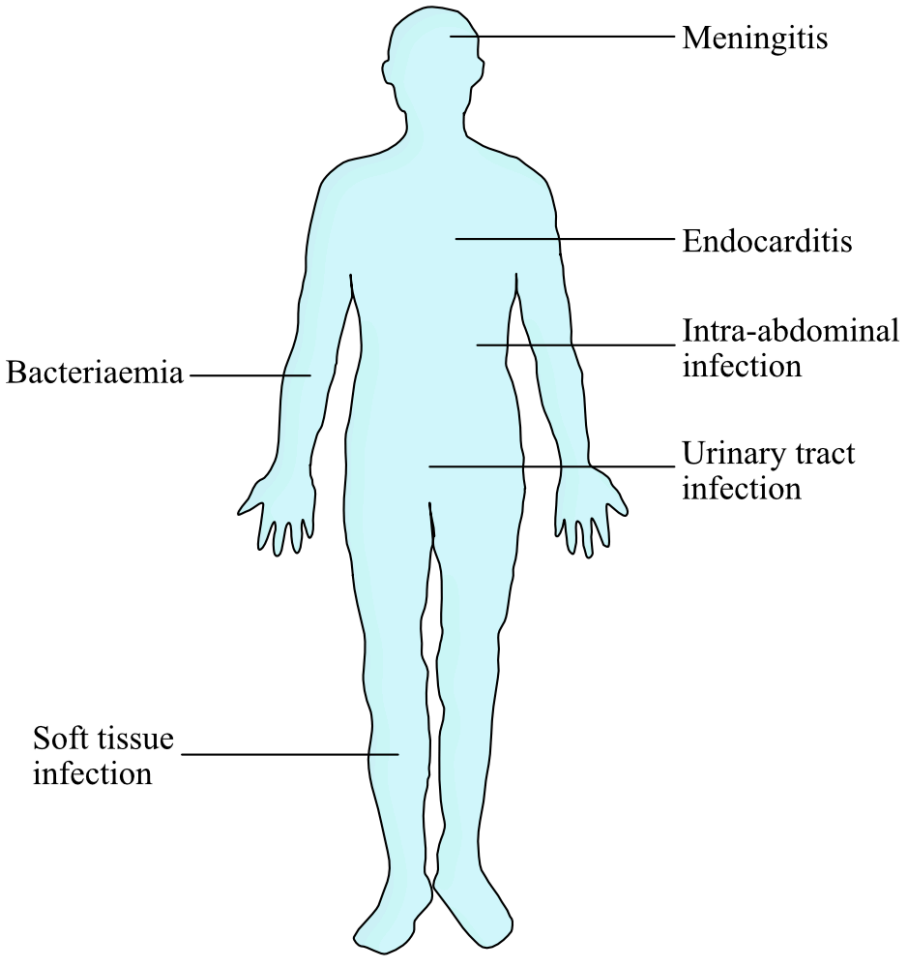

Figure 1.7: Diseases caused by *Enterococcus faecium* in human hosts.

1.6.2 Enterococcal Pathogenesis

Historically *E. faecalis* has been the predominant human pathogen in the *Enterococcus* genus which has meant that virulence has been primarily studied in this organism. However, over recent years many homologues of the *E. faecalis* virulence factors have been identified in *E. faecium.* The virulence factors of *E. faecium* fall into two broad categories, those that are extracellularly secreted and those that are not. The extracellularly secreted proteins can have markedly different functions, GelE is a gelatinase which breaks down host derived peptides such as gelatine and collagen (Rathnayake, Hargreaves and Huygens, 2012; Landete *et al.*, 2018) while cytolysin (Cyl) is a secreted protein that forms a pore in the surface of both prokaryotic and eukaryotic cells causing cell death (Van Tyne, Martin and Gilmore, 2013). Another secreted protein SagA, allows *E. faecium* cells to bind to the extracellular matrix and has also been shown in some clinical isolates to aid in biofilm formation (Teng *et al.*, 2003; Paganelli *et al.*, 2015; Pedicord *et al.*, 2016; Rangan *et al.*, 2016; Griffin *et al.*, 2021). As well as the secreted virulence factors, *E. faecium* has a large arsenal of surface associated molecules which contribute to virulence. The production of capsular polysaccharides has been associated with a number of functions including masking cells from detection by the host immune system and biofilm formation (Hancock and Gilmore, 2002; Ramos, Sansone and Morales, 2021). *E. faecium* possesses both lipoteichoic acids and cell wall teichoic acids which are also involved in biofilm formation and evasion of the host immune system. *E. faecium* has many LPxTG anchored cell surface proteins including Esp, Acm, Scm, PrpA, EcbA and SgrA, which are involved in binding to the host extracellular matrix and biofilm formation (Hendrickx *et al.*, 2009; Guzmán Prieto *et al.*, 2015; Freitas *et al.*, 2018). Two types of pili are present in *E. faecium*, PilA and PilB, as with the LPxTG anchored proteins the pili aid in cell attachment and biofilm formation

(Hendrickx *et al.*, 2008). Phosphotransferase systems have also been linked to virulence in enterococci, while their mechanism of action is unclear, the phosphotransferase BepA has been associated with *in vivo* biofilm formation (Paganelli *et al.*, 2016).

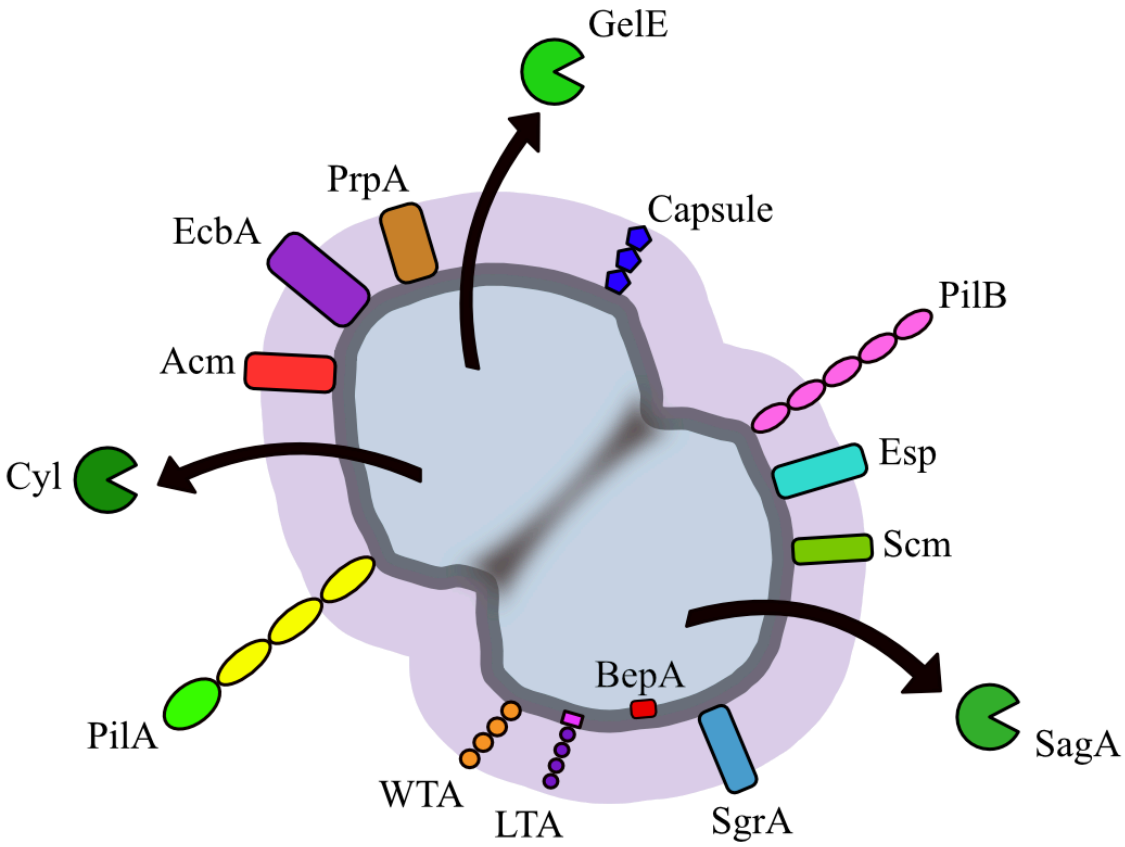

Figure 1.8: *Enterococcus faecium* virulence factors including both cell-associated and extracellular secreted factors.

1.6.3 Antibiotic resistance in *Enterococcus faecium*

Around 45,000 deaths per year have been directly attributed to antibiotic-resistant *Enterococcus faecium* and a further 190,000 deaths have been associated with antibiotic-resistant *E. faecium* (Murray *et al.*, 2022). *E. faecium* is resistant to a number of antibiotic drugs through both intrinsic and acquired mechanisms (Hollenbeck and Rice, 2012). Strains belonging to the hospital associated clade A1 are often resistant to the beta-lactam antibiotic ampicillin. These resistant isolates encode a D, D-transpeptidase called PBP5 which has a reduced binding affinity for beta-lactam antibiotics (Gagetti *et al.*, 2019).

As well as being intrinsically resistant to beta-lactam drugs, *E. faecium* is also intrinsically resistant to aminoglycosides. Low-level resistance is generally attributed to the integrity of the enterococcal cell envelope and/or the *aac(6')-Ii* gene which is found in the majority of *E. faecium* strains (Costa *et al.*, 1993; Adamecz *et al.*, 2021). For high-level resistance to aminoglycosides, *E. faecium* strains have horizontally acquired genes. Different aminoglycoside resistance genes provide resistance to different aminoglycoside drugs, with *ant(6')-Ia* conferring resistance to streptomycin while *aac(6')-aph(2'')* confers resistance to kanamycin (Adamecz *et al.*, 2021). A single aminoglycoside resistance gene can also confer resistance to a number of drugs, such is the case for *ant(4'')-Ia* which confers resistance to amikacin, kanamycin and tobramycin (Chow, 2000). These aminoglycoside resistance genes can often be found on large conjugative plasmids demonstrating their ability to transfer between strains (Tanimoto and Ike, 2008). Clinical *E. faecium* strains often have three or more aminoglycoside resistance genes making them pan-resistant (Chow, 2000). Previously, vancomycin was the drug of choice for treating beta-lactam- and aminoglycoside-resistant *E. faecium* strains. However, vancomycin resistance emerged in the late 1980s and has increased over the last 30 years (Rubinstein and Keynan, 2014). It is thought that the use of

another glycopeptide drug, avoparcin, as a growth promoter in agriculture may have driven the emergence of vancomycin resistance (Bager *et al.*, 1997). Unlike other antibiotic resistance genes which require only one gene to confer resistance, vancomycin requires an array of genes (Figure 1.9). These genes include *vanR* and *vanS* that code for a two-component regulatory systems which senses the presence of vancomycin. Upon binding of vancomycin, VanS is dimerised and autophosphorylated which in turn leads to the phosphorylation of VanR. The phosphorylated VanR can then bind to its own promoter and the promoter of the *vanHAX* operon where it works as an activator (Stogios and Savchenko, 2020). The first gene in this operon *vanH* is a D-hydroxyacid dehydrogenase which catalyses the formation of D-lactate from pyruvate creating a pool of D-lactate for the subsequent steps in the pathway (Matelska *et al.*, 2018). The next gene in the operon is *vanA,* which functions as a D-alanyl-D-lactate ligase that catalyses the formation of D-alanyl-D-lactate. This substrate is then used by MurF in the peptidoglycan synthesis pathway (Roper *et al.*, 2000). This D-alanyl-D-lactate ligase gene is variable with nine variants currently known (*vanA, vanB, vanC, vanD, vanE, vanG, vanL, vanM* and *vanN),* with *vanA* and *vanB* being the most clinically relevant (Sattari-Maraji *et al.*, 2019). The final gene that is essential for vancomycin resistance is *vanX*, which is a D-alanyl-D-alanine peptidase that selectively cleaves D-Ala-D-Ala dipeptides into single D-alanine subunits. The cleavage of the dipeptide reduces the amount available for incorporation into the cell wall thus increasing the amount of D-alanyl-D-lactate which is incorporated instead (Bussiere *et al.*, 1998). As well as the *vanRSHAX* genes which are essential for vancomycin resistance, there are also two accessory genes called *vanY* and *vanZ* which are not required for resistance to vancomycin. VanY is a D-D-carboxypeptidase that targets lipid II containing the terminal D-alanyl-D-alanine motif, in a similar manner to VanX, VanY limits the amount of the

native cell wall precursor that is available for incorporation into the cell wall (Stogios and Savchenko, 2020). The final protein VanZ is not required for vancomycin resistance but is required for resistance to the glycopeptide antibiotic teicoplanin through a mechanism that remains to be elucidated (Vimberg *et al.*, 2020).

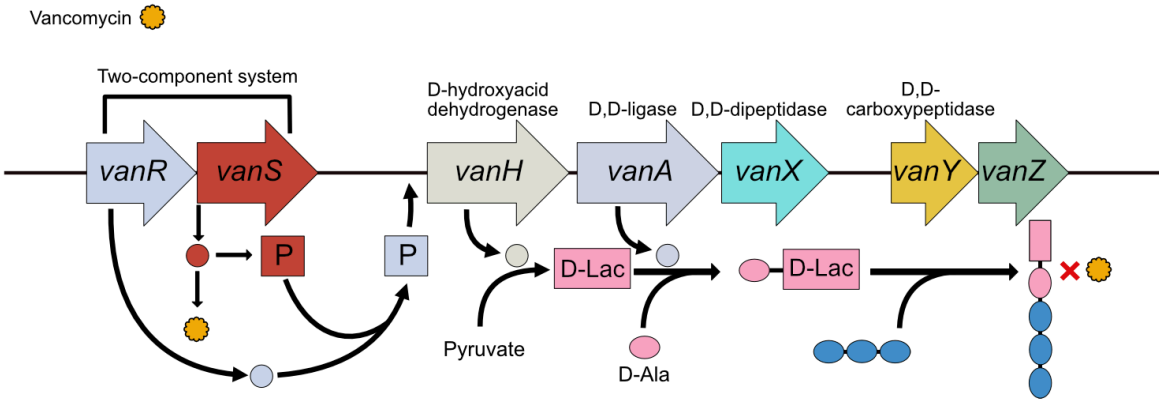

Figure 1.9: Vancomycin resistance mechanism. Vancomycin is sensed by VanS which leads to the activation of the *vanHAX* genes. The VanHAX proteins replace the terminal D-Ala D-Ala motif of the peptidoglycan precursors which D-Ala D-Lac which has a reduced biding affinity for vancomycin thus conferring resistance.

As well as vancomycin resistant and vancomycin susceptible *E. faecium* there are also strains that are vancomycin variable. These strains are phenotypically susceptible to vancomycin but contain the genes required for phenotypic resistance (Downing *et al.*, 2015). Vancomycin variable *E. faecium* (VVE) strains often contain insertions and deletions in the promoter regions of the *vanR* and *vanH* genes (Wagner *et al.*, 2021). Although VVE strains are phenotypically susceptible to vancomycin, they can rapidly revert back to a high-level resistant phenotype when exposed to a low concentration of vancomycin (Thaker *et al.*, 2015). The mechanisms behind the reversion to a resistant phenotype often involve the excision of IS elements from the resistance region and formation of constitutive promoters (Sivertsen *et al.*, 2016).

The widespread dissemination of vancomycin-resistant *E. faecium* has necessitated the use of more recently developed antibiotics to treat infections caused by these strains. Linezolid

is an oxazolidinone antibiotic that was first licenced in the early 2000s and is used to treat drug resistant Gram-positive pathogens, particularly vancomycin resistant *Enterococcus faecium* (Hashemian, Farhadi and Ganjparvar, 2018). Two different mechanisms for resistance have emerged in *E. faecium*, the first is a *G2576T* point mutation in the 23S rRNA gene leads to a target site modification in the ribosome and reduces binding of linezolid (Auckland *et al.*, 2002). Linezolid resistance genes can also be horizontally acquired from other *E. faecium* strains. One such gene *optrA,* encodes an ABC transporter that actively removes linezolid from the cell (Wang *et al.*, 2015). Another horizontally acquired gene Cfr, is an rRNA methyltransferase which methylates 23 rRNA gene at position A2503 blocking the linezolid ribosomal binding site (Long *et al.*, 2006). Additionally, the *poxtA* gene encodes an ATP-binding cassette protein that provides ribosomal protection from linezolid, likely by displacement (Antonelli *et al.*, 2018).

1.6.4 Treatment of *E. faecium* infections

Owing to their extensive repertoire of antibiotic resistance genes, the treatment options for infections caused by *Enterococcus faecium* are limited to a handful of antibiotic drugs (Zhou *et al.*, 2020). For bacteraemia caused by vancomycin-susceptible *E. faecium*, intravenous infusion of the glycopeptide antibiotic vancomycin is the first-choice treatment. Similarly, in cases of endocarditis caused by vancomycin-susceptible *E. faecium*, vancomycin infusion is the recommended treatment, but this can also be supplemented with gentamicin (Turco *et al.*, 2021). Bacteriaemia and endocarditis infections caused by vancomycin-resistant *E. faecium* cannot be treated with vancomycin and are instead treated with linezolid or daptomycin. As with vancomycin-susceptible cases of endocarditis, treatment with linezolid or daptomycin can also be augmented with gentamicin (O'Driscoll and Crank, 2015). The glycopeptide antibiotic teicoplanin can also be used for the treatment of vancomycin-susceptible and *vanB*-type resistant *E. faecium* strains as they do not carry the *vanZ* gene that confers resistance to teicoplanin. However, teicoplanin cannot be used to treat *vanA*-type *E. faecium* infections as they often carry this gene (Escolà-Vergé *et al.*, 2019). Asymptomatic urinary tract infections are not regularly treated with antibiotics. However, fosfomycin is efficacious in the treatment of UTIs caused by vancomycin-resistant *E. faecium* strains (Abbott *et al.*, 2020). As *E. faecium* isolates are often multidrug resistant it is important to practice good antimicrobial stewardship when treating these infections. Treatment of *E. faecium* should be informed by susceptibility testing of isolates to ensure an appropriate treatment is administered.

1.6.5 Enterococcal Plasmids

Enterococcus species have plasmids belonging to four main groups namely RepA_N, Inc18, RCR and Rep_3 (Arredondo-Alonso *et al.*, 2020). RepA_N plasmids are found in many low-

GC Gram-positive organisms but have a very narrow host range (Weaver *et al.*, 2009). The well characterised pheromone-responsive plasmids (PRP) pAD1 and pCF10 of *E. faecalis* belong to this group. The transfer of these plasmids is mediated by the production of a peptide hormone in plasmid-free strains which stimulates conjugation in the plasmid-containing strain. Plasmid transfer rates can be exceptionally high in PRPs reaching rates of 1 transconjugant per 10 donor cells (Sterling *et al.*, 2020). PRPs can transfer antibiotic resistance determinants, with the prototypical PRP pCF10 carrying the tetracycline resistance gene *tetM* and pBRG1 carrying the *vanA* vancomycin resistance operon (Magi *et al.*, 2003; Dunny, 2007). Other notable members of the RepA_N plasmids include the pRUM plasmids which are non-conjugative but often carry vancomycin resistance genes in both *E. faecalis* and *E. faecium* and pHTβ which also carries vancomycin resistance genes and is highly conjugative (Tomita and Ike, 2008; Hegstad *et al.*, 2010). Inc18 plasmids are broad host-range plasmids that often carry antibiotic resistance genes in both *E. faecalis* and *E. faecium*. The prototypical Inc18 plasmid pIP501 can be transferred to a wide variety of species including streptococci, staphylococci and lactococci amongst others (Kohler, Vaishampayan and Grohmann, 2018). It has also been shown that vancomycin resistance in *S. aureus* can be caused by the transfer of Inc18 plasmids carrying vancomycin resistance genes from enterococci (Rossi *et al.*, 2014). Rolling circle replication (RCR) plasmids are found in many Gram-positive organisms but are not common in enterococci. However, some RCR plasmids carry antibiotic resistance genes such as pMV158 which carries the tetracycline resistance gene *tetL* (Ruiz-Masó *et al.*, 2017). The Rep_3 plasmids are small plasmids that often carry bacteriocin genes and can also carry antibiotic resistance genes such as the tetracycline resistance gene *tetL* in pAMα1 and *aadE* aminoglycoside resistance gene in pEF418 (Clewell *et al.*, 2002; Wardal *et al.*, 2017).

1.6.6 Incidence of enterococci in UK hospitals

Although enterococcal species are ESKAPE pathogens, mandatory reporting is not required in UK hospitals. However, a voluntary surveillance program was in place between 2014 and 2018 (Public Health England, 2019). The total bacteraemia cases caused by *Enterococcus* spp. was 6,126 in 2014 and has increased year-on-year to 8,136 cases in 2018 (Figure 1.10). The surveillance program recorded bacteraemia caused by 14 different species of *Enterococcus* but 12 of the 14 species accounted for only around 5% of the total cases. *E. faecalis* and *E. faecium* caused the majority of bacteraemia cases, *E. faecalis* caused 2,549 cases and *E. faecium* caused 2,295 cases in 2014. Cases caused by *E. faecalis* and *E. faecium* have also increased year-on-year but have made up the same proportion of cases each year approximately 43% and 38% respectively.

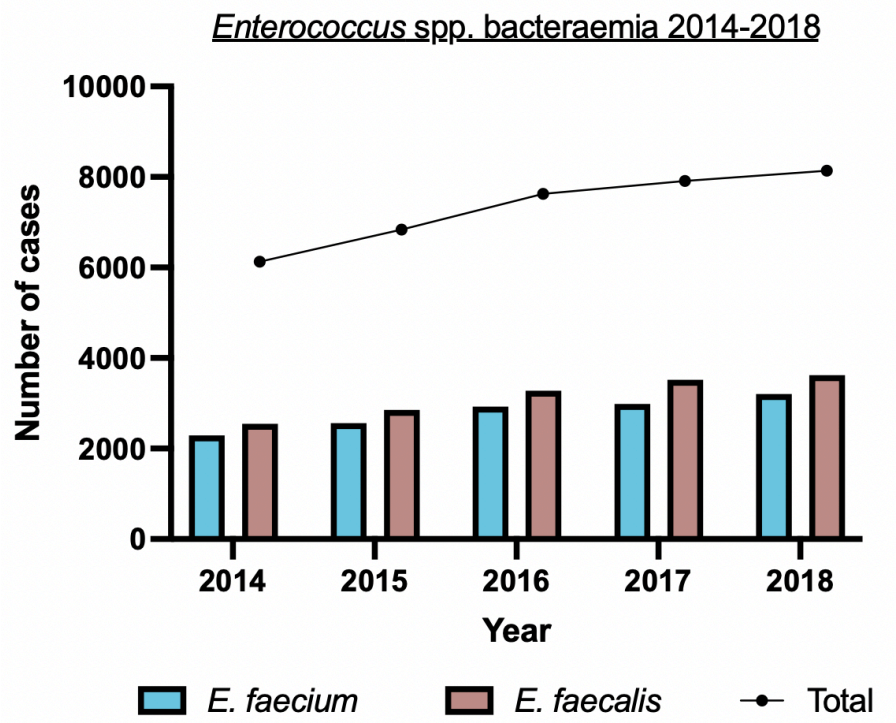

Figure 1.10: Prevalence of *Enterococcus* spp. bacteraemia in England, Wales and Northern Ireland from 2014 to 2018. Created from data from Health Protection Report, Volume 13, Number 33, 23/09/19, "Laboratory surveillance of Enterococcus spp. Bacteraemia in England, Wales and Northern Ireland: 2018". Total cases (black dotted line), *E. faecium* (blue) and *E. faecalis* (brown).

1.6.7 Incidence of antibiotic resistant *Enterococcus* spp. in UK hospitals

The antibiotic resistance profile of *E. faecalis* and *E. faecium* causing bacteraemia in England and Northern Ireland was measured between 2015 and 2018 (Public Health England, 2019). *E. faecium* had extremely high rates of ampicillin resistance with 91% of isolates being resistant from 2015 to 2018 while resistance rates remained at around 2% for *E. faecalis* across the same time period (Figure 1.11). The rate of vancomycin resistance was also considerably higher for *E. faecium* with a yearly rate of around 23%. In comparison, only around 1% of *E. faecalis* isolates were found to be vancomycin-resistant every year. Approximately 25% of *E. faecium* strains causing bacteraemia were found to be resistant to teicoplanin while only 2% of *E. faecalis* were.

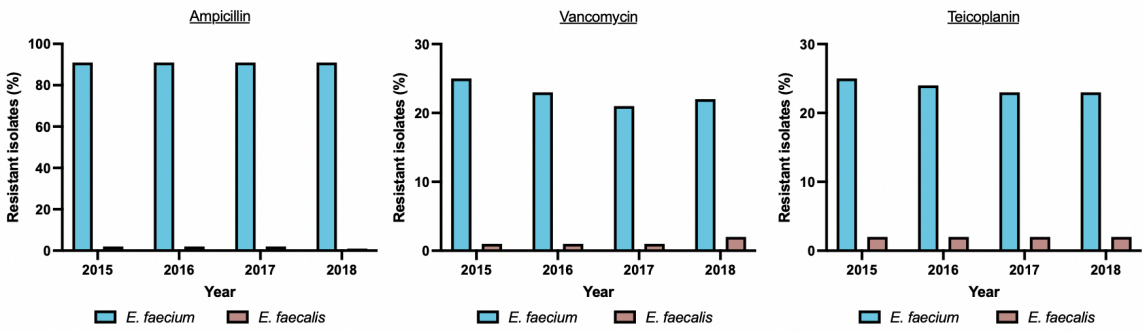

Figure 1.11: Percentage of *Enterococcus faecium* and *Enterococcus faecalis* species causing bacteraemia in England and Wales between 2015 and 2018 that are antibiotic resistant. *E. faecium* (blue) and *E. faecalis* (brown).

1.6.8 Incidence of VRE in Europe

The European Centre for Disease Control has been monitoring rates of VRE in the European Union since 2001 (TESSy & ECDC, 2021). The percentage of *E. faecium* which are resistant to vancomycin has risen across Europe over the last 20 years from an average of 4% in 2001 to an average of 18% in 2020 (Figure 1.12). However, VRE rates have varied dramatically across Europe. The Scandinavian countries of Finland, Norway and Sweden have maintained low rates of VRE over this period with around 1-2% isolates being resistant to vancomycin. On the other hand, the Eastern European countries including Latvia, Lithuania, Slovakia and Romania have reported significant increases in VRE rate with 30-60% of isolates being resistant to vancomycin in 2020.

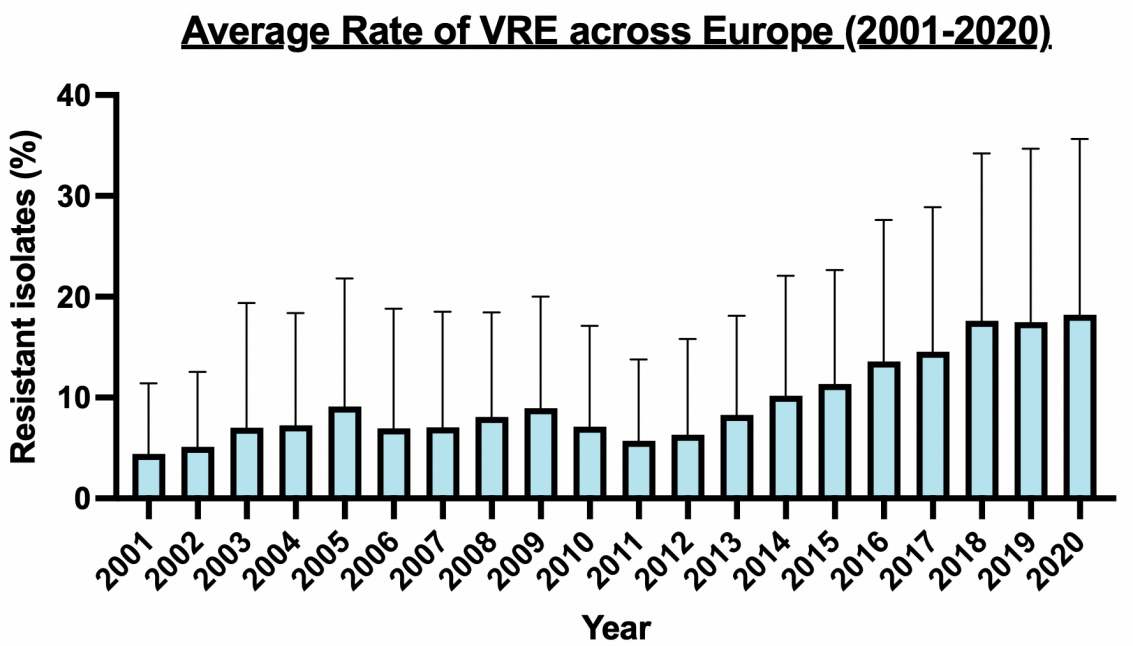

Figure 1.12: Mean percentage of *Enterococcus faecium* isolates which are resistant to vancomycin across Europe from 2001 to 2020. Error = Standard deviation. Generated from data gathered in (TESSy & ECDC, 2021)

1.7 Aims

With the research presented in this thesis, I aimed:

- To identify whether antibiotic use in aquaculture is driving antibiotic resistance in Bangladesh.
- To investigate the diversity of *Enterococcus faecium* in a Birmingham hospital and to determine whether a VRE outbreak was clonal in nature.
- To use transcriptomics and functional genomics to identify genes that are differentially expressed and genes that contribute to increased fitness in the presence of vancomycin.

Chapter 2
Materials and Methods

2.1 General methods

2.1.1 Buffers and solutions

2.1.1.1 General Buffers

Phosphate Buffered Saline (PBS) (AppliChem): 1 tablet per 100 ml of deionised water.

5x Tris Borate EDTA (TBE) (National Diagnostics)

2.1.1.2 *Enterococcus* competent cells protocol

Medium A (50 ml): 2.5 ml of 20% (w/v) glycine (Alfa Aesar), 5 ml of 1 M sucrose (VWR) & 42.5 ml of BHI broth (VWR)

Medium B (50 ml): 25 ml of 2x concentrated BHI (VWR) & 25 ml of 1 M sucrose (VWR)

Wash buffer (50 ml): 25 ml of 1 M sucrose (VWR) & 25 ml of 20% (v/v) glycerol (Fisher Bioreagents)

2.1.1.3 *E. coli* competent cells protocol

Transformation Storage Solution (TSS) buffer pH 6.5 (100 ml): 2 g LB powder (Sigma-Aldrich), 10 g PEG 3350 (Alfa Aesar), 5 ml DMSO (Alfa Aesar) & 0.5 g $MgCl_2$ (Merck Life Science Ltd).

2.1.2 Growth Media and Antibiotics

All *E. coli* strains were grown in Lysogeny Broth (LB) and on LB agar (Table 2.1). *E. faecium* strains were grown in Brain Heart Infusion (BHI) broth and on BHI agar plates. Iso-Sensistest broth was used in all minimum inhibitory concentration assays (see 2.1.20 Broth microdilution MIC).

Table 2.1: Growth media used in this thesis.

Media	Components	Supplier
LB (400 ml)	LB powder (8 g)	Sigma-Aldrich
LB agar (400 ml)	LB agar powder (14 g)	Sigma-Aldrich
BHI broth (400 ml)	BHI broth powder (14.8 g)	VWR
2x BHI broth (400 ml)	BHI broth powder (29.6)	VWR
BHI agar (400 ml)	BHI agar powder (20.8 g)	Millipore
Iso-Sensitest broth (400 ml)	Iso-Sensitest broth powder (9.36 g)	Oxoid

LB = Lysogeny Broth, BHI = Brain Heart Infusion
All media were prepared according to the manufacturer's instructions.

Antibiotic concentrations varied in each experiment and will be described in each relevant section. A list of all antibiotics used in this thesis and the solvent used is provided below (Table 2.2).

Table 2.2: Antibiotics used in this thesis.

Antibiotic	Solvent	Diluent	Supplier
Spectinomycin dihydrochloride	ddH_2O	ddH_2O	Alfa Aesar
Erythromycin	EtOH (100%)	ddH_2O	Serva
Vancomycin hydrochloride	ddH_2O	ddH_2O	Alfa Aesar
Rifampicin	DMSO	ddH_2O	Fisher Chemical
Fusidic acid sodium salt	ddH_2O	ddH_2O	Arcos Organics
Ampicillin sodium salt	ddH_2O	ddH_2O	Fisher Bioreagents
Ciprofloxacin hydrochloride	ddH_2O	ddH_2O	MP Biomedicals
Gentamicin sulfate	ddH_2O	ddH_2O	Alfa Aesar

ddH_2O = Double-distilled water, EtOH = Ethanol, DMSO = Dimethyl sulfoxide

2.1.3 Plasmid sequencing

Plasmids were sanger sequenced using the Eurofins TubeSeq service. Prior to sequencing plasmids were quantified using the Qubit broad range kit (Invitrogen) and diluted to 50 – 100 μg/ml in 15 μl. A full list of primers used for Sanger sequencing can be found in Tables 2.7 and 2.8. FASTA files were visualised and analysed using SnapGene software (from GSL Biotech; available at snapgene.com).

2.1.4 Polymerase Chain Reaction

Polymerase Chain Reaction (PCR) was used to confirm successful electroporation of plasmids into *E. faecium* and transformation of plasmids into *E. coli*. A standard PCR reaction consisted of 25 μl of DreamTaq 2x Mastermix (Thermo Fisher Scientific), 0.25 μl (0.5 μM) of both forward and reverse primers, a bacterial colony and 24.5 μl of PCR-grade water (VWR). The reaction was incubated according to the following program which was carried out in a Mastercycler Pro Thermal Cycler (Eppendorf):

Table 2.3: Colony PCR program

Step	Time	Temperature	Number of cycles
Initial denaturation	3 minutes (mins)	95°C	1
Denaturation	30 seconds (s)	95°C	25-30
Annealing	30 s	Tm of primers -5°C	
Extension	1 min per kb	72°C	
Final Extension	10 mins	72°C	1
Hold	Indefinite	10°C	1

Phusion polymerase (New England Biolabs (NEB)) was used for the amplification of products used in cloning reactions. PCR reactions using Phusion polymerase were carried out as follows: 10 μl of 5x Phusion HF buffer (NEB), 1 μl of 10 mM dNTPs (NEB), 2.5 μl of 10 μM forward and reverse primers, 10 ng of template DNA and 0.5 μl of Phusion DNA polymerase were prepared on ice and mixed well. The reaction was incubated according to

the following program which was carried out in a Mastercycler Pro Thermal Cycler (Eppendorf):

Table 2.4: Cloning PCR program

Step	Time	Temperature	Number of cycles
Initial denaturation	30 s	98°C	1
Denaturation	10 s	98°C	25-30
Annealing	25 s	Tm of primers +3°C	
Extension	20 s per kb	72°C	
Final Extension	10 mins	72°C	1
Hold	Indefinite	10°C	1

2.1.5 Agarose Gel Electrophoresis

DNA fragments between 200 and 20,000 bp were visualised on 1% (w/v) agarose gels. Typically, 500 mg of agarose (Sigma) was dissolved in 50 ml of 1% (v/v) TBE (National diagnostics) (0.089 M Tris base, 0.089 M boric acid (pH 8.3) and 2 mM Na_2EDTA). 5 µl of SYBR Safe® (EDVOTEK) was added to the TBE agarose mixture prior to dissolving. The agarose was dissolved by heating to 100°C for 1 min. The gel mixture was cooled for 10 mins at room temperature before being poured into the gel tray and set for 15 mins. When DreamTaq 2x Mastermix had been used, 10 µl of PCR product was loaded directly onto the gel. However, in the case of a PCR which used Phusion polymerase, 2 µl of DNA loading dye was mixed with 10 µl of the sample and again 10 µl was loaded onto the gel. Gels were run at 125 V for 45 mins and imaged on an Amersham Imager 680 gel imager (General Electric Healthcare) using blue light epifluorescence and an automatic exposure time.

2.1.6 Competent cells and bacterial transformation

2.1.6.1 *E. coli* competent cells

All *E. coli* cells used for transformations were made competent by the one step polyethylene glycol method (Chung, Niemela and Miller, 1989). A single *E. coli* colony was inoculated into 5 ml of LB broth and grown for 16 hours at 37ºC with shaking (200 RPM). 1 ml of the bacterial culture was transferred to 99 ml of LB broth and grown at 37ºC with shaking (200 RPM) until an OD_{600} = 0.4. The culture was transferred to two 50 ml Falcon tubes (Corning) and centrifuged at 1000 x g for 10 mins at 4ºC. The supernatant was discarded, and the cell pellet was resuspended in 5 ml of ice cold TSS buffer (see 2.1.1.3 *E. coli* competent cells protocol). The cells were then aliquoted into 1.5 ml Eppendorf tubes and immediately stored at -80ºC. Competent *E. coli* cells were transformed by incubating 100 µl of cells with 10 µl of ligation mixture or 500 ng of plasmid DNA on ice for 30 mins. The cells were further incubated at 42ºC for 45 s and then on ice for 3 mins. 900 µl of LB was added to the cells and they were recovered at 37ºC for 1 hour with shaking (200 RPM). 100 µl of the transformed cells was spread on a LB agar plate containing an appropriate antibiotic for selection and incubated at 37ºC for 16 hours.

2.1.6.2 *E. faecium* competent cells

E. faecium cells were made electrocompetent using the protocol described in (X Zhang *et al.*, 2012). *E. faecium* cells were transformed as in (X Zhang *et al.*, 2012), electroporation was performed in a Eporater (Eppendorf) at 2.5 kV.

2.1.7 Plasmid DNA isolation

Plasmid DNA < 10 kb was isolated using the GeneJET Plasmid Miniprep Kit (Thermo Scientific™). The strain of interest was grown in the relevant media for 16 hours at 37ºC

with shaking (200 RPM). Cells were collected by centrifugation at 4,000 RPM for 10 mins. The pelleted cells were resuspended in 250 µl of Resuspension solution by pipetting up and down. For the extraction of Enterococcal plasmid DNA, lysozyme was added to a concentration of 10 mg/ml and incubated at 37ºC for 1 hour. To lyse the cells 250 µl of Lysis solution was added, and the tube was inverted 10 times to mix. 350 µl of Neutralisation solution was then added, and the tube was inverted 10 times to mix. The solution was centrifuged at 13,000 x g for 5 mins before 755 µl of the supernatant was removed and applied to the Thermo Scientific GeneJET Spin Column. The spin column was centrifuged at 13,000 x g for 1 min and the flow through was discarded. 500 µl of wash solution was then added to the column and centrifuged for a further 1 min at 13,000 x g. The flow-through was discarded and the previous step was repeated once more. The empty column was centrifuged at 13,000 x g for 2 mins to remove any residual wash buffer. The column was then placed in a fresh 1.5 ml microcentrifuge tube and 32 µl of nuclease-free water was added directly to the filter. After incubation for 2 mins at room temperature the column was centrifuged at 13,000 x g for 2 mins, the column was then discarded and the flow-through containing the purified plasmid DNA was retained.

2.1.8 PCR purification and gel extraction

Both PCR purification and gel extraction used the GeneJet PCR Purification Kit (Thermo Scientific™). For gel extraction the agarose gel piece was weighed and an equal amount (µl:mg) of binding buffer was added, the reaction was then incubated at 65ºC for 10 mins to melt the agarose. For PCR purification, an equal volume (µl:µl) of binding buffer was added to the PCR reaction. Up to 800 µl of the gel solution / PCR reaction solution was then added to a GeneJet purification column and centrifuged at 13,000 x g for 1 min. The flow through

was discarded and 700 µl of Wash buffer was added to the column, the column was centrifuged for a further 1 min at 13,000 x g and again the flow through was discarded. The column was placed in a fresh 1.5 ml microcentrifuge tube and 32 µl of nuclease-free water was applied directly to the filter. The column was incubated at room temperature for 2 mins before being centrifuged at 13,000 x g for 2 mins. The column was discarded, and the purified DNA was retained.

2.1.9 Restriction digest

All restriction enzymes were purchased from New England Biolabs (USA). High Fidelity (HF) enzymes were purchased where possible to make double-digest reactions more straightforward. Restriction enzyme reactions were set up according to the manufacturer's instructions. A typical restriction digest reaction consisted of 1 µl (20 U) of the relevant restriction enzyme, 1 µg of DNA, 5 µl of the appropriate restriction enzyme buffer and the reaction was made up to 50 µl with nuclease-free water. The restriction enzyme buffer used with HF enzymes was always CutSmart buffer however for non-HF enzymes this could be NEB Buffer 1.1, 2.1 or 3.1. The reactions were incubated for greater than 1 hour independent of whether they were HF or not. The enzymes were inactivated according to the manufacturer's instructions for the particular enzyme. For double-digest reactions where the enzymes could not use the same restriction enzyme buffer, a single-digest reaction was carried out followed by purification (see 2.1.8 PCR purification and gel extraction) and then the second digestion was performed.

2.1.10 Ligation reactions

T4 DNA ligase (New England Biolabs) was used in all ligation reactions. A typical ligation reaction consisted of 2 µl of T4 DNA ligase buffer (New England Biolabs, USA), vector and insert DNA at a molar ratio of 1:3, nuclease-free water up to 20 µl and 1 µl of T4 DNA ligase. The reaction was incubated at 16ºC for 16 hours before inactivation at 65ºC for 10 mins. Ligation reactions which were unsuccessful had the molar ratio of vector to insert DNA changed from 1:3 to 1:5 and the previous steps were repeated.

2.1.11 Oligonucleotide annealing

To create short inserts that could be ligated into plasmids and adapters for sequencing, oligonucleotides were designed which could be annealed together. 5 µl of the forward and reverse primer (1 nmol/µl) were mixed with 0.5 µl of 100x Tris-EDTA, 0.5 µl of 5 M NaCl and 39 µl of nuclease-free water (New England Biolabs). The reaction was incubated for 95ºC for 10 mins, the metal block was then removed from the heat block and allowed to cool to room temperature. The annealed oligonucleotides were then stored at -20ºC for 1 year.

2.1.12 CRISPR protocol

A CRISPR-Cas9 system was used to insert and delete genes from the genomes of *E. faecium* strains (de Maat *et al.*, 2019).

2.1.12.1 Electroporation of pVPL3004 into the strain of interest

Plasmid pVPL3004, containing the *cas9* gene and tracrRNA, was electroporated into the strain of interest and selected for on BHI agar with erythromycin 50 µg/ml at 30ºC for 24-48 hours (see 2.1.6.2 *E. faecium* competent cells), colonies were then screened for successful plasmid uptake by colony PCR (see 2.1.4 Polymerase Chain Reaction).

2.1.12.2 Cloning of CRISPR sequence into pVDM1001

A CRISPR sequence was designed by identifying a unique 30 bp stretch of DNA followed by the protospacer adjacent motif (PAM) sequence "NGG" in the gene of interest. An oligonucleotide was then ordered that contained the CRISPR sequence with an additional AAAC at the 5' end and a G at the 3' end. A second oligonucleotide was ordered with the reverse complement of the CRISPR sequence and an additional AAAAC at the 5' end. The oligonucleotides were annealed together to form a CRISPR insert (see 2.1.11 Oligonucleotide annealing). Plasmid pVDM1001 was digested with BsaI-HF (New England Biolabs) and was column purified (see 2.1.8 PCR purification and gel extraction). The annealed CRISPR insert possessed BsaI-compatible overhangs which allowed the linearised pVDM1001 plasmid and CRISPR insert to be directly ligated together (see 2.1.10 Ligation reactions). The ligation mixture was transformed into *E. coli* EC1000 and selected on LB agar with spectinomycin 100 μg/ml for 16 hours at 37^{o}C, colonies were then screened for successful insertion by colony PCR (see 2.1.4 Polymerase Chain Reaction).

2.1.12.3 Construction of gene insertion or deletion plasmids

DNA templates were created which could be used to insert or delete a gene from the genome. 500 bp stretches of sequence were taken from upstream and downstream of the gene being targeted. For the deletion template, a gBlock (Integrated DNA Technologies) was ordered which had the two flanking regions directly fused together. For the gene insertion template, the same flanking regions were taken but the gBlock was ordered with restriction sites between the two flanking regions, this template was termed the universal repair template (URT). The templates were then amplified and pVDM1001 containing the appropriate CRISPR sequence was digested with SmaI (New England Biolabs) (see 2.1.9 Restriction digest) and ligated together with the gBlock construct (see 2.1.10 Ligation reactions). The

ligation mixture was transformed into *E. coli* EC1000 and selected for on LB agar with spectinomycin 100 μg/ml for 16 hours at 37ºC (see 2.1.6 Competent cells and bacterial transformation), colonies were then screened for successful insertion by colony PCR (see 2.1.4 Polymerase Chain Reaction). For insertion plasmids, pVDM1001 containing the URT was digested with an appropriate restriction enzyme (see 2.1.9 Restriction digest) and ligated together with the gene of interest (from a gBlock or amplicon) (see 2.1.10 Ligation reactions). The ligation mixture was then transformed into *E. coli* EC1000 and selected for on LB agar with spectinomycin 100 μg/ml for 16 hours at 37ºC (see 2.1.6 Competent cells and bacterial transformation), colonies were then screened for successful insertion by colony PCR (see 2.1.4 Polymerase Chain Reaction). The insertion or knockout plasmid was then electroporated into the *E. faecium* strain of interest carrying pVPL3004 (see 2.1.6.2 *E. faecium* competent cells) and selected for on BHI agar with 200 μg/ml spectinomycin and 70 μg/ml erythromycin at 30ºC. Colonies were then screened for successful knockout or insertion by colony PCR (see 2.1.4 Polymerase Chain Reaction).

2.1.12.4 Curing of pVDM1001 and pVPL3004

Colonies which contained the successful insertion or deletion were inoculated into 5 ml of BHI with 200 μg/ml spectinomycin and 70 μg/ml erythromycin and incubated at 30ºC for 16 hours. 4 μl of the culture was then inoculated into 40 ml of BHI broth and incubated at 37ºC for 16 hours. The subculturing was repeated a further three times, after the third day the culture was diluted 10^{-6} in PBS, 100 μl spread on BHI agar and incubated at 37ºC for 16 hours. 50 colonies were then patched onto BHI agar, BHI agar containing 200 μg/ml spectinomycin and BHI agar containing 50 μg/ml erythromycin and incubated at 37ºC for

16 hours. Colonies which grew on BHI but not on the spectinomycin or erythromycin plates were cultured in 5 ml of BHI at 37ºC for 16 hours and were saved as glycerol stocks.

2.1.13 Conjugation protocol

The donor and recipient *E. faecium* strains were first grown for 16 hours in 5 ml of BHI broth at 37ºC with shaking (200 RPM). 1 ml of the bacterial culture was mixed with 19 ml of brain heart infusion (BHI) broth and cultured at 37ºC until it had reach OD_{600} = 1. 1ml aliquots of the donor and recipient strains were mixed together and cells were pelleted by centrifugation at 14,000 x g for 2 mins. The cells were washed by adding 1 ml of phosphate buffered saline (PBS) and resuspending the cells. The cells were re-pelleted by further centrifugation at 14,000 x g for 2 mins and the supernatant was discarded. The cell pellets were re-suspended in 30 µl of BHI broth and spotted on 0.45 µm Whatman nitrocellulose filters (GE Healthcare) which were placed on BHI agar plates. These plates were then incubated at 37ºC for 16 hours with the agar facing upwards. After incubation, the filters were removed from the plates and the cells were resuspended in 1 ml of BHI broth. 100 µl of the cell suspension was plated on BHI agar containing an antibiotic which selected for the donor and transconjugant bacteria, an antibiotic which selected for the recipient and transconjugant bacteria and multiple antibiotics which selected for only the transconjugant bacteria. These plates were incubated at 37ºC for 24 hours before the colonies were counted and the conjugation rate was calculated.

2.1.14 Genomic DNA extraction for short-read sequencing

Chromosomal DNA was extracted using the Wizard® Genomic DNA Purification Kit (Promega). The bacteria that were to have their genomic DNA extracted were grown for 16 hours in the relevant media at 37ºC with shaking (200 RPM). 1 ml of cells were pelleted by

centrifugation at 16,000 x g for 2 minutes. The supernatant was discarded, and the cells were resuspended in 480 µl of 50 mM EDTA (Invitrogen). 120 µl of 10 mM lysozyme (Sigma-Aldrich) was then added to the cells and they were incubated at 37ºC for 60 minutes. The cells were pelleted by centrifugation at 16,000 x g, the supernatant was removed and 600 µl of nuclei lysis solution was added and mixed by pipetting up and down. The cell suspension was then incubated at 80ºC for 5 mins before being cooled to room temperature. 3 µl of RNase solution was added to the cells and incubated at 37ºC for 60 mins, the solution was again cooled to room temperature. To precipitate the proteins in the sample, 200 µl of the protein precipitation solution was added, and the cells were vortexed for 20 s. The cells were incubated on ice for 5 mins and then centrifuged at 16,000 x g for 3 mins to pellet the precipitated protein. To precipitate the DNA, the supernatant was transferred to a tube containing 600 µl of room temperature isopropanol and mixed by inverting 10 times. The DNA was pelleted by centrifugation at 16,000 x g for 2 mins and the supernatant was discarded. 600 µl of 70% (v/v) ethanol was added to the pellet and mixed by inverting 10 times. The solution underwent another round of centrifugation at 16,000 x g for 2 minutes and the ethanol was aspirated so that no liquid was left on the interior of the tube. The pellet was left to air dry for 10 mins before being rehydrated in 50 µl of nuclease-free water at 65ºC for 1 hour. Extracted DNA was stored at -80ºC until used.

2.1.15 Genomic DNA extraction for long-read sequencing

A single colony of *Enterococcus faecium* was inoculated into 5 ml of BHI broth and grown at 37ºC for 16 hours with shaking (200 RPM). 1.5 – 2 ml of the culture was pelleted by centrifugation at 17,000 x g for 1 minute to collect 1 x 10^9 – 5 x 10^9 cells. The pellet was then resuspended in 300 µl of STET buffer (https://openwetware.org/wiki/Yu:STET)

containing 50 mg/ml lysozyme and incubated at 37ºC for 30 mins. DNA extraction was carried out using the Monarch® HMW DNA Extraction Kit for Tissue (NEB). 300 µl of HMW gDNA Tissue lysis buffer (NEB) was added to the sample and mixed by inverting 10 times, followed by 20 µl of proteinase K (NEB) and mixed again by inverting 20 times. The sample was incubated at 56ºC for 30 mins in a thermo-mixer HCM100-Pro with shaking at 1400 rpm. 10 µl of RNase A (NEB) was added to the sample, mixed by inverting 10 times and incubated for a further 10 mins at 56ºC with shaking at 1400 RPM. 300 µl of Protein separation solution was then added, and the sample was mixed by inverting for 1 min. The sample was then centrifuged at 16,000 x g for 10 mins to separate the protein from DNA. This split the sample into a large upper clear phase and a lower clear phase, the large upper clear phase (approximately 800 µl) was transferred to a fresh tube containing 550 µl of 100% isopropanol and 2 DNA capture beads. The tube was slowly inverted 30 times to precipitate the DNA and to wrap it around the DNA capture beads. The liquid was then removed from the tube without disturbing the DNA wrapped around the beads. 500 µl of gDNA Wash Buffer (NEB) was quickly added to the beads and the tube was inverted 5 times to mix. The wash step was repeated once more and then the beads were tipped into a bead strainer to remove any residual wash buffer by pulse-spinning in a centrifuge for ~ 1 s. The beads were then transferred to a new 2 ml tube and 100 µl of 10 mM Tris-HCl (pH 8.5) (Sigma-Aldrich) was added quickly, the sample was then incubated at 56ºC for 10 mins with shaking at 300 RPM. The liquid and beads were poured into a bead strainer placed in a fresh 1.5 ml and centrifuged at 12,000 x g for 30 s to remove the DNA from the beads. The DNA was pipetted up and down 10 times to mix, it was then left overnight at room temperature and pipetted a further 10 times to mix. To ensure the integrity of the high molecular weight DNA it was stored at 4ºC for a maximum of one week.

2.1.16 Measurement of DNA concentrations

DNA concentrations between 0.2 – 100 ng were measured using the Qubit™ dsDNA HS Assay Kit (Invitrogen) and concentrations between 2 – 1000 ng were measured using the Qubit™ dsDNA BR Assay Kit (Invitrogen). The procedure for both kits was the same however the specific reagents for each kit were different. A working solution was prepared by diluting Qubit® dsDNA BR/HS Reagent 1 in 200 in Qubit® dsDNA BR/HS Buffer. Standards 1 and 2 were made by mixing 190 µl of the working solution with 10 µl of Qubit® dsDNA BR/HS Standard #1 and 10 µl of Qubit® dsDNA BR/HS Standard #2 in a Qubit™ Assay Tube. To measure the DNA, 2 µl of DNA was mixed with 198 µl of working solution in a Qubit™ Assay Tube. The mixed solutions were vortexed briefly and incubated for 2 mins at room temperature. The assays were then read using a Qubit 4 fluorometer (Invitrogen).

2.1.17 Long-read sequencing

2.1.17.1 DNA repair and end-preparation

Long-read sequencing was performed using the MinION sequencer (Oxford Nanopore). To obtain the longest reads possible the Ligation Sequencing Kit (SQK-LSK109) (Oxford Nanopore) was used. Genomic DNA was extracted from the *E. faecium* strains using the Monarch® HMW DNA Extraction Kit for Tissue (New England Biolabs) (see 2.1.15 Genomic DNA extraction for long-read sequencing). 1 µg of genomic DNA was prepared for the ligation sequencing kit by diluting in 48 µl of nuclease-free water. A DNA repair and end preparation reaction was carried out by mixing 48 µl of DNA, 3.5 µl of NEBNext FFPE DNA Repair Buffer, 2 µl of NEBNext FFPE DNA Repair Mix, 3.5 µl of Ultra II End-prep reaction buffer and 3 µl of Ultra II End-prep enzyme mix. The reaction was mixed by flicking the tube and briefly spinning down. The reaction was incubated at 20ºC for 5 minutes followed by 65ºC for 5 mins in a Mastercycler Pro thermal cycler (Eppendorf). The reaction was transferred to a 1.5 ml LoBind tube (Eppendorf) and 60 µl of AMPure XP beads (Beckman Coulter) were added and mixed by flicking the tube. The reaction was incubated on a Hula mixer (Invitrogen) at room temperature for 5 mins. The AMPure XP beads were pelleted on a magnet and the supernatant was removed and discarded. The pellet was washed by adding 200 µl of 70% (v/v) ethanol. Without disturbing the pellet, the ethanol was removed and discarded and a further 200 µl of 70% (v/v) ethanol was added. The ethanol was removed, and the pellet was allowed to air dry for 30 s. The tube was removed from the magnetic rack and 25 µl of nuclease-free water was added, this was incubated for 2 mins at room temperature. The AMPure XP beads were pelleted using a magnet until the liquid was clear, 25 µl of the clear liquid was removed to a clean 1.5 ml LoBind tube.

2.1.17.2 Barcode ligation

The native barcodes were then ligated to the end-repaired DNA. 500 μg of end-repaired DNA was first diluted to a final volume of 22.5 μl in nuclease-free water and mixed with 2.5 μl of a single native barcode (taken from either EXP-NBD104 or EXP-NBD114 barcode kits (Oxford Nanopore)) and 25 μl of Blunt/TA Ligase Master Mix (NEB). The reaction was incubated for 10 mins at room temperature and then cleaned up with AMPure XP beads as the previous section however the DNA was eluted in 26 μl of nuclease-free water.

2.1.17.3 Adapter ligation

DNA from each barcode ligation reaction was pooled in equimolar quantities to a total DNA amount of 700 ng and the volume adjusted to 65 μl with nuclease-free water. The pooled DNA was mixed with 5 μl of Adapter Mix II (Oxford Nanopore), 20 μl of NEBNext Quick Ligation Reaction Buffer (5X) (NEB) and 10 μl of Quick T4 DNA Ligase (NEB) then incubated for 10 mins at room temperature. The DNA was then cleaned up using AMPure XP beads as previously described however during the wash steps 250 μl of Long Fragment Buffer (Oxford Nanopore) was used instead of 70 % (v/v) ethanol and the DNA was eluted with 15 μl of Elution Buffer (Oxford Nanopore).

2.1.17.4 Library preparation for loading

The sequencing library was prepared by mixing 37.5 μl of Sequencing Buffer (Oxford Nanopore) with 25.5 μl of Loading Beads and 12 μl of the adapter ligated DNA. The sequencing library was then loaded onto a pre-primed R.9.4.1 flowcell (Oxford Nanopore) and sequenced for 72 hours or until no pores remained actively sequencing.

2.1.18 Short-read sequencing

Illumina short-read sequencing was carried out by microbesNG (Birmingham). DNA was extracted from the bacterial isolates by incubating with lysozyme and RNase A for 25 mins at 37ºC. Proteinase K and more RNase A were added and incubated at 65ºC for 5 mins. The genomic DNA was purified using SPRI beads and the eluted DNA was resuspended in Elution buffer. The Nextera XT Library Prep Kit (Illumina) was used for library preparation according to the manufacturer's instructions apart from the following alterations. Two nanograms of DNA was used as input to the kit and the PCR extension time was increased from 30 s to 1 min. The libraries were sequenced on an Illumina HiSeq 2500 using a 250 bp paired end protocol. The demultiplexed sequencing reads were provided with the adapter and barcodes removed.

2.1.19 Broth Microdilution Minimum Inhibitory Concentration Assay

The strains to be tested were grown for 16 hours in Iso-Sensitest broth (Oxoid) at 37ºC with shaking (200 RPM). The antibiotics to be tested were dissolved in the relevant solvent at a concentration 2x greater than the highest antibiotic concentration to be tested. 50 µl of Iso-Sensitest broth was added to columns 2 – 11 and 100 µl to column 12 of a round bottomed 96-well plate (Corning). 100 µl of the antibiotic solution was added to column 1. 50 µl of the antibiotic solution was withdrawn from column 1 and added to column 2 and mixed by pipetting up and down. This was repeated for columns 2 to 10 and the media removed from column 10 was discarded. The absorbance (OD_{600}) of the bacterial culture was measured and diluted to OD_{600} = 0.01 in Iso-Sensitest broth. 50 µl of the diluted culture was added to columns 1 – 11 and the 96-well plate was incubated for 16 hours at 37ºC. The MIC was determined as the lowest concentration of antibiotic with no visible growth.

2.1.20 Glycerol stocks

Bacterial isolates were stored long-term by creating glycerol stocks. A single colony was picked and inoculated into 5 ml of an appropriate medium and grown at 37ºC for 16 hours with shaking (200 RPM). 750 µl of the culture was then mixed with 750 µl of 50% (v/v) glycerol to create a stock with a final glycerol concentration of 25% (v/v). The glycerol stocks were stored at -80ºC.

2.1.21 Bioinformatic Analyses

Bioinformatic analyses were carried out on the Cloud Infrastructure for Microbial Bioinformatics (CLIMB) servers unless otherwise stated. Specific bioinformatic techniques will be described in detail in the chapter-specific methods. The programming language "R" (r-project.org) was used extensively throughout this thesis for the analysis and visualisation of data. Specific R packages are referenced within the text. GraphPad Prism v8.3.1 was also used for statistical analysis and visualisation of data.

2.2 Chapter 3 methods

2.2.1 Site selection

Paired freshwater surface water and sediment samples were collected in Bangladesh from 24 surface water sites across three districts (Mymensingh, Shariatpur and Dhaka). These sites spanned both rural and urban areas with different population densities. Samples were collected from 11 aquaculture ponds in the rural areas of two districts (Mymensingh and Shariatpur) with high commercial aquaculture activity. These ponds all had a history of antibiotic use within the past three months of collection. Six ponds with no history of antibiotic use were also sampled from these rural areas. In Mymensingh, 3 ponds used for domestic purposes were selected, while in Shariatpur, these were aquaculture ponds with no prior antibiotic use, which were used for culturing fingerlings. Antibiotic use information for the ponds was collected from local dealers who were responsible for supplying fish feed for these ponds. In addition to rural surface water sites, 7 water bodies (rivers, lakes and public ponds) were sampled in Dhaka. The public ponds were heavily used for domestic purposes and, while some had history of casual (non-commercial) fish cultivation, none of them had any prior antibiotic use.

2.2.2 Sample collection

Samples were named using the following scheme; water (W) or sediment (S) followed by aquaculture (A) or control (C; ponds without antibiotic use). Sample sites were designated using (M) Mymensingh, (S) Shariatpur or (D) Dhaka and a number was included to differentiate samples. Further metadata on the samples, including temperature, pH and dissolved oxygen levels was collected. Water samples were collected by submerging a sterile 500 ml Nalgene plastic bottle approximately 15 cm below the water's surface. Bottles were capped before being removed from the water. The water samples were filtered through a

0.22 μm Sterivex-GP filter (Millipore) until water would no longer pass through the filter. The filter units were then capped and stored in a cool box and transported to the laboratory within 12 hours of sampling. In addition to the water samples, approximately 10 g of sediment was taken from either the bed of the pond or from the bank 30 – 50 cm below the surface of the water. The sediment samples were stored in sterile 50 ml Falcon tubes and were transported with the water samples.

2.2.3 Selective culturing for coliforms in surface water and sediment samples

Water and sediment samples were screened for the presence of ESBL-producing coliforms by quantitative plating on Brilliance ESBL agar (Oxoid). For sediment, 0.1 g of the sample was mixed with 0.9 ml of sterile saline solution, and 50 μl was spread onto the agar plates. For water samples, 50 μl of the undiluted sample was spread onto the media. Plates were incubated for 48 hours (h) at 37°C. In accordance with the manufacturer's instructions, blue, pink, and green colonies were designated coliforms and counted.

2.2.4 DNA extraction and Illumina sequencing

DNA was extracted from the Sterivex filters and sediment samples using the DNeasy PowerWater Kit (Qiagen) and the DNeasy PowerSoil kit (Qiagen), respectively, in accordance with the manufacturer's instructions. DNA concentrations were quantified using the Qubit dsDNA HS assay kit (Thermo Fisher) with all samples yielding more than 0.2 ng/μl. Negative control runs were performed for both kits by isolating DNA from sterile, distilled water: these yielded no detectable DNA. Metagenomic DNA libraries were prepared using the Nextera XT Library Prep Kit (Illumina). The libraries were pooled and sequenced on the HiSeq 2500 sequencing platform (Illumina) using a 150 bp paired-end protocol. Paired reads were adapter trimmed and both duplicates and reads less than 50 bp

were removed using Trimmomatic v.0.30 with Q15 as the sliding-window quality cut-off (Bolger, Lohse and Usadel, 2014). The short-read sequencing data for this project has been deposited under accession number PRJEB39306.

2.2.5 Taxonomic Profiling

To perform taxonomic profiling, the paired-end sequencing reads were mapped against clade specific markers using the MetaPhlAn2 package (Truong *et al.*, 2015). The MetaPhlAn2 package was run with default parameters. The utility script merge_metaphlan_tables.py was used to merge all of the output files into a single tab delimited file.

2.2.6 Source-sink analysis

Raw sequence reads from projects PRJNA254927, PRJEB7626 and PRJEB6092, which had previously been used as sources for source-sink analysis (McGhee *et al.*, 2020), were downloaded from the European Nucleotide Archive (ENA). These sequences represented freshwater, soil and gut metagenomes respectively. Adapters were removed from the sequence reads using fastp (Chen *et al.*, 2018). Taxonomic counts were created for these metagenomic sequences and the 48 samples in this study by kraken2 v.2.0.9 (Wood, Lu and Langmead, 2019) and Bracken v.2.6.0 (Lu *et al.*, 2017) using a database containing bacterial, archaeal, viral and fungal sequences. A metadata table was created which described the environment that the sample was from and designated it as either a source or a sink. The taxonomic count table and the metadata table were used as input to the R package FEAST v.0.1.0 (Shenhav *et al.*, 2019) which determined the proportion that each source contributed to each sink.

2.2.7 Resistome profiling

Antibiotic resistance genes were identified using the ShortBRED package v.0.9.5 (Kaminski *et al.*, 2015). The CARD database (Alcock *et al.*, 2020) (downloaded 1st July 2019) and the UniRef90 database (downloaded 4 July 2019) were used by ShortBRED-Identify to construct a marker database which the metagenomic reads could be mapped against. ShortBRED-Quantify.py was then used to map these paired-end reads against the database (cutoffs, ≥95% identity and ≥95% coverage). The relative abundance in reads per kilobase of reference sequence per million sample reads (RPKM) was generated for each resistance gene family in the database. The RPKMs were summed for antibiotic resistance genes belonging to the same class and visualised with the pheatmap package (https://cran.r-project.org/web/packages/pheatmap/pheatmap.pdf) in R (R Core Team, 2017).

2.2.8 Reconstruction of plasmids from metagenomic datasets

Metagenomic sequencing reads were assembled using the MEGAHIT v1.1.3 assembler using default parameters (Li *et al.*, 2015). Contigs produced by MEGAHIT were then classified as plasmid or chromosomal by trained neural networks in the PlasFlow v.1.1 program (Krawczyk, Lipinski and Dziembowski, 2018). Contigs designated to be of plasmid origin were queried against the CARD database by ABRicate v.0.9.8 (https://github.com/tseemann/abricate) to identify the presence of antibiotic resistance genes. Resistance genes were identified which had at least 95% identity and 50% coverage compared to the CARD database. Plasmid contigs were similarly queried against the PlasmidFinder database (Carattoli *et al.*, 2014) to identify replication genes. Plasmids were circularised by comparing 300 bp from either end of putative plasmid-containing contigs using BLASTn (Altschul *et al.*, 1990). When ends were found to overlap, one copy of the overlapping sequence was removed to generate a complete, circularized plasmid sequence.

To avoid spurious overlaps, overlapping regions were manually inspected to ensure that they were not part of translocatable elements.

2.2.9 Statistical analyses

The Shannon Diversity Index of the samples was calculated in R v.3.4.3 using the diversity function of the vegan package v.2.5-7 (Oksanen *et al.*, 2019). Non-metric multidimensional scaling (NMDS) was also performed in R using the metaNMDS function of the vegan package. Permutational multivariate analysis of variance (PERMANOVA) was performed on a Bray-Curtis distance matrix of species abundance in R using the adonis function of the vegan package. Correlation between total ARG abundance and human gut bacterial contribution was calculated using the lm function in base R. Additional tests for determining statistical significance were performed as described in the text and implemented in GraphPad Prism v.8.3.1.

2.3 Chapter 4 methods

2.3.1 Collection of isolates

E. faecium isolates were collected from Heartlands Hospital in Birmingham, UK over a period of 5 years (2016-2020). The first set of isolates (n=39) were collected during an outbreak of *Enterococcus faecium* associated bacteriaemia in 2016 and 2017. The second set of isolates (n=21) were routine screening isolates collected in 2019 and 2020. Anonymised metadata was collected for all of the patients. As these isolates were collected as part of outbreak surveillance and routine screening, no ethical approval was required.

2.3.2 Genome assembly

Adapters were removed from the short-read data and reads were trimmed using Trimmomatic v.0.39 (Bolger, Lohse and Usadel, 2014) with a sliding window cut-off of Q15. Short-reads were assembled with shovill v.1.0.5 (Seeman T, shovill, https://github.com/tseemann/shovill) using default parameter. SPAdes v.3.14.0 (Bankevich *et al.*, 2012) was used within shovill to assemble the reads. Isolates that were both long and short read sequenced were hybrid assembled using both sets of reads. Unicycler v.0.4.8 (Wick *et al.*, 2017) was used with default parameters to obtain closed genomes and closed plasmid sequences. The quality metrics of the assemblies were determined using Quast v.5.0.2 (Gurevich *et al.*, 2013). The average nucleotide identity to *E. faecium* strain E745 was calculated using FastANI v.1.32 (Jain *et al.*, 2018). The *E. faecium* genomes were annotated using Prokka v.1.14.6 (Seemann, 2014).

2.3.3 Sequence type analysis

The *E. faecium* isolate sequence types were determined by querying their short-read assemblies against PubMLST (Jolley and Maiden, 2010) using mlst v.2.18.0 (Seemann T, mlst, https://github.com/tseemann/mlst).

2.3.4 Phylogenetic analysis

A phylogenetic tree was constructed for the clinical *E. faecium* isolates by mapping the short-reads against the *E. faecium* E745 reference genome (GCA_001750885.1) using snippy v.4.6.0 (Seemann T, 2015 snippy: fast bacterial variant calling from NGS reads https://github.com/tseemann/snippy). A core genome SNP alignment was created using snippy-core. The core SNP alignment was passed to Gubbins v.2.4.1 (Croucher *et al.*, 2015) to mask recombination. A phylogenetic tree was then constructed using RaxML v.8.1.15 (Stamatakis, 2014) implementing the GTRGAMMA DNA substitution model.

2.3.5 Short-read resistance genes

The antibiotic resistance gene profiles for the clinical *E. faecium* isolates were determined by querying the short-read assemblies against the Resfinder database (2020) (Zankari *et al.*, 2012) using ABRicate v.0.9.8 (Seemann T, ABRicate, https://github.com/tseemann/abricate). A minimum identity cut-off of 95% and minimum coverage cut-off of 50% was used. To identify resistance to linezolid, the *E. faecium* short reads were queried using LRE-Finder v.1.0.0 (Hasman *et al.*, 2019).

2.3.6 Reversion of VVE to a resistant phenotype

A single VVE susceptible isolate (HHI2, HHI3 or OI25) was inoculated into 5 ml of BHI broth and grown for 16 hours at 37^{o}C with shaking (200 RPM). The bacterial culture was then diluted 1/100 into 5 ml of BHI broth containing 8 μg/ml vancomycin. The culture was

grown at 37°C with shaking (200 RPM) and the cultures were observed every 24 hrs until growth was observed (broth had become turbid). The cultures were then diluted 10^{-6} and 100 µl was spread onto BHI agar plates containing 8 µg/ml vancomycin. Two colonies from each plate were cultured in BHI broth containing 8 µg/ml vancomycin for 16 hours at 37°C with shaking (200 RPM). Glycerol stocks were then made, and the bacterial isolates were stored at -80°C for further use.

2.3.7 Terminator analysis

The Rho-independent terminator of the ribosomal RNA gene operon was identified in isolate OI25 by taking 100 nucleotides downstream of the 5S stop codon and submitting them to the RNAfold Web Server (Lorenz *et al.*, 2011). The output was then manually inspected to identify the typical A-tail, Loop, T-tail structure of a Rho-independent terminator (Kingsford, Ayanbule and Salzberg, 2007).

2.3.8 RT-qPCR

2.3.8.1 Vancomycin challenge

The *E. faecium* strains were first streaked onto BHI agar and incubated at 37°C for 48 hours. A single colony was used to inoculate 5 ml of BHI broth and incubated at 37°C for 16 hours with shaking (200 RPM). The *E. faecium* cultures were then diluted 1/100 by inoculating 200 µl in 19.8 ml of BHI broth, they were then grown at 37°C with shaking (200 RPM) to an OD_{600} = 0.5. 2 ml of the culture was transferred to a 2 ml Eppendorf tube for RNA extraction. Vancomycin was added to the remaining culture to a final concentration of 8 µg/ml and incubated for a further hour at 37°C with shaking (200 RPM). A further 2 ml of the culture was then removed from each culture to a 2 ml Eppendorf tube for RNA extraction.

2.3.8.2 RNA extraction

The Total RNA miniprep kit (New England Biolabs) was used to extract total RNA from the *E. faecium* cells. The bacterial cells were harvested from 2 ml of culture by centrifugation at 16,000 x g for 30 s and the supernatant was discarded. The cell pellet was resuspended in 700 µl of 1X DNA/RNA Protection reagent (New England Biolabs) and transferred to a lysing matrix B tube (MP Biomedicals). The cells were lysed using a FastPrep-24 5G device at a setting of 6 m/s for 40 s. The homogenate was then transferred to a fresh 1.5 ml Eppendorf tube and the debris was pelleted by centrifugation at 16,000 x g for 2 mins. The supernatant was transferred to a fresh tube 1.5 ml Eppendorf tube and an equal volume of RNA lysis buffer (New England Biolabs) was added and vortexed briefly. 800 µl of the mixture was added to a gDNA Removal column (New England Biolabs) and then centrifuged at 16,000 x g for 30 s. For samples with more than 800 µl, the column was reloaded, and the centrifugation step was repeated. An equal volume of 95% (v/v) ethanol was added to the flow-through from the gDNA removal column and mixed by pipetting. The mixture was then added to an RNA purification column (New England Biolabs), centrifuged for 30 s at 16,000 x g and the flow-through was discarded. 500 µl of RNA Wash Buffer (New England Biolabs) was applied to the column, centrifuged for 30 s at 16,000 x g and the flow-through was discarded. 5 µl of DNase I was combined with 75 µl of DNase I Reaction Buffer and applied directly to the matrix of the RNA purification column, it was then incubated at room temperature for 15 mins. 500 µl of RNA Priming Buffer was added to the column, centrifuged at 16,000 x g for 30 s and the flow-through was discarded. 500 µl of RNA Wash Buffer was then added to the column and centrifuged at 16,000 x g for 30 s. Another 500 µl of RNA Wash Buffer was added to the column and centrifuged at 16,000 x g for 2 mins. The RNA purification column was then moved to a new 1.5 ml Eppendorf

tube, 50 µl of nuclease-free water was added directly to the membrane and centrifuged for 30 s. The TURBO DNA-*free*™ kit (Invitrogen) was then used to remove any residual traces of DNA in the extracted RNA. 5 µl of 10X TURBO DNase™ buffer and 1 µL of TURBO DNase™ Enzyme were added to the RNA and mixed gently by flicking the tube. The RNA sample was then incubated at 37°C for 30 mins. The reaction was stopped by adding 6 µl of DNase Inactivation Reagent and flicking to mix. The samples were incubated at 25°C for 5 mins and were mixed by flicking every minute. The samples were centrifuged at 10,000 x g for 1.5 mins to pellet the DNase Inactivation Reagent and the supernatant was transferred to a fresh 1.5 ml RNase-free microfuge tube. The eluted RNA was stored at -80°C until further use.

2.3.8.3 cDNA synthesis

Maxima First Strand cDNA Synthesis Kit for RT-qPCR (Thermo Scientific™) was used for cDNA synthesis. A reaction was set up with included 4 µl of 5X Reaction Mix, 2 µl of Maxima Enzyme Mix, 100 ng of RNA and the reaction was made up to 20 µl with nuclease-free water. The reaction was mixed by pipetting and pulse centrifuged to collect the liquid at the bottom of the tube. The reaction was then incubated at 25°C for 10 mins followed by 50°C for 15 mins, the reaction was stopped by heating to 85°C for 5 mins. A second reaction was carried out which was identical to the first, except the Maxima Enzyme Mix was replaced with nuclease-free water, this was used as the no reverse transcriptase control for the subsequent qPCR reactions. The synthesised cDNA was stored at -80°C until further use.

2.3.8.4 qPCR

qPCR reactions were set up as follows: 10 µl PrimeTime® Gene Expression Master Mix (2X) (Integrated DNA Technologies (IDT)), 1 µl PrimeTime® qPCR Assay (20X) (IDT), 2

µl cDNA and 7 µl Nuclease-free water. PrimeTime qPCR assays (IDT) which contain the forward primer, reverse primer and probe were created for *vanRS*, *vanHAX* and *tufA* (Table 2.7) using the PrimerQuest™ tool (IDT). qPCR reactions were carried out in biological triplicate and technical duplicate. No template controls and no reverse transcriptase controls were included to ensure no genomic DNA was carried over (Ct > 30). The qPCR reactions were transferred to a MicroAmp® Optical 96-Well reaction plate (Applied Biosystems) and sealed with an Optical Adhesive Cover (Applied Biosystems). The qPCR reaction was carried out in a QuantStudio 1 Real-Time PCR system (Applied Biosystems) using the following program:

Table 2.5: RT-qPCR program.

Step	Time	Temperature	Number of cycles
Polymerase activation	3 mins	95ºC	1
Denaturation	15 s	95ºC	40
Annealing/Extension	1 min	60 ºC	

2.3.9 Variant calling

Variants were called between the VVE revertant isolates and the VVE parent strains using snippy v.4.6.0 (Seemann T, 2015 snippy: fast bacterial variant calling from NGS reads https://github.com/tseemann/snippy).

2.3.10 Identification of plasmids in short-read data

Plasmid replicons were identified in the *E. faecium* isolates by querying the short-read assemblies against the PlasmidFinder database (Carattoli *et al.*, 2014) using ABRicate v.0.9.8 (Seemann T, ABRicate, https://github.com/tseemann/abricate). A minimum identity cut-off of 95% and minimum coverage cut-off of 50% was used.

2.3.11 RT-PCR of rRNA-*vanHAX* junctions

RNA was extracted from the OI25 revertant isolates and reversed transcribed into cDNA as in (2.3.8.2 RNA extraction) and (2.3.8.3 cDNA synthesis). PCR was performed from the 23S rRNA gene across the junction into each of the *vanHAX* genes using primers Junc_fwd, Junc_vanH_rev, Junc_vanA_rev and Junc_vanX_rev (Table 2.7) according the protocol in (2.1.4 Polymerase Chain Reaction). The PCR products were visualised on a 1% (w/v) agarose gel.

2.4 Chapter 5 methods

2.4.1 Preparation of samples for RNA-seq

E. faecium strains E745 and E8202 (Table 2.10) were streaked onto BHI agar plates and incubated for 20 hrs at 37ºC to obtain single colonies. 5 ml of BHI broth was then inoculated with a single colony of either E745 or E8202 (cultures were prepared in triplicate), they were grown for 16 hrs at 37ºC with shaking (200 RPM). 20 ml of BHI broth and 20 ml of BHI broth plus vancomycin (E745: 128 μg/ml and E8202: 64 μg/ml) were inoculated with 200 μl of the overnight culture and grown to an OD_{600} = 1. 1500 μl of culture was transferred to a 1.5 ml Eppendorf tube and centrifuged at 16,000 x g for 30 s, the supernatant was discarded, and the previous step was repeated with a further 1500 μl. Tubes were centrifuged briefly and any remaining BHI broth was removed. The cell pellets were then flash frozen by submerging the tubes in a slurry of dry ice and 100 % ethanol. The cell pellets were stored at -80ºC until use.

2.4.2 RNA extraction, library preparation and sequencing

RNA extraction, library preparation and sequencing were carried out by GENEWIZ From Azenta Life Sciences (New Jersey, USA). RNA was extracted using the RNeasy Plus Universal kit (Qiagen) following the manufacturers protocol. RNA quality was measured using the Qubit 2.0 fluorometer and Agilent TapeStation, samples were taken forward if they had an RNA integrity number (RIN) > 6, DIV_{200} score > 70 and DNA concentration < 10% of the RNA concentration. Ribosomal RNA was removed from the samples using the Ribo-Zero rRNA Depletion kit (Illumina). The sequencing library was prepared using the NEBNext® Ultra™ II RNA Library Prep Kit for Illumina® (NEB). The quality of the sequencing library was assessed using the Qubit 2.0 fluorometer and KAPA qPCR assay.

The prepared libraries were mixed in equimolar concentrations and sequenced on the Illumina HiSeq 4000 sequencing platform.

2.4.3 RNA data analysis

The quality of the sequencing data was assessed using FastQC v.0.11.9 (Andrews S, FastQC, https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). DNA assemblies were downloaded from Genbank for *E. faecium* strains E745 (GCA_001750885.1) and E8202 (GCA_900639535.1) and annotated using PROKKA v.1.14.6 (Seemann, 2014) using the *Enterococcus* specific database. Kallisto v.0.44.0 (Bray *et al.*, 2016) was used to pseudo-align the RNA-seq reads and against the *E. faecium* transcripts and quantify the reads per gene. The annotation files in genbank format were converted to a kallisto input using the script genbank_to_kallisto.py (https://github.com/AnnaSyme/genbank_to_kallisto.py). Voom/limma v.3.50.1 (Law *et al.*, 2014) in Degust v.4.1.5 (David R. Powell. Degust: interactive RNA-seq analysis, DOI: 10.5281/zenodo.3258932.) was used to analyse differential gene expression (DGE) between the vancomycin treated samples and the controls. Further visualisation of the DGE results was performed in R, v.3.4.3.

2.4.4 Functional annotation of significant genes identified by RNA-seq or Tn-seq

The genome assemblies of *E. faecium* E745 and *E. faecium* E8202 were downloaded from Genbank and submitted to the eggNOG-mapper webserver v.2.1.7 (Cantalapiedra *et al.*, 2021). The genes were then manually curated into different functional categories by the COG letter that they were assigned. The coding sequences of genes which were significantly upregulated or downregulated in the RNA-seq analysis or significant in the Tn-seq analysis were also submitted to the eggNOG-mapper webserver. Again, the genes were manually curated into different functional categories. The percentage of significant genes in each

functional category was calculated by dividing the number of genes in that category identified as being significant by the total number of genes in that category and multiplying by 100.

2.4.5 Vancomycin treatment of the Tn-seq libraries

The transposon mutant libraries for *E. faecium* strains E745 (Zhang *et al.*, 2017), E8202 and Aus0004 (both generated by Dr Vincent de Maat, University Medical Centre Utrecht, The Netherlands (de Maat, 2022)) were recovered by adding 20 µl of the transposon library to 20 ml of BHI broth and growing for 16 hours at 37°C with shaking (200 RPM). 20 µl of the cultured library was then added to 20 ml of BHI broth and grown at 37°C with shaking until OD_{600} = 1, similarly 20 µl of the cultured library was added to 20 ml of BHI broth that was supplemented with ½ MIC vancomycin (E745 = 128 µg/ml, Aus0004 = 8 µg/ml & E8202 = 64 µg/ml) and grown at 37°C with shaking. This was carried out in quadruplicate for the control and the treatment groups.

2.4.6 Library preparation

2.4.6.1 DNA digestion with MmeI

Genomic DNA was extracted from 2 ml of each culture (see 2.1.14 Genomic DNA extraction for short-read sequencing). 2 µg of genomic DNA was digested with 10 U of MmeI (NEB) at 37ºC for 4 hours before being dephosphorylated with 1 U of Calf Intestinal Alkaline Phosphatase (Invitrogen™) at 50ºC for 30 mins. The DNA was then re-purified by adding 200 µl of phenol:chloroform:isoamyl alcohol (Sigma) to the sample. The sample was then vortexed, centrifuged at 17,000 x g for 5 minutes and the top layer was transferred to a fresh 2 ml tube containing 200 µl of chloroform:isoamyl alcohol (Sigma). As with the previous step, the samples were vortexed, centrifuged at 17,000 x g for 5 minutes and the top layer

was removed to a fresh 2 ml tube. 20 µl of 3M NaAc pH 5.3 (Thermo Scientific™), 0.5 µl of glycogen (10 mg/ml) (VWR) and 500 µl of 100% ethanol were added and mixed well. The samples were incubated at -20ºC for 16 hours and then centrifuged at 17,000 x g for 10 mins at 4°C to pellet the DNA. The DNA pellet was washed with 500 µl of 70% (v/v) EtOH, centrifuged at 17,000 x g for 5 mins at 4°C, the supernatant was discarded, and the pellet was air dried for 10 mins. The DNA pellet was then resuspended in 20 µl of nuclease-free water.

2.4.6.2 Adapter annealing and 5' phosphorylation

Primers containing a barcode sequence (A-L) and a partial TruSeq sequence, were annealed together to form adapters (Table 2.8). 5 µl of 1 nmol/µl PBGSF29-Barcode, 5 µl of the corresponding PBGSF30-Barcode, 0.5 µl 100x Tris-EDTA (Sigma-Aldrich), 0.5 µl 5 M NaCl (Sigma-Aldrich) and 39 µl of nuclease-free water (NEB) were mixed together and incubated at 95ºC for 10 mins in a heat block. The metal tube holder was removed and allowed to cool to room temperature over a time period of ~45 mins. The adapters were then 5'-phosphorylated by mixing 2 µl of the annealed adapter (100 pmol/µl), 2 µl of 10x T4 DNA ligase buffer (NEB), 0.5 µl 10 U/µl T4 polynucleotide kinase (3' phosphatase minus) (NEB) and 15.5 µl dH2O. The reaction was initially incubated at 37ºC for 5 mins and then incubated at 70ºC for 10 mins in a heat block, the heat block was then removed and cooled to room temperature.

2.4.6.3 Adapter ligation to genomic DNA

Next the adapters were ligated to the digested genomic DNA by mixing 100 ng of dephosphorylated MmeI restriction fragments, 0.2 µl (10 pmol/µl) phosphorylated & annealed adapter (adapters with different barcodes were used for each sample), 2 µl 10x T4

DNA ligase buffer (NEB), 0.2 µl (10 U/µl) T4 DNA ligase (NEB) and the reaction was made up to 20 µl with nuclease-free water. The ligation reaction was incubated at 16ºC for 60 mins.

2.4.6.4 PCR amplification, sequencing adapter addition and sequencing

The following PCR reaction was set up: 26 µl nuclease-free water, 10 µl 5x Phusion HF buffer (NEB), 1 µl 10 mM dNTP mix (NEB), 5 µl (4 pmol/µl) PBGSF40 primer, 5 µl (4 pmol/µl) PBGSF31 primer, 2.5 µl ligation mixture and 0.5 µl Phusion DNA polymerase (NEB). The PCR reaction was performed according to the following program: 72°C 1 min, 98°C 30 s, 25 cycles (98°C 10 s, 57°C 30 s, 72°C 10 s) and 72°C for 5 mins. The PCR reactions were cleaned up using the MinElute Reaction Cleanup Kit (Qiagen) in the following reaction. The PCR reaction was mixed with 50 µl of nuclease-free water, 300 µl of ERC buffer and transferred to a MinElute column. The column was centrifuged at 13,000 x g for 1 min and the flow through was discarded. 750 µl of PE buffer was then added to the column, it was once again centrifuged at 13,000 x g for 1 min and the flowthrough was discarded. The column was then centrifuged at 17,000 x g for 1 min to completely dry the membrane. The column was moved to a fresh 1.5 ml Eppendorf tube and 10 µl of nuclease-free water was added directly to the membrane. After 1 min, the column was centrifuged at 17,000 x g for 1 min and the flowthrough containing the DNA was retained. The DNA concentration was measured (see 2.1.16 Measurement of DNA concentrations). Equimolar quantities of the barcoded DNA were mixed together and sequenced on an Illumina NovaSeq platform using a SE50 protocol.

2.4.7 Tn-seq data analysis

Raw reads were first demultiplexed and the barcodes removed using Cutadapt v.3.0 (Martin, 2011). The reads were then cropped to 16 bp to remove the transposon sequence and leave only genomic DNA using Trimmomatic v.0.39 (Bolger, Lohse and Usadel, 2014). The reads were aligned to the parent strains genome using the BWA algorithms aln and samse v.0.7.17-r1188 (Li and Durbin, 2009). Reads were counted per gene using the Bio-Tradis v.1.4.1 scripts tradis_plot and tradis_gene_insert_sites, reads were not included which mapped to the first and last 10% of the genes. The outputs of Bio-Tradis were processed in R v3.4.3. DESeq2 v.1.30.0 was then used to calculate the fold-change of reads between the control experiment and the vancomycin experiment, a Benjamini-Hochberg adjusted *P* value was also produced to correct for multiple testing (Love, Huber and Anders, 2014). Genes which had an adjusted *P* value < 0.05 and a $\log_2$ fold-change less than -1 were considered to contribute to increased fitness in the presence of vancomycin.

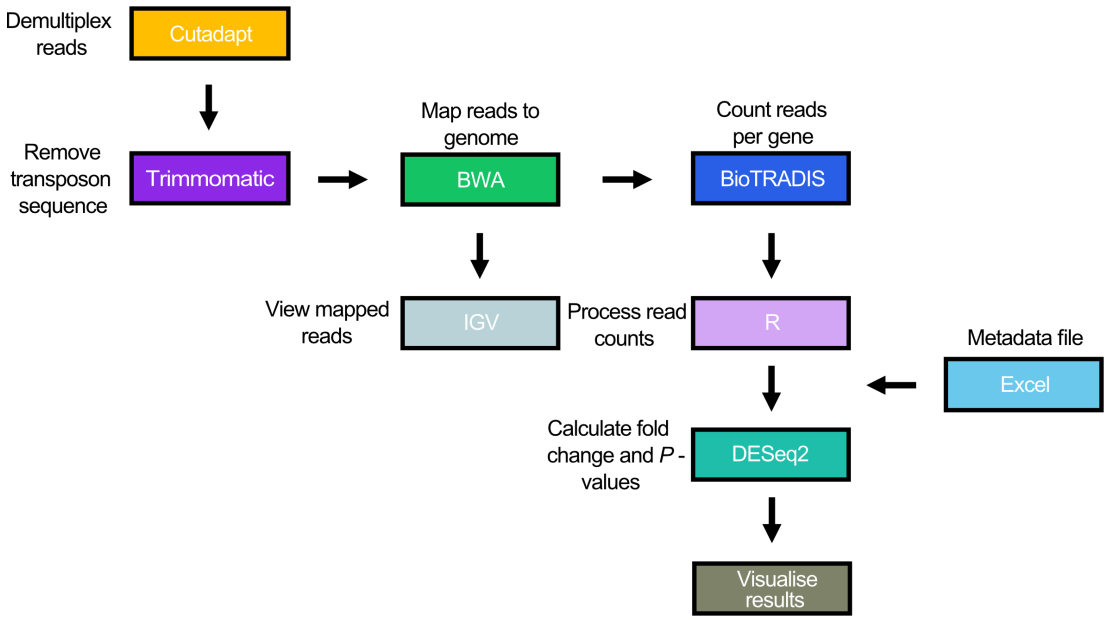

Figure 2.1: An outline of the bioinformatics pipeline used to analyse and visualise data from the Tn-seq experiment.

2.4.8 Phylogenetic analysis

A phylogenetic tree was constructed of the 1644 *E. faecium* isolates using mash distances in Mashtree v.1.1.2 (Katz *et al.*, 2019). A phylogenetic tree was constructed for the *Enterococcus* genus using MEGAX v.10.01.08 (Kumar *et al.*, 2018). The 16S rRNA genes were extracted from the genomes using Barrnap v.0.9 (Seeman T, Barrnap, https://github.com/tseemann/barrnap) and MEGAX was used to create an alignment of the 16S rRNA genes and construct a phylogenetic tree. A custom database was created that contained the nucleotide sequences of *asnB*, *avrA*, *glpO*, *gshAB*, *mgs* and *vanX.* The assemblies of 1644 *E. faecium* clade A isolates (Arredondo-Alonso *et al.*, 2020) were queried against the custom database using ABRicate v.0.9.8 (Seemann T, ABRicate, https://github.com/tseemann/abricate) with minimum identity and minimum coverage cut-offs of 95% and 50% respectively. Similarly, the assemblies of 24 species of *Enterococcus*, 3 species from genera closely related to *Enterococcus* and 2 outgroup species were queried against the database using the same method.

Table 2.6: Species used in phylogenetic analyses.

Species	Accession number
Enterococcus	
E. asini	NZ_ASVU01000001.1
E. avium	NZ_KE136500.1
E. caccae	NZ_KE136472.1
E. casseliflavus	NZ_GG670403.1
E. cecorum	NZ_ASWI01000001.1
E. columbae	NZ_KE136491.1
E. dispar	NZ_ASWK01000001.1
E. durans	NZ_KE136512.1
E. faecalis OG1RF	NC_017316.1
E. faecalis V583	NC_004668.1
E. faecium Aus0004	CP003351.1
E. faecium EnGen0015	NZ_KB029496.1
E. faecium EnGen0043	NZ_KB029937.1
E. gallinarum	NZ_GG670299.1
E. gilvus	NZ_ASWH01000001.1
E. haemoperoxidus	NZ_KE136479.1
E. hirae	NZ_ASVZ01000001.1

E. italicus	NZ_GL622241.1
E. malodoratus	NZ_KE136481.1
E. moraviensis	NZ_ASWB01000001.1
E. mundtii	NZ_ASWC01000001.1
E. pallens	NZ_ASWD01000001.1
E. phoeniculicola	NZ_ASWE01000001.1
E. raffinosus	NZ_ASWF01000001.1
E. saccharolyticus	NZ_KE136522.1
E. sulfureus	NZ_ASWO01000001.1
E. villorum	NZ_KE136487.1
Closely related species	
Melissococcus plutonius	NC_015516.1
Tetragenococus halophilus	NC_016052.1
Vagococcus lutrae	NZ_AYSH01000001.1
Outgroup species	
Carnobacterium maltaromaticum	NC_019425.2
Lactococcus garvieae	NZ_ASWT01000001.1

2.4.9 Gene deletions

Genes *avrA*, *avrB*, *asnB*, *glpO* and *gshAB* were deleted using a CRISPR-Cas9 system (see 2.1.12 CRISPR protocol). Primers used to create the CRISPR sequence and amplify the deletion templates can be found in Table 2.8. The sequences of the deletion templates can be found in Table 8.1.

2.4.10 Conjugation of a gentamicin-tagged vancomycin resistance plasmid into the *E. faecium* deletion mutants

A conjugation assay was carried out between the E745 transposon library and *E. faecium* 64/3 according to the standard conjugation protocol (see 2.1.13 Conjugation protocol). The donor library and transconjugants were selected with gentamicin 300 μg/ml, the recipient and transconjugants were selected with rifampicin 25 μg/ml and fusidic acid 25 μg/ml and the transconjugants were selected with gentamicin 300 μg/ml, rifampicin 25 μg/ml and fusidic acid 25 μg/ml. A further conjugation assay was performed between 64/3 pvanA-G and the deletion mutants that did not contain the vancomycin resistance plasmid. The donor strain and transconjugants were selected with gentamicin 300 μg/ml, the recipient and

transconjugants were selected with ciprofloxacin 64 μg/ml and the transconjugants were selected with gentamicin 300 μg/ml and ciprofloxacin 64 μg/ml.

2.4.11 Inverse PCR

To map the insertion site of the gentamicin transposon in the *vanA* plasmid, genomic DNA was extracted from *E. faecium* 64/3 pvanA-G using the Wizard Genomic DNA Purification kit (Promega) (see 2.1.14 Genomic DNA extraction for short-read sequencing). 1 μg of genomic DNA was digested with HaeIII (New England Biolabs) at 37ºC for 1 hour, the enzyme was then heat inactivated at 80ºC for 20 mins. The digested DNA was then self-ligated to form a circular product, a 100 μl ligation reaction was set up which contained 10 μl of T4 DNA ligase buffer (New England Biolabs), 250 ng of digested genomic DNA, nuclease-free water (New England Biolabs) up to 100 μl and 5 μl of T4 DNA ligase (New England Biolabs). The ligation reaction was incubated at room temperature for 5 hours before being heat inactivated at 65ºC for 10 mins. The region adjacent to transposon insertion site was amplified by PCR with primers IPCR_HaeIII_F and IPCR_HaeIII_R (Table 2.8) using Phusion High-Fidelity DNA Polymerase (New England Biolabs) and an input DNA concentration of 50 ng. The cycling conditions were as follows: 98ºC for 30 s followed by 35 cycles of 98ºC for 10 s, 60ºC for 30 s and 72ºC for 5 mins; and a final extension at 72ºC for 10 mins.

2.4.12 Growth kinetics in the presence of vancomycin

The knock-out strains were grown in the presence and absence of ½ MIC vancomycin to assess their fitness when exposed to vancomycin. *E. faecium* strains were streaked onto BHI agar and incubated at 37ºC for 16 hours. Single colonies were then used to inoculate 5 ml of BHI broth and incubated for 16 hours at 37ºC with shaking (200 RPM). The liquid cultures

were initially diluted to a concentration of 2 x 10^5 with BHI broth. For the wells with no vancomycin present, 50 µl of BHI broth and 50 µl of the diluted culture were added to a flatbottomed 96-well plate. For strains exposed to ½ MIC vancomycin, 50 µl of BHI broth with 1x MIC vancomycin and 50 µl of diluted culture was added to the 96-well plate. The outer wells of the 96-well plate were filled with 200 µl of sterile water to minimise evaporation and the lid was replaced. The 96-well plate was incubated at 37^{o}C for 15 hours with shaking at 240 RPM. The absorbance of each well was measured at 600 nm every 10 mins for the 15-hour period.

2.5 Primers

Table 2.7: Chapter 4 Primers.

Primer name	Sequence (5’ -> 3’)	Description	Source
vanRS & *vanHAX* RT-PCR			
tufA_fwd	GGTGACGATGTTCCTGTAGTT	PrimeTime qPCR Probe assay targeting the *tufA* gene in *E. faecium*.	This project
tufA_probe	TGAAAGCTCTAGAAGGCGACGCTT		
tufA_rev	CGTTCTGGAGTTGGGATGTATT		
vanHAX_fwd	ATATAAAGCGCTCGGCTGTAG	PrimeTime qPCR Probe assay targeting the *vanHAX* operon in *E. faecium*.	This project
vanHAX_probe	TAACGGCCGCATTGTACTGAACGA		
vanHAX_rev	TGAAACCGGGCAGAGTATTG		
vanRS_fwd	AAGCTGGCCGAACAAAGA	PrimeTime qPCR Probe assay targeting the *vanRS* operon in *E. faecium*.	This project
vanRS_probe	ACGTTGTTATGTACTTGGCGCACG		
vanRS_rev	CGTCAAGCAGGCTCAAATAAC		
OI25 chromosomal insertion			
Junc_fwd	CGGACTGATACTAATCGATCG	Primer used to check for junction between rRNA and *vanHAX* genes in cDNA.	This project
Junc_vanH_rev	GGAATGCATCTGCCTCATC	Primer used to check for junction between rRNA and *vanH* gene in cDNA.	This project
Junc_vanA_rev	AGATTTTACCGATACGTCATGC	Primer used to check for junction between rRNA and *vanA* gene in cDNA.	This project
Junc_vanX_rev	CAACGAACACCGTGTACTAT	Primer used to check for junction between rRNA and *vanX* gene in cDNA.	This project

Table 2.8: Chapter 5 Primers.

Primer name	Sequence (5' -> 3')	Description	Source
Transposon sequencing library preparation			
PBGSF29 ATCACG	TTCCCTACACGACGCTCTTCCGATCTATCACGNN	Oligonucleotides used to form adapter A	(Zhang *et al.*, 2017)
PBGSF30 ATCACG	P-CGTGATAGATCGGAAGAGCGTCGTGTAGGGAAAGAGT-P		
PBGSF29 CGATGT	TTCCCTACACGACGCTCTTCCGATCTCGATGTNN	Oligonucleotides used to form adapter B	(Zhang *et al.*, 2017)
PBGSF30 CGATGT	P-ACATCGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGT-P		
PBGSF29 TTAGGC	TTCCCTACACGACGCTCTTCCGATCTTTAGGCNN	Oligonucleotides used to form adapter C	(Zhang *et al.*, 2017)
PBGSF30 TTAGGC	P-GCCTAAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGT-P		
PBGSF29 TGACCA	TTCCCTACACGACGCTCTTCCGATCTTGACCANN	Oligonucleotides used to form adapter D	(Zhang *et al.*, 2017)
PBGSF30 TGACCA	P-TGGTCAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGT-P		
PBGSF29 ACAGTG	TTCCCTACACGACGCTCTTCCGATCTACAGTGNN	Oligonucleotides used to form adapter E	(Zhang *et al.*, 2017)
PBGSF30 ACAGTG	P-CACTGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGT-P		
PBGSF29 GCCAAT	TTCCCTACACGACGCTCTTCCGATCTGCCAATNN	Oligonucleotides used to form adapter F	(Zhang *et al.*, 2017)
PBGSF30 GCCAAT	P-ATTGGCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGT-P		
PBGSF29 CAGATC	TTCCCTACACGACGCTCTTCCGATCTCAGATCNN	Oligonucleotides used to form adapter G	(Zhang *et al.*, 2017)
PBGSF30 CAGATC	P-GATCTGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGT-P		
PBGSF29 ACTTGA	TTCCCTACACGACGCTCTTCCGATCTACTTGANN	Oligonucleotides used to form adapter H	(Zhang *et al.*, 2017)
PBGSF30 ACTTGA	P-TCAAGTAGATCGGAAGAGCGTCGTGTAGGGAAAGAGT-P		
PBGSF29 GATCAG	TTCCCTACACGACGCTCTTCCGATCTGATCAGNN	Oligonucleotides used to form adapter I	(Zhang *et al.*, 2017)
PBGSF30 GATCAG	P-CTGATCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGT-P		
PBGSF29 TAGCTT	TTCCCTACACGACGCTCTTCCGATCTTAGCTTNN	Oligonucleotides used to form adapter J	(Zhang *et al.*, 2017)
PBGSF30 TAGCTT	P-AAGCTAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGT-P		
PBGSF29 GGCTAC	TTCCCTACACGACGCTCTTCCGATCTGGCTACNN	Oligonucleotides used to form adapter K	(Zhang *et al.*, 2017)
PBGSF30 GGCTAC	P-GTAGCCAGATCGGAAGAGCGTCGTGTAGGGAAAGAGT-P		

PBGSF29 CTTGTA	TTCCCTACACGACGCTCTTCCGATCTCTTGTANN	Oligonucleotides used to form adapter L	(Zhang *et al.*, 2017)
PBGSF30 CTTGTA	P-TACAAGAGATCGGAAGAGCGTCGTGTAGGGAAAGAGT-P		
PBGSF31	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT	Primer to amplify barcoded DNA	(Zhang *et al.*, 2017)
PBGSF40	CAAGCAGAAGACGGCATACGAGATAGACCGGGGACTTATCATCCAACCTGT	Primer to amplify barcoded DNA and add Illumina adapter	(Zhang *et al.*, 2017)

Gene deletion

asnB_crispr_fwd_2	AAACGAATTCAAAATTCATCCGTGTGATTGCTCAG	Oligonucleotide used to form *asnB* cripsr insert	This project
asnB_crispr_rev_2	AAAACTGAGCAATCACACGGATGAATTTTGAATTC		
asnB_amp_fwd	ACTATTAGTACTGCAAAAGAAAAGCATGGGATG	Primer used to amplify *asnB* gBlock	This project
asnB_amp_rev	CTGATCAGTACTAAAAAAATGGTATGAAGATGCTAGTGAATC		
asnB_check_fwd	CCAATGTACCATAAGTAAAAATCATCAAC	Primer used to check for successful KO of *asnB*	This project
asnB_check_rev	CATTATTAGATGAAGGGTACCCTGA		
avrB_crispr_fwd	AAACGAAATTCGATCGATAAAGTGAATACACAAGG	Oligonucleotide used to form *avrB* cripsr insert	This project
avrB_crispr_rev	AAAACCTTGTGTATTCACTTTATCGATCGAATTTC		
avrB_amp_fwd	ACTATTGATATCATTTCAATGAAGGAAGAGATAAGATGC	Primer used to amplify *avrB* gBlock	This project
avrB_amp_rev	CTGATCGATATCTGGTCATTTTCCATCTTGACT		
avrB_check_fwd	GTCCGTAGTAAAGTAAGGAAACAAG	Primer used to check for successful KO of *avrB*	This project
avrB_check_rev	CTTCTGAAAATTCTTTTTTGAACTCATCC		
glpO_crispr_fwd	AAACAATCAAAGCAAAATTAGTGATCAATACGAGG	Oligonucleotide used to form *glpO* cripsr insert	This project
glpO_crispr_rev	AAAACCTCGTATTGATCACTAATTTTGCTTTGATT		
glpO_amp_fwd	ACTATTAGTACTACAAGCGAAGATGAGGTC	Primer used to amplify *glpO* gBlock	This project
glpO_amp_rev	CTGATCAGTACTTACGTTCGTTCCAGGTG		
glpO_check_fwd	CTTCAAATGTTGCAAAAAGCTAGC	Primer used to check for successful KO of *glpO*	This project
glpO_check_rev	GTTGATTGCATATCCTGTTGG		
gshAB_crispr_fwd	AAACTTCTTACAATTATTCCTGATGTACATGTTAG	Oligonucleotide used to form *gshAB* cripsr insert	This project
gshAB_crispr_rev	AAAACTAACATGTACATCAGGAATAATTGTAAGAA		
gshAB_amp_fwd	ACTATTAGTACTTAGAAAAAATAAAACGAGTAATCCAGCG	Primer used to amplify *gshAB* gBlock	This project
gshAB_amp_rev	CTGATCAGTACTTATGGCACACCTTTAGTGAAC		

gshAB_check_fwd	GCAGGCAGATGTTTTACTAGTT	Primer used to check for successful KO of *gshAB*	This project
gshAB_check_rev	GGTATTTTTTATGGCACAAGATGATGT		
mgs_cripsr_fwd	AAACTGAACGTGTTATTGATACATTACGCAGCTAG	Oligonucleotide used to form *mgs* cripsr insert	This project
mgs_cripsr_rev	AAAACTAGCTGCGTAATGTATCAATAACACGTTCA		
mgs_amp_fwd	ACTATTAGTACTTACACATTTATTTGGTTTGGTCATGTT	Primer used to amplify *mgs* gBlock	This project
mgs_amp_rev	CTGATCAGTACTCCGACCATGATCGTCAG		
mgs_check_fwd	TCAAACTGTCTCCGGATCAA	Primer used to check for successful KO of *mgs*	This project
mgs_check_rev	ATTCTTTGAATAAAGGCAGTACTTGA		
vanX_cripsr_fwd	AAACCTAAGCGTGCTGTAAACTGTTTTATGCAATG	Oligonucleotide used to form *vanX* cripsr insert	This project
vanX_cripsr_rev	AAAACATTGCATAAAACAGTTTACAGCACGCTTAG		
vanX_amp_fwd	ACTATTAGTACTGTGTGAAAAAAGTCAATAGCGC	Primer used to amplify *vanX* gBlock	This project
vanX_amp_rev	CTGATCAGTACTTCAAGGGACAGAACATCG		
vanX_check_fwd	GTGGCAGCTACGTTTACC	Primer used to check for successful KO of *vanX*	This project
vanX_check_rev	GTTGTAGTCACAGATAAAGCCC		
avrA_crispr_fwd	AAACCAACTGACAGCTCAAGCGAAACAAGTAATAG	Oligonucleotide used to form *avrA* cripsr insert	This project
avrA_crispr_rev	AAAACTATTACTTGTTTCGCTTGAGCTGTCAGTTG		
avrA_amp_fwd	ACTATTAGTACTGAAAAAGAAGGTGGTGCTC	Primer used to amplify *avrA* gBlock	This project
avrA_amp_rev	CTGATCAGTACTAAGGATACAATGTATTTCTATCGACC		
avrA_check_fwd	TCATGTTGCGCTAATCGTTG	Primer used to check for successful KO of *avrA*	This project
avrA_check_rev	GATGAACAATCCGACAGGTT		

Gene overexpression

avrA_OE_Fwd	GGATATCTGCAGTCAATGAAGGAAGAGATAAG TAGATGCGAAAAGTAGCACCTATTTT	Primers used to amplify the *avrA* gene, add a ribosome binding site and PstI site to the 5' end. A his-tag and glycine linker are added to the 3' end.	This project
avrA_OE_Rev	ATGATGATGATGATGATGTCCTCCTCCTCCTCCTCCTCCTTGAGAATCACCTGCTTGC		
avrB_OE_Fwd	GGATATCTGCAGCAGCAAGCAGGTGATTCTCA TAGATGACATTTTTACTAGACCTTATCAACG	Primers used to amplify the *avrB* gene, add a ribosome binding site and PstI site to the 5' end. A his-tag and glycine linker are added to the 3' end.	This project
avrB_OE_Rev	ATGATGATGATGATGATGTCCTCCTCCTCCTCCTCCTCCTGAATGTCTTCTTCTTTGGCTAC		
gshAB_OE_Fwd	GGATATCTGCAGTATGAATTGGAGGTCACTTA TAGATGATGAATTTTAAGCAATTATTATTGCATGTC	Primers used to amplify the *gshAB* gene, add a ribosome binding site and PstI site to the 5' end. A his-tag and glycine linker are added to the 3' end.	This project
gshAB_OE_Rev	ATGATGATGATGATGATGTCCTCCTCCTCCTCCTCCTCCAAATACTTCAGGATATAACAATTTCAATACATTC		

asnB_OE_Fwd	GGATATCTGCAGTATTAGGATGTGAAAGTTAA TAGATGTGCGGTATTGTAGGTTT	Primers used to amplify the *asnB* gene, add a ribosome binding site and PstI site to the 5' end. A his-tag and glycine linker are added to the 3' end.	This project
asnB_OE_Rev	ATGATGATGATGATGATGTCCTCCTCCTCCTCCTCCTCCAACAGTCACATAGTCGATTGC		
glpO_OE_Fwd	GGATATCTGCAGTGAATAAGGAGGAAATCCAT TAGATGTTTTCAAACAAGACAAGACAAGA	Primers used to amplify the *glpO* gene, add a ribosome binding site and PstI site to the 5' end. A his-tag and glycine linker are added to the 3' end.	This project
glpO_OE_Rev	ATGATGATGATGATGATGTCCTCCTCCTCCTCCTCCTCCTTTCTCCCCCTTTAGTTCC		
mgs_OE_Fwd	GGATATCCCGGGCTAGTTTGGAGGCGAAAAAC TAGATGAAAATCGGCTTTTTTACCGA	Primers used to amplify the *mgs* gene, add a ribosome binding site and SmaI site to the 5' end. A his-tag and glycine linker are added to the 3' end.	This project
mgs_OE_Rev	ATGATGATGATGATGATGTCCTCCTCCTCCTCCTCCTCCTCTTTTAAATGAATTTAATTTTATTTTGATTCGTTCAATC		
vanX_OE_Fwd	GGATATCTGCAGGCGTTAAAGGGGTGATAAGCTAGATGGAAATAGGATTTACTTTTTTAGATGAAATAGT	Primers used to amplify the *vanX* gene, add a ribosome binding site and PstI site to the 5' end. A his-tag and glycine linker are added to the 3' end.	This project
vanX_OE_Rev	ATGATGATGATGATGATGTCCTCCTCCTCCTCCTCCTCCTTTAACGGGGAAATCAAAATAGCT		
Tn-seq_OE_Rev	CGAATCGCATGCTTAATGATGATGATGATGATGTCCTC	Primer used to amplify the OE amplicons and add a SphI site at the 3' end.	This project

Inverse PCR

IPCR_HaeIII_F	AAACAGGAATTTATCGAAAATGGT	Primers used to determine the location of the mariner transposon insertion.	(X Zhang *et al.*, 2012)
IPCR_HaeIII_R	CCCCCTGAAATCCTTACAT		

General Check primers

vanA-F	AAGATTCCGTACTGCAGCCT	Primers to check for the presence of *vanA*	This project
vanA-R	GGATAGCTACTCCCGCCTTT		
Cas9_check_fwd	GGGCGGTGATCACTGATGAATATA	Primers to check for successful electroporation of pVPL3004	(de Maat *et al.*, 2019)
Cas9_check_rev	ACCAATAATTCCTCAGTACCATCCAT		
pWS3_check_fwd	TGTGCTGCAAGGCGATTAAGTTGG	Primers to check for presence of pVDM1001 and to check for insertions in pVDM1001.	(de Maat *et al.*, 2019)
pWS3_check_rev	AAAGCGTCCCTACTACTGACAGCTTC		

P = 5' Phosphorylation

2.6 Plasmids

Table 2.9: Plasmids used / created in this thesis.

Plasmid name	Description	Source
pVPL3004	Plasmid which expresses Cas9 and tracrRNA. Ery^r	(Oh and van Pijkeren, 2014)
pVDM1001	Plasmid for cloning the CRISPR sequence and Insertion/Knockout template of the gene of interest. Spec^r	(de Maat *et al.*, 2019)
pVDM1001-asnB-KO	Plasmid containing the *asnB* CRISPR sequence and the *asnB* knockout construct. Spec^r	This project
pVDM1001-avrA-KO	Plasmid containing the *avrA* CRISPR sequence and the *avrA* knockout construct. Spec^r	This project
pVDM1001-avrB-KO	Plasmid containing the *avrB* CRISPR sequence and the *avrB* knockout construct. Spec^r	This project
pVDM1001-glpO-KO	Plasmid containing the *glpO* CRISPR sequence and the *glpO* knockout construct. Spec^r	This project
pVDM1001-gshAB-KO	Plasmid containing the *gshAB* CRISPR sequence and the *gshAB* knockout construct. Spec^r	This project
pVDM1001-mgs-KO	Plasmid containing the *mgs* CRISPR sequence and the *mgs* knockout construct. Spec^r	This project
pMSP3535	*E. faecium* overexpression vector. Ery^r	(Bryan *et al.*, 2000)
pMSP3535-asnB-OE	*E. faecium* overexpression vector carrying the *asnB* gene. Ery^r	This project
pMSP3535-avrA-OE	*E. faecium* overexpression vector carrying the *avrA* gene. Ery^r	This project
pMSP3535-avrB-OE	*E. faecium* overexpression vector carrying the *avrB* gene. Ery^r	This project
pMSP3535-glpO-OE	*E. faecium* overexpression vector carrying the *glpO* gene. Ery^r	This project
pMSP3535-gshAB-OE	*E. faecium* overexpression vector carrying the *gshAB* gene. Ery^r	This project
pMSP3535-mgs-OE	*E. faecium* overexpression vector carrying the *mgs* gene. Ery^r	This project
pMSP3535-vanX-OE	*E. faecium* overexpression vector carrying the *vanX* gene. Ery^r	This project

2.7 Bacterial isolates

Table 2.10: Bacterial isolates used / created in this thesis.

Name	Sample-source	Phenotypic vancomycin resistance	Citation
E. faecium isolates			
E. faecium E745	Rectal swab	Resistant	(Zhang *et al.*, 2017)
E. faecium E8202	Hospitalised patient	Resistant	(Arredondo-Alonso *et al.*, 2020)
E. faecium Aus0004	Rectal swab	Resistant	(Lam *et al.*, 2012)
E. faecium 64/3	Rectal swab	Susceptible	(Bender *et al.*, 2015)
E. faecium OI1	Blood culture	Resistant	This project
E. faecium OI2	Blood culture	Resistant	This project
E. faecium OI3	Blood culture	Resistant	This project
E. faecium OI4	Blood culture	Resistant	This project
E. faecium OI5	Blood culture	Resistant	This project
E. faecium OI6	Blood culture	Resistant	This project
E. faecium OI7	Blood culture	Resistant	This project
E. faecium OI8	Blood culture	Resistant	This project
E. faecium OI9	Blood culture	Resistant	This project
E. faecium OI10	Blood culture	Resistant	This project
E. faecium OI11	Blood culture	Resistant	This project
E. faecium OI12	Blood culture	Resistant	This project
E. faecium OI13	Blood culture	Resistant	This project
E. faecium OI14	Blood culture	Resistant	This project
E. faecium OI15	Blood culture	Resistant	This project
E. faecium OI16	Blood culture	Resistant	This project
E. faecium OI17	Blood culture	Susceptible	This project
E. faecium OI18	Blood culture	Susceptible	This project
E. faecium OI19	Blood culture	Resistant	This project
E. faecium OI20	Blood culture	Susceptible	This project
E. faecium OI21	Rectal swab	Resistant	This project
E. faecium OI22	Blood culture	Susceptible	This project
E. faecium OI23	Rectal swab	Resistant	This project
E. faecium OI24	Blood culture	Susceptible	This project
E. faecium OI25	Blood culture	Susceptible	This project
E. faecium OI25 revertant 1	Blood culture / Laboratory	Resistant	This project
E. faecium OI25 revertant 2	Blood culture / Laboratory	Resistant	This project
E. faecium OI26	Rectal swab	Resistant	This project
E. faecium OI27	Blood culture	Susceptible	This project
E. faecium OI28	Blood culture	Susceptible	This project
E. faecium OI29	Rectal swab	Resistant	This project
E. faecium OI30	Blood culture	Susceptible	This project
E. faecium OI31	Blood culture	Susceptible	This project
E. faecium OI32	Blood culture	Resistant	This project
E. faecium OI33	Blood culture	Susceptible	This project
E. faecium OI34	Rectal swab	Resistant	This project
E. faecium OI35	Blood culture	Resistant	This project
E. faecium OI36	Blood culture	Resistant	This project
E. faecium OI37	Blood culture	Susceptible	This project
E. faecium OI38	Blood culture	Resistant	This project
E. faecium OI39	Blood culture	Susceptible	This project
E. faecium HHI1	Blood culture	Susceptible	This project
E. faecium HHI2	Blood culture	Susceptible	This project
E. faecium HHI2 revertant 1	Blood culture / laboratory	Resistant	This project
E. faecium HHI2 revertant 2	Blood culture / laboratory	Resistant	This project
E. faecium HHI3	Blood culture	Susceptible	This project

E. faecium HHI3 revertant 1	Blood culture / laboratory	Resistant	This project
E. faecium HHI3 revertant 2	Blood culture / laboratory	Resistant	This project
E. faecium HHI4	Rectal swab	Resistant	This project
E. faecium HHI5	Rectal swab	Resistant	This project
E. faecium HHI6	Rectal swab	Resistant	This project
E. faecium HHI7	Urine	Resistant	This project
E. faecium HHI8	Rectal swab	Resistant	This project
E. faecium HHI9	Rectal swab	Resistant	This project
E. faecium HHI10	Rectal swab	Resistant	This project
E. faecium HHI11	Rectal swab	Resistant	This project
E. faecium HHI12	Rectal swab	Resistant	This project
E. faecium HHI13	Rectal swab	Resistant	This project
E. faecium HHI14	Rectal swab	Resistant	This project
E. faecium HHI15	Rectal swab	Resistant	This project
E. faecium HHI16	Rectal swab	Resistant	This project
E. faecium HHI17	Rectal swab	Resistant	This project
E. faecium HHI18	Rectal swab	Resistant	This project
E. faecium HHI19	Rectal swab	Resistant	This project
E. faecium HHI20	Rectal swab	Resistant	This project
E. faecium HHI21	Rectal swab	Resistant	This project
E. faecium E745 Δ*avrA*	Laboratory	Susceptible	This project
E. faecium E745 Δ*avrB*	Laboratory	Resistant	This project
E. faecium E745 Δ*asnB*	Laboratory	Susceptible	This project
E. faecium E745 Δ*glpo*	Laboratory	Resistant	This project
E. faecium E745 Δ*gshAB*	Laboratory	Susceptible	This project

E. coli isolates

E. coli DH5α	Laboratory	NA	(Taylor, Walker and McInnes, 1993)
E. coli EC1000	Laboratory	NA	(Leenhouts *et al.*, 1996)

Chapter 3
Metagenome-Wide Analysis of Rural and Urban Surface Waters and Sediments in Bangladesh Identifies Human Waste as a Driver of Antibiotic Resistance

3.1 Introduction

The prevalence of antibiotic-resistant bacteria causing infections is increasing globally, but the clinical issues, including significant morbidity and mortality, posed by these bacteria are particularly alarming in low- and middle-income countries (LMICs) (Laxminarayan *et al.*, 2013; Lim *et al.*, 2016; Founou, Founou and Essack, 2017; Gandra *et al.*, 2019). Proposed drivers for the high burden of drug-resistant infections in LMICs include the unregulated sales of antibiotics and their misuse in clinical medicine, agriculture, and aquaculture; an inadequate sewerage infrastructure; poor governance; and low investments in health care (Collignon *et al.*, 2018; Chokshi *et al.*, 2019).

One of the challenges of studying antimicrobial resistance (AMR) is to disentangle the spread of resistant bacteria and antibiotic resistance genes between humans, animals, and the wider environment (Woolhouse *et al.*, 2015). For this reason, AMR is increasingly being studied from a collaborative and cross-disciplinary perspective that has been termed "One Health" (McEwen and Collignon, 2018). The One Health concept for studying the spread of AMR is particularly relevant for LMICs due to the crucially important role of agriculture and aquaculture in the livelihoods of billions of people in many of these countries, especially the poorest ones (Robinson *et al.*, 2016). Asia is home to an estimated 74% of the world's 570 million farms (Lowder, Skoet and Raney, 2016), and, in 2016, 89% of the global aquaculture production was estimated to originate from this continent (Garlock *et al.*, 2020). However, there are still major knowledge gaps on the spread of AMR in Asia from a One Health perspective.

Bangladesh is an LMIC in South Asia, where antibiotic-resistant infections are common among both hospitalized patients and the non-hospitalized population (Ahmed, Rabbi and

Sultana, 2019). The country has a number of unique characteristics that may contribute to the rapid spread of AMR. The capital city of Bangladesh, Dhaka, has a population of around 16 million people, with a population density that ranks among the highest of any megacity. Less than 20% of the households in Dhaka are directly connected to sewerage infrastructure (Peal *et al.*, 2014). The prevalence of carriage of multidrug-resistant *Escherichia coli* among healthy humans is relatively high in Bangladesh, as it is in other LMICs (Mamun, Shears and Hart, 1993; Monira *et al.*, 2017; Nji *et al.*, 2021).

Antibiotic-resistant bacteria that colonize the human gut can be passed into rivers, lakes, and coastal areas through the release of untreated wastewater, the overflow of pit latrines during monsoon season, or by practices such as open defecation (Akter, Kurisu and Hanaki, 2017; Hussain *et al.*, 2017). These contaminated environments are often used for bathing and the washing of clothes and food preparation equipment, thus facilitating human gut colonization by antibiotic-resistant bacteria (Amin *et al.*, 2019). The antibiotic resistance genes carried by human-associated bacteria are often found on mobile genetic elements, including plasmids, which are capable of transfer within and between bacterial species. Plasmid-mediated transfer of antibiotic resistance genes thus contributes to the rapid dissemination of antibiotic resistance genes in microbial ecosystems, including those in the human gut and the environment (Martínez, 2008; Li *et al.*, 2019).

While a prescription is legally required to purchase antibiotics in Bangladesh, antibiotics can be readily acquired from many of the 200,000 drug stores across Bangladesh (Rousham *et al.*, 2019). In rural Bangladesh, aquaculture is widespread, with more than 2 million tons of freshwater fish produced in 2017 from inland freshwater fisheries (*FAO yearbook. Fishery and Aquaculture Statistics 2016*, 2018). A survey performed from 2011 to 2012 revealed

that antibiotics are widely used in Bangladeshi aquaculture for disease prevention and growth promotion. The most prominent classes of antibiotics employed are the tetracyclines, but other antibiotic classes, including β-lactams and sulphonamides, are also used (Ali *et al.*, 2016). The use of antibiotics in Bangladesh is regulated in line with the European Union standards for antibiotic use in aquaculture, but Bangladesh has been found to be in breach of these regulations several times (Lulijwa, Rupia and Alfaro, 2020). The causes of antibiotics overuse in aquaculture are multifactorial: pharmaceutical companies provide food that is premixed with antibiotics without the farmers' knowledge, farmers administer antibiotics too often because they do not understand the instructions, and prophylactic use of antibiotics may be used to reduce the chance of damaging losses in production caused by disease (Kawsar *et al.*, 2018). The combination of a densely populated country, intensive antibiotic usage in aquaculture, and the potential for the dissemination of antibiotic-resistant bacteria through surface water thus provides a unique opportunity to study the spread of AMR from a One Health perspective in Bangladesh.

In this chapter, I use a combination of quantitative bacterial culture and metagenomic shotgun sequencing methods to disentangle pathways that contribute to the dissemination of antibiotic resistance. Specifically, we describe the abundance and diversity of microorganisms and antibiotic resistance genes in surface water in rural and urban settings in Bangladesh.

3.2 Hypothesis

The use of antibiotics in rural aquaculture is a major driver of antibiotic resistance in Bangladesh.

3.3 Aims and objectives

1. To quantify the number of ESBL-producing coliforms present in sediment and water samples in Bangladesh.
2. To identify and quantify the bacteria present in water and sediment samples in Bangladesh using metagenomic sequencing.
3. To identify and quantify the antibiotic resistance genes present in these samples.
4. To correlate the abundance of antibiotic resistance genes with the level of gut associated bacteria in these samples.
5. To identify plasmids, particularly those carrying antibiotic resistance genes, in these samples.

3.4 Results

3.4.1 Sample collection across urban and rural sites in Bangladesh.

Freshwater surface water and sediment samples were collected from 24 sites across 3 districts in Bangladesh (Mymensingh, Shariatpur, and Dhaka; Figure 3.1). These sites spanned both rural and urban areas with different population densities. Among rural sites, ponds used for aquaculture with a history of antibiotic use (n = 11) and ponds with no history of antibiotic use (n = 6) were sampled. Further information on sampling locations and protocols is provided in Materials and Methods. We used culture-dependent and culture-independent methods to study the abundance of antibiotic resistance genes and the diversity of microbiotas across the different sites.

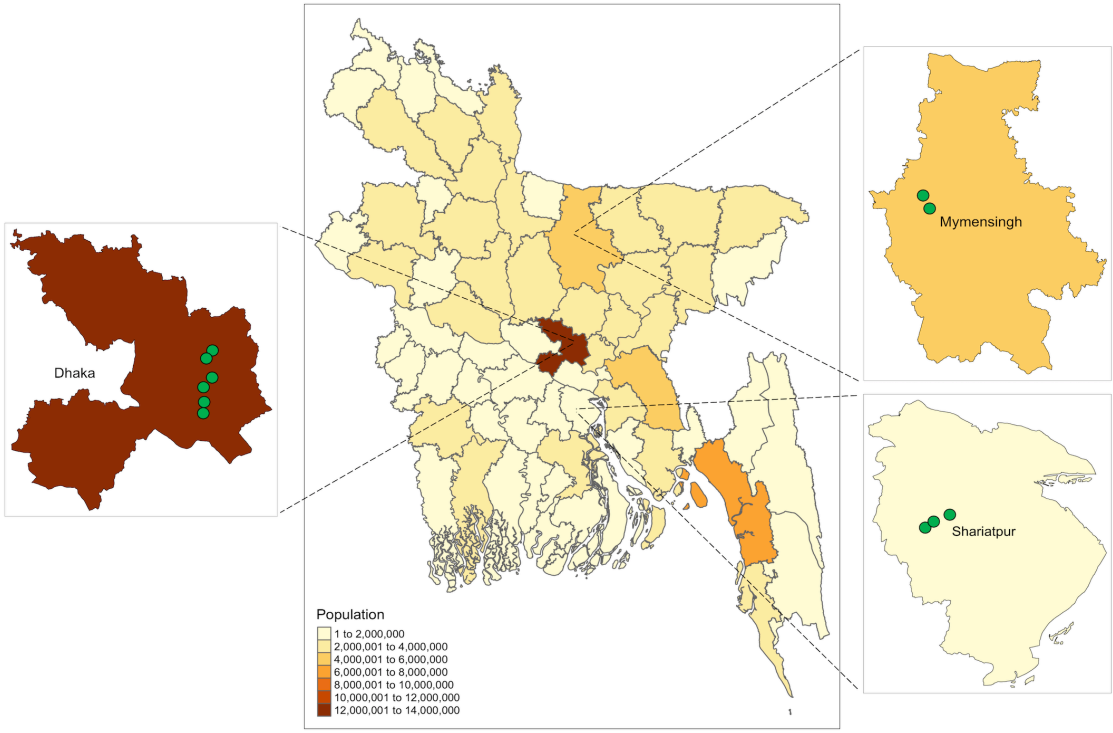

Figure 3.1: Map of Bangladesh showing the districts that the samples were collected from and the population of each district (obtained through https://data.humdata.org/dataset/bangladesh-administrative-level-0-3-population-statistics). Green circles represent sampling locations.

3.4.2 ESBL-producing coliforms were more prevalent in urban samples than in rural samples

We quantitatively determined the burden of extended-spectrum beta-lactamase (ESBL)-producing coliforms in the water and sediment samples from the different sampling locations and found that ESBL-producing coliforms were detected in significantly more urban samples (12/14) than rural samples (15/34) (Fisher exact test; $P = 0.01$). However, there was no statistically significant difference in the viable counts of ESBL-producing coliforms in urban and rural samples (Figure 3.2).

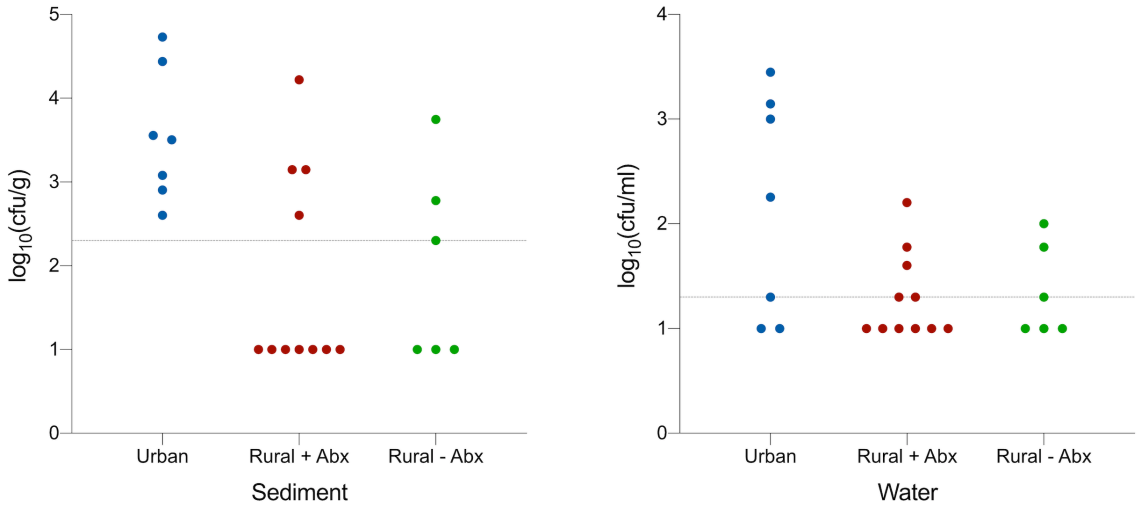

Figure 3.2: The abundance of ESBL-producing coliforms isolated from sediment $\log_{10}$(CFU/g) and surface water $\log_{10}$(CFU/ml) in urban sites and rural settings with antibiotic use (+Abx) and without antibiotic use (-Abx) across Bangladesh. The horizontal dashed lines represent the detection limit of 200 CFU/g for sediment and 20 CFU/ml for water. Samples with ESBL-producing coliforms below the detection limit were plotted at $\log_{10}$(CFU/ml) of 1.

3.4.3 Microbiotas of surface water and sediments are distinct, with higher levels of human gut bacteria in urban samples.

Shotgun metagenomic sequencing was used to study the diversity and composition of the microbial communities in the different samples. An important determinant shaping the communities was the sample type, with distinct (permutational multivariate analysis of variance ([PERMANOVA], $P < 0.001$) clustering of sediment and water samples (Figure 3.3).

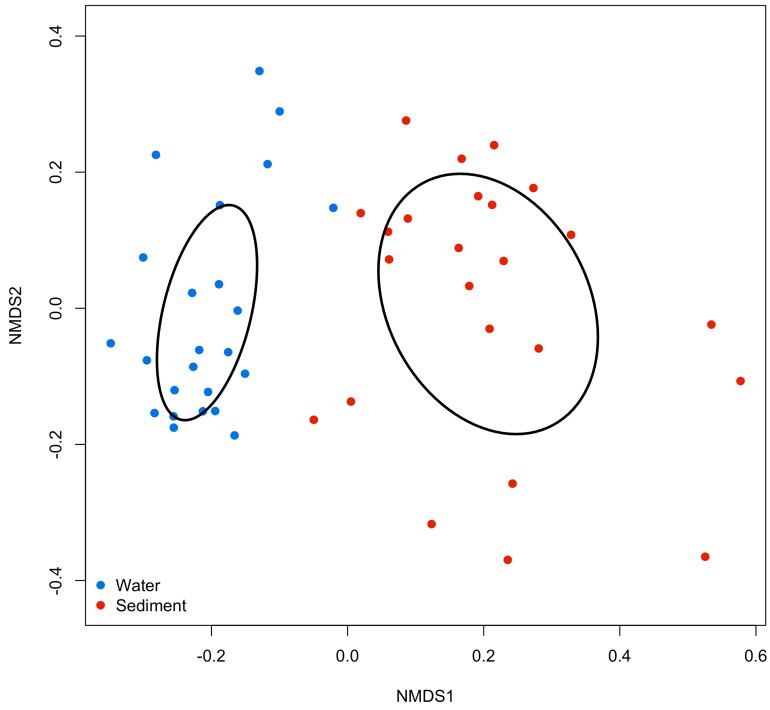

Figure 3.3: Nonmetric multidimensional scaling (NMDS) analysis of a Bray-Curtis distance matrix of species abundance. Stress, 0.15. Ellipses represent standard deviation.

The sediment samples were dominated by the phylum *Proteobacteria* (73.8%; standard deviation [SD], 27.1), while in the water samples, *Cyanobacteria* (60.9%; SD, 29.6) was the dominant phylum (Figure 3.4). However, among the different sample types (water and sediment), these phyla were not inevitably dominant, as in five of the nine sediment samples collected in Mymensingh, the abundance of *Euryarchaeota* was greater than 50%, while in five Dhaka water samples, *Proteobacteria* were present at levels greater than 45%. Water

sample WAM6 had very high levels (>60%) of bacteriophage DNA. The sediment samples were dominated by typical soil bacteria such as *Pseudomonas*, *Azoarcus*, and *Anaeromyxobacter*, while the water samples were dominated by cyanobacteria such as *Cyanobium*, *Microcystis*, and other typical aquatic bacterial species from the phyla *Proteobacteria* and *Actinobacteria* (Figure 3.6). Three bacteriophages (*Mycobacterium* phage Rizal, *Microcystis aeruginosa* phage Ma LMM01, and an Epsilon15-like virus) were also identified at different sampling sites. It was apparent that three of the Dhaka water samples contained bacteria which are typically found within the gastrointestinal tract, including *Escherichia coli*, *Streptococcus infantarius*, *Bifidobacterium adolescentis*, and *Prevotella copri*.

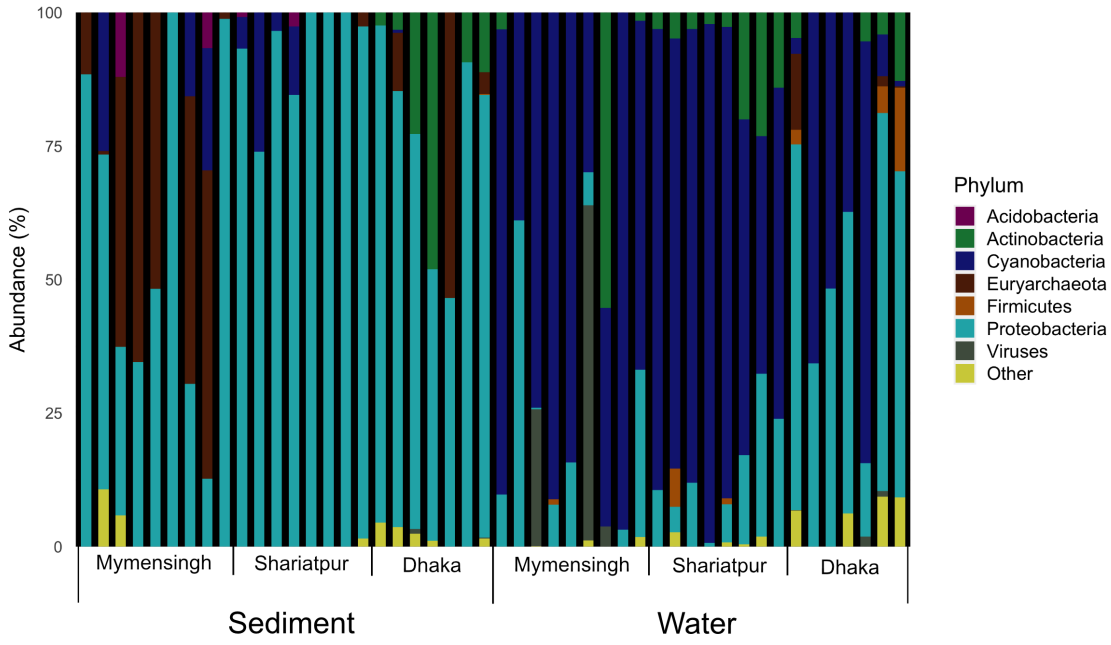

Figure 3.4: Relative abundance (%) of phyla across the 48 samples from sediment and surface water.

Through microbial source-tracking analysis of our shotgun sequencing data using the FEAST algorithm (Shenhav *et al.*, 2019), we found that the urban water samples had a significantly greater contribution from gut bacteria (Kruskal-Wallis, $P < 0.01$) than the rural samples without previous antibiotic use (Figure 3.5).

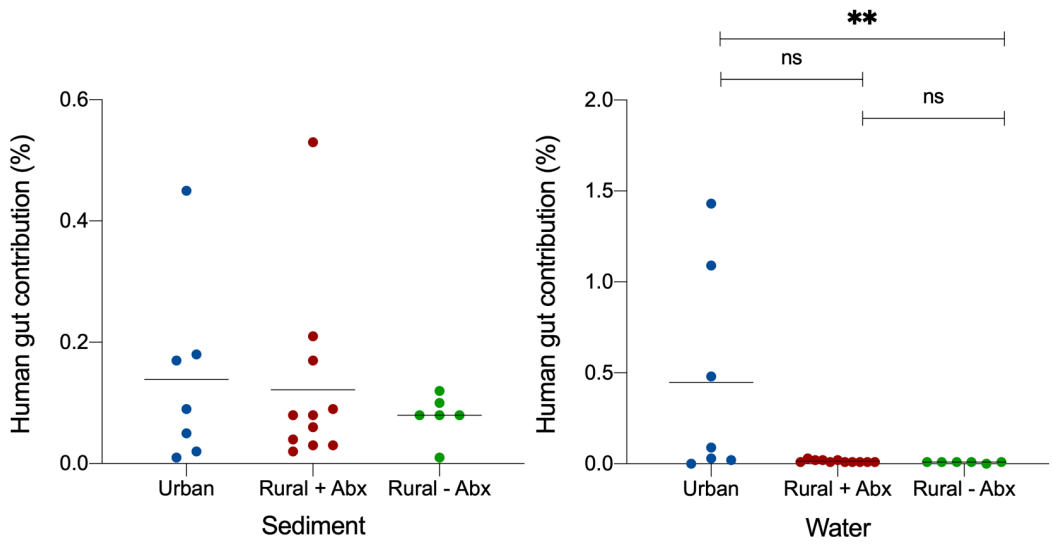

Figure 3.5: Source-sink analysis, percentage contribution of human gut bacteria to the bacterial composition of the water and sediments samples. Kruskal-Wallis test; ** $P < 0.01$.

Based on species-level data (Figure 3.6), the urban sediment samples were significantly more diverse than both the rural samples with and without previous antibiotic use (Browne-Forsythe and Welch, $P < 0.05$). There was no significant difference in diversity between either of the rural sediment sample types (Figure 3.7). On the other hand, the rural water samples without previous antibiotic use were significantly more diverse than the rural samples with previous antibiotic use (Browne-Forsythe and Welch, $P < 0.005$), but there was no significant difference between the urban water samples and either of the rural sample types.

Figure 3.6: Heatmap of species-level abundances for surface water ($n = 24$) and sediment ($n = 24$) samples. Species-level abundance was determined by the analysis of shotgun metagenomic sequencing data by the MetaPhlAn2 package. Colours reflect the relative abundance of each individual taxon. The samples and species are clustered by Bray-Curtis dissimilarity.

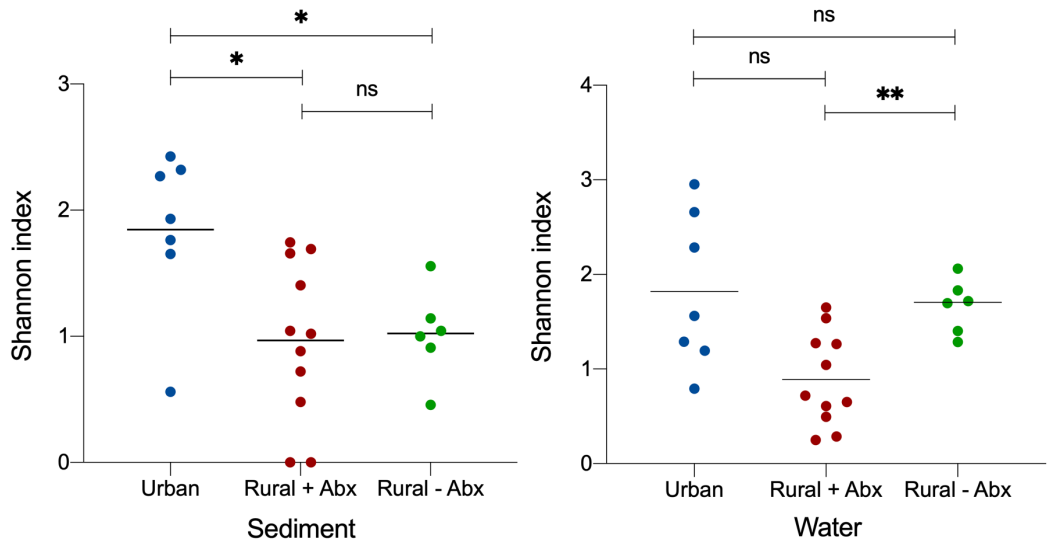

Figure 3.7: Shannon diversity values of species present in sediment and water samples from across Bangladesh. Brown-Forsythe ANOVA. * $P < 0.05$; ** $P < 0.005$; ns, not significant ($P > 0.05$).

3.4.4 Urban samples carry the highest antibiotic resistance gene loads.

A total of 114 different antibiotic resistance genes (ARGs) that confer resistance to 16 antibiotic classes were identified in the 48 samples from sediment and surface water. The urban samples had the greatest number of ARGs (n = 99) followed by the rural samples with previous antibiotic use (n = 49), while the rural samples with no previous antibiotic use had the fewest resistance genes (n = 36) (Figure 3.8). There was a large overlap between the ARGs present in the different sample types, with the urban and rural positive antibiotic samples sharing the greatest number of resistance genes (n = 24). There were 17 ARGs shared between all three sample types, including five different beta-lactamase genes belonging to the bla_{OXA} and bla_{RSA} families.

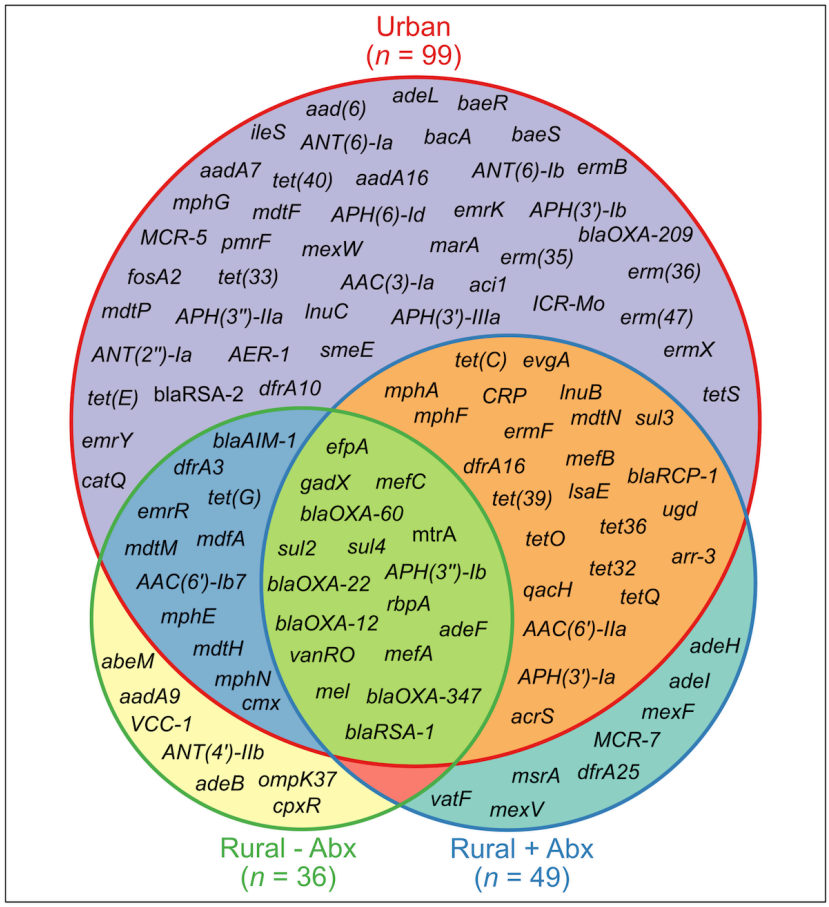

Figure 3.8: Distribution of antibiotic resistance genes across urban, rural without prior antibiotic use, and rural with prior antibiotic use sample types. Circles are proportional to the number of antibiotic resistance genes (*n*) present within each sample type.

The abundance of antibiotic resistance genes varied 1,525-fold between sites, with sample SAM6 (rural sediment sample with previous antibiotic exposure collected in Mymensingh) having the lowest abundance (0.078 reads per kilobase of reference sequence per million sample reads [RPKM]) and sample WD7 (surface water sample collected in Dhaka) having the highest ARG abundance (120.45 RPKM). Of the paired sediment and water samples, the ARG abundance was, on average, 3 times greater in the water samples than the sediment samples (Wilcoxon, $P < 0.0001$). The urban sediment samples collected from around the city of Dhaka were found to have a significantly greater (Kruskal-Wallis, $P < 0.05$) total ARG abundance (median RPKM, 4.01; interquartile range [IQR], 0.95 to 12.79) than the

rural samples with prior antibiotic use (median RPKM, 0.60; IQR, 0.20 to 1.27) (Figure 3.9). However, the urban sediment samples were not significantly different from the rural samples without antibiotic use (median RPKM, 0.72; IQR, 0.64 to 1.36). There was also no statistically significant difference (Kruskal-Wallis, $P < 0.99$) between ARG abundance in rural sediment with prior antibiotic use versus sediment from rural sites in which antibiotics had not been used. ARG levels in the water samples reflected that of the sediment samples, with the total ARG abundance in urban samples (median RPKM, 37.08; IQR, 5.71 to 97.74) being significantly higher (Kruskal-Wallis, $P < 0.05$) than the rural samples with previous antibiotic use (median RPKM, 4.30; IQR, 2.39 to 7.60) but not significantly different from the rural samples with no previous antibiotic use (median RPKM, 5.09; IQR, 1.80 to 11.68). As with the sediment samples, there was no significant difference found between either of the rural sample types (Kruskal-Wallis, $P > 0.99$).

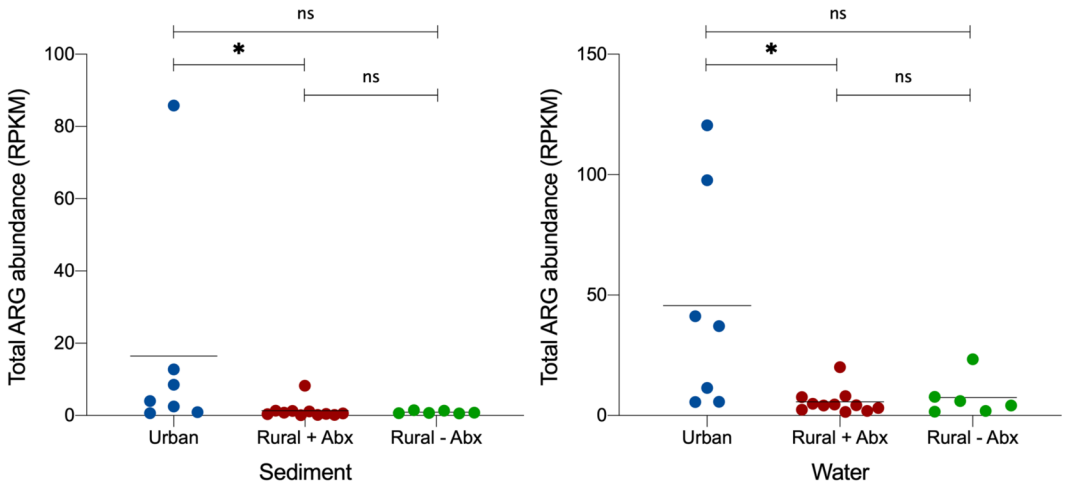

Figure 3.9: Abundance in reads per kilobase of reference sequence per million sample reads (RPKM) of antibiotic resistance genes (ARGs) in each sample (sediments and surface water; urban, rural with antibiotic use, and rural without antibiotic use). Kruskal-Wallis test, *, $P < 0.05$.

The individual antibiotic resistance genes were collated into 16 classes that cover resistance to specific antibiotics and a separate class for genes conferring antibiotic efflux mechanisms

(Figure 3.10). Efflux genes were present in 47 of 48 samples, making it the most widespread ARG class. Other abundant antibiotic resistance classes were resistant to sulphonamides, macrolides, and aminoglycosides. Urban water samples WD2, WD6, WD7, and WD1 and an urban sediment sample SD7 clustered together, with high levels of resistance genes from these classes.

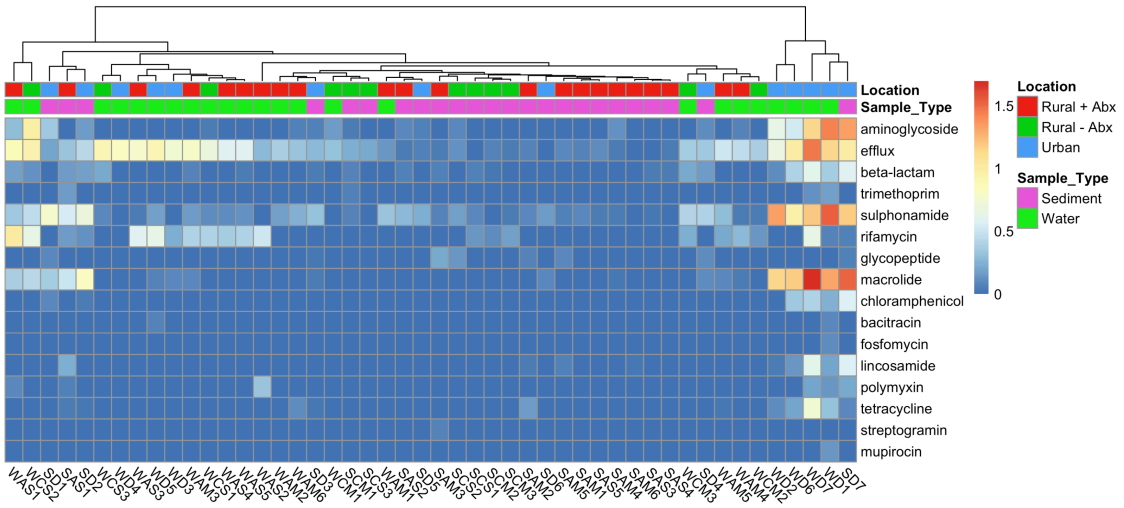

Figure 3.10: Heatmap representing the summed abundance (log_{10}-transformed RPKM) of antibiotic resistance gene classes present in water and sediment samples from surface water sites across Bangladesh.

3.4.5 Abundance of human gut bacteria predicts levels of antibiotic resistance genes.

There was a statistically significant correlation (R^2 = 0.73; P = 8.9x10^{-15}) between the aggregated abundance of ARGs and the levels of human gut bacteria, as identified by FEAST, across our study (Figure 3.11A). We also determined whether the levels of ESBL-producing coliforms are correlated with the total abundance of ARGs and observed a relatively weak but statistically significant correlation (R^2 = 0.12; P = 0.009) (Figure 3.11B).

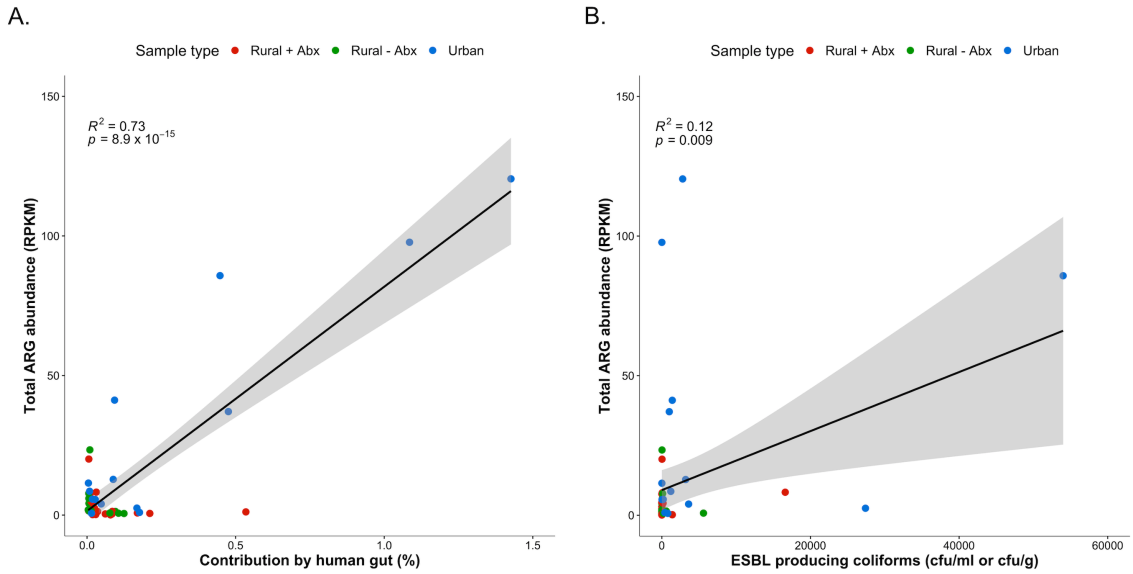

Figure 3.11: Correlation between total ARG abundance and other factors. (A) Correlation between the total antibiotic resistance gene (ARG) abundance (reads per kilobase of reference sequence per million sample reads [RPKM]) and the percentage of bacteria contributed from the human gut within each sample. $R^2 = 0.73$; $P = 8.9 \times 10^{-15}$. (B) Correlation between the total ARG abundance (RPKM) and the number of ESBL-producing coliforms (CFU/ml) in each sample. $R^2 = 0.12$; $P = 0.009$. Gray area represents the 95% confidence interval.

3.4.6 Urban sites were enriched in plasmids carrying antibiotic resistance genes.

As antibiotic resistance genes were particularly abundant in water samples, we performed a metagenomic assembly of the short-read data from the surface water samples to recover complete plasmid sequences and study their potential association with antibiotic resistance. The metagenomic assemblies were queried against the PlasmidFinder database (Carattoli *et al.*, 2014) to identify contigs which contained plasmid replication (*rep*) genes. Eleven contigs in our data set contained rep genes (Table 3.1). Seven Gram-negative replicons were found, which were related to representatives of the P and Q incompatibility groups or to small theta- or rolling circle-replicating plasmids. A single Gram-positive replicon, *repUS43*, was identified in sample WD1.

Table 3.1: Features of plasmid contigs.

Sample	Contig	Rep type	Length (bp)	Coverage (%)[+]	Identity (%)[+]
WCM1	k141_156157	Col156	5212	100	94.81
	k141_206349*	Col(BS512)	2254	100	100
	k141_35625	Col8282	3712	100	80.88
WD1	k141_304072*	IncQ1	8676	100	100
	k141_320207	IncQ	1420	51.33	88.31
	k141_593572	repUS43	1262	50.41	96.05
	k141_711213	IncQ1	24244	78.39	77.48
	k141_728573	Col(pWES)	1941	93.26	80.95
WD2	k141_452869	IncQ	2296	78.39	77.48
WD7	k141_315908	IncP6	1166	100	99.63
	k141_77466	IncQ	4072	78.39	77.48

* Plasmid contig which could be circularised.
[+] Coverage and identity are of the closest *rep* gene in the PlasmidFinder database.

Two plasmid contigs, k141_206349 (2,113 bp) and k141_304072 (8,535 bp), could be circularized. The latter plasmid, which we named pWD1, contained the sulphonamide resistance gene *sul2* (Figure 3.12).

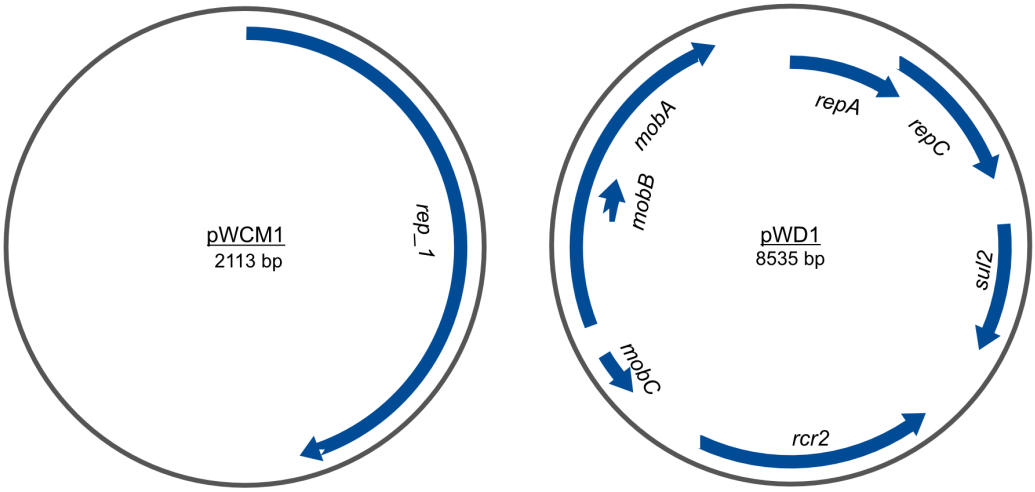

Figure 3.12: Plasmid maps of pWCM1 and pWD1 which were circularised from contigs k141_206349 and k141_304072.

pWD1 was found to have 99.97% identity over 81% of its sequence to the canonical broad-host-range mobilizable plasmid RSF1010 (Scholz *et al.*, 1989) (Figure 3.13).

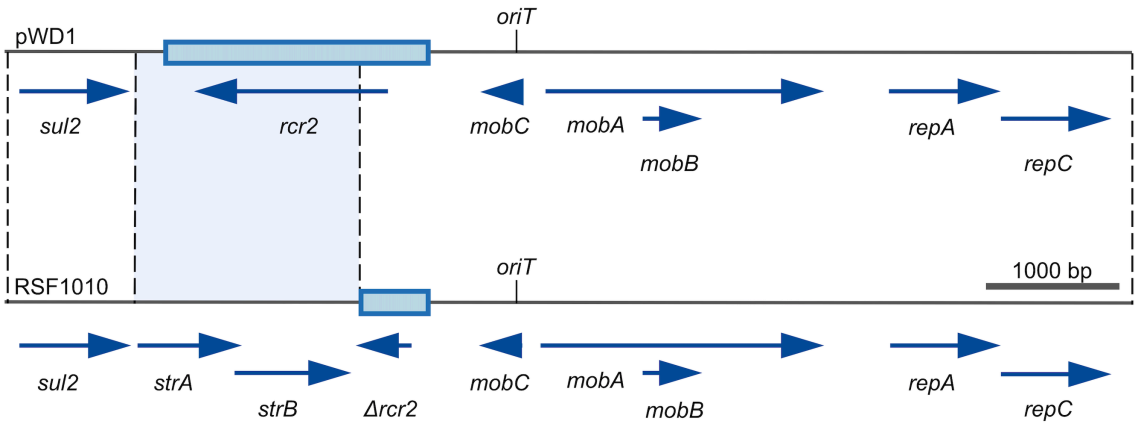

Figure 3.13: Comparison of plasmids pWD1 and RSF1010. Plasmid sequence is shown as a black line with the positions of genes indicated by labelled arrows below and the location of *oriT* shown above. The mobile element CR2 is shown as a thicker blue box. The light-blue shading highlights the region that differs between the plasmids and includes the *strAB* genes in RSF1010. Drawn to scale from GenBank accession nos. MW363525 and M28829 for pWD1 and RSF1010, respectively.

As metagenomic assemblies are often fragmented and plasmid replication genes may not be on the same contigs as ARGs that are carried on another region of the plasmid, we employed PlasFlow (Krawczyk, Lipinski and Dziembowski, 2018) to classify contigs in our metagenomic assembly as either chromosomal or plasmid. We identified a total of 93 plasmid contigs containing ARGs. The urban sediment samples contained significantly more plasmid contigs with ARGs than either of the rural sample types (Kruskal-Wallis, $P < 0.001$), whereas the urban water samples had significantly more ARG-bearing plasmid contigs than the rural samples with no previous antibiotic use (Kruskal-Wallis, $P < 0.05$) (Figure 3.14). There was no significant difference in the number of ARG-containing plasmid contigs between rural samples with and without prior antibiotic use.

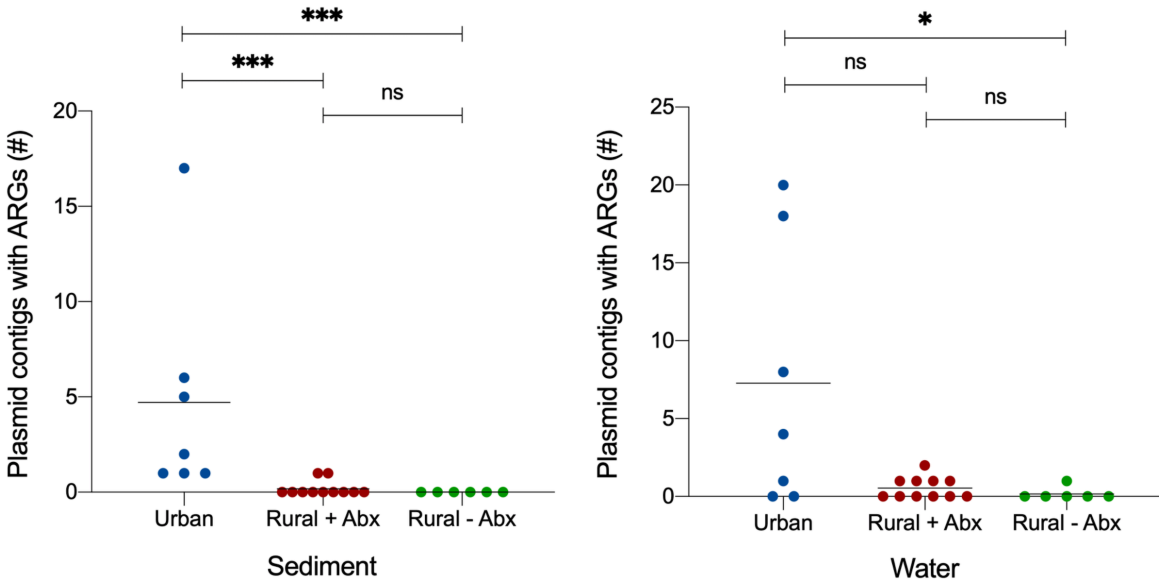

Figure 3.14: Number of plasmid contigs, as identified by PlasFlow, carrying antibiotic resistance genes in urban and rural (with and without antibiotic use [Abx]) surface water and sediment samples. Kruskal Wallis test; *, $P < 0.05$; ***, $P < 0.001$.

Of the 93 contigs identified which contained ARGs, 78 contigs contained only 1 resistance gene, with the remaining 15 contigs containing 2 or more ARGs (Table 3.2). All of the contigs that contained multiple resistance genes were found in urban samples and were closely related to known proteobacterial plasmids.

Table 3.2: Closest matches in the blastn database to the contigs of plasmid origin carrying antibiotic resistance genes.

Sample	Contig	Closest match	Bacterium	Identity (%)	Coverage (%)
WCM3	k141_292096	pPm14C18	*Proteus mirabilis*	99.88	100.00
SD1	k141_106889	pSCU-397-2	*Escherichia coli*	99.11	99.00
	k141_227167	pAN70-1	*Alcaligenes faecalis*	99.96	100.00
	k141_704017*	pNFYY023-1	*Comamonas testosteroni*	99.94	91.00
	k141_836485	pG5A4Y217	*Escherichia coli*	99.34	58.00
	k141_99417	pKP14812-MCR1	*Klebsiella pneumoniae*	97.57	40.00
WD1	k141_128693	p33	*Escherichia coli*	100.00	76.00
	k141_134061	Unnamed plasmid	*Butyricimonas faecalis*	89.14	28.00
	k141_139895	pCP8-3-IncFIB	*Escherichia coli*	95.43	99.00
	k141_160447*	pCAV1335-92	*Klebsiella oxytoca*	98.76	100.00
	k141_205613	p1	*Klebsiella pneumoniae*	99.76	90.00
	k141_256831*	pGENC284	*Enterobacter hormaechei* subsp. X*iangfangensis*	100.00	100.00
	k141_304072	pSL7202-3	*Salmonella enterica* subsp. *enterica* serovar *Typhimurium*	99.98	81.00
	k141_344452	pRWC72a	Uncultured bacterium	98.60	100.00
	k141_377411*	pUCLAOXA232-5	*Klebsiella pneumoniae*	100.00	41.00
	k141_391604	pTZC1	*Cutibacterium acnes*	98.57	91.00
	k141_44508	pJF-786	*Enterobacter cloacae*	99.83	86.00
	k141_467424*	pYH12207-3	*Acinetobacter piscicola*	100.00	92.00
	k141_510896	pG5A4Y217	*Escherichia coli*	99.46	51.00
	k141_604491	pSTN0717-64-1	*Enterobacter hormaechei*	99.75	100.00
	k141_719363	pSAN1-06-0624	*Salmonella enterica* subsp. *enterica* serovar *Anatum*	100.00	100.00

	k141_719904	pAeme6	*Aeromonas media*	99.77	100.00
	k141_804869	pMRGN207	*Escherichia coli*	99.59	70.00
	k141_881861	pBS228	*Pseudomonas aeruginosa*	99.94	92.00
	k141_903595	pWCX23_1	*Aeromonas hydrophila*	99.87	100.00
	k141_91069	pPN3F2_1	*Shewanella aestuarii*	99.23	68.00
SD2	k141_324783	pEC422_1	*Escherichia coli*	99.91	81.00
	k141_325504	Unnamed plasmid	*Klebsiella michiganensis*	100.00	33.00
	k141_343850	RW109	*Pseudomonas aeruginosa*	99.92	58.00
	k141_461478*	pAN70-1	*Alcaligenes faecalis*	100.00	100.00
	k141_701410	pGENC284	*Enterobacter hormaechei* subsp. *Xiangfangensis*	100.00	100.00
	k141_743709	pN1566_2	*Salmonella enterica* subsp. *enterica* serovar *Schwarzengrund*	99.77	100.00
WD2	k141_109523	pVB82_1	*Acinetobacter baumannii*	99.94	92.00
	k141_113493*	pYH12207-3	*Acinetobacter piscicola*	99.93	78.00
	k141_500987	p24358-1	*Salmonella enterica* subsp. *enterica* serovar *Bredeney*	99.86	100.00
	k141_723703*	pNFYY023-1	*Comamonas testosteroni*	100.00	99.00
	k141_740211	pVCGX2	*Vibrio campbellii*	99.94	100.00
	k141_740911	pMH17-012N_3	*Citrobacter freundii*	99.19	95.00
	k141_76560	p3	Novosphingobium sp. *ES2-1*	99.93	72.00
	k141_777886	p1681-tetX	*Empedobacter falsenii*	98.40	98.00
SD3	k141_1668271	pOXA58_010030	*Acinetobacter defluvii*	99.82	100.00
	k141_625842	pCF39S	*Pseudomonas aeruginosa*	99.94	99.00
SD4	k141_546527	pEI-2234-3	*Edwardsiella ictaluri*	61.00	100.00
SD5	k141_806265	pG5A4Y217	*Escherichia coli*	99.78	74.00
WD5	k141_583117	pGENC284	*Enterobacter hormaechei* subsp. *Xiangfangensis*	99.81	59.00

SD6	k141_104572	pAH01-4	*Escherichia coli*	99.74	100.00
WD6	k141_124684	p116753-FIIK	*Klebsiella pneumoniae*	100.00	82.00
	k141_158642*	pPm14C18	*Proteus mirabilis*	99.90	51.00
	k141_173719	pWP7-S18-ESBL-04	*Klebsiella* sp. *WP7-S18-ESBL-04*	99.97	72.00
	k141_243506*	pYPR31	*Providencia rettgeri*	100.00	100.00
SD7	k141_100623	p63039	*Myroides odoratimimus*	99.89	100.00
	k141_204766	pKP20194a-p3	*Klebsiella pneumoniae*	100.00	100.00
	k141_225349	p1	*Neisseria gonorrhoeae*	98.74	100.00
	k141_239404	p24358-1	*Salmonella enterica* subsp. *enterica* serovar *Bredeney*	99.97	99.00
	k141_24536	pA2293-Ct2	*Klebsiella pneumoniae*	99.64	100.00
	k141_270502	pNA6	Uncultured bacterium	99.93	30.00
	k141_333949	pLraf_19_5_1	*Lactococcus raffinolactis*	94.90	37.00
	k141_383255	pHNCF11W-130kb	*Escherichia fergusonii*	100.00	100.00
	k141_393768	pVCGX2	*Vibrio campbellii*	99.83	100.00
	k141_451891	pRErm46	*Rhodococcus hoagii*	99.00	89.00
	k141_464290	pC16KP0065-1	*Klebsiella pneumoniae*	100.00	39.00
	k141_479837	pEI-2234-3	*Edwardsiella ictaluri*	99.90	100.00
	k141_556637	pRGRH0399	Uncultured bacterium	94.84	80.00
	k141_557329	pHDC14-2.133K	*Enterococcus hirae*	98.45	88.00
	k141_569663*	pNFYY023-1	*Comamonas testosteroni*	99.98	80.00
	k141_574437	Plasmid 2	*Salmonella enterica* subsp. *enterica* serovar *Typhi*	99.44	96.00
	k141_94024*	pAb-C63_1	*Acinetobacter baumannii*	99.95	79.00
WD7	k141_113036	pEI-2234-3	*Edwardsiella ictaluri*	100.00	100.00
	k141_119966	pHNCF11W-130kb	*Escherichia fergusonii*	100.00	100.00
	k141_139664	pMS2H5VEB-1	*Klebsiella pneumoniae*	99.88	95.00

	k141_16407	pBS228	*Pseudomonas aeruginosa*	99.79	71.00
	k141_174331	pRSB222	Uncultured bacterium	97.35	83.00
	k141_262055	pG5A4Y217	*Escherichia coli*	90.98	85.00
	k141_323094	pN1566_2	*Salmonella enterica* subsp. *enterica* serovar *Schwarzengrund*	100.00	100.00
	k141_337873	pMH17-012N_3	*CItrobacter freundii*	98.83	92.00
	k141_362572	pALTS33	Uncultured bacterium	93.81	54.00
	k141_38198	pTet	*Campylobacter jejuni* subsp. *jejuni*	99.89	100.00
	k141_389765	pCAP01	*Capnocytophaga ochracea*	89.11	35.00
	k141_390832	pMR0211	*Providencia stuartii*	99.96	95.00
	k141_47250	pE211-2	*Enterococcus faecalis*	99.41	100.00
	k141_51571*	p3iANG	*Vibrio cholerae*	97.91	85.00
	k141_5646*	pYH12207-3	*Acinetobacter piscicola*	100.00	96.00
	k141_7584	p63039	*Myroides odoratimimus*	99.89	26.00
	k141_93260*	pSY153-MDR	*Pseudomonas putida*	99.97	96.00
	k141_98708	p345-185	*Vibrio harveyi*	94.32	77.00
SAS1	k141_249579	p3	*Novosphingobium* sp. *ES2-1*	99.92	85.00
WAS1	k141_49220	p4130-KPC	*Pseudomonas aeruginosa*	99.93	100.00
SAS2	k141_781158	pEI-2234-3	*Edwardsiella ictaluri*	100.00	100.00
WAS3	k141_1155247	p3	*Novosphingobium* sp. *ES2-1*	90.90	92.00
WAM1	k141_388101	pP72_e	*Phaeobacter inhibens*	98.95	99.00
	k141_448754	pEI-2234-3	*Edwardsiella ictaluri*	99.65	100.00
WAM3	k141_457349	pCF39S	*Pseudomonas aeruginosa*	100.00	88.00
WAM6	k141_67005	pHH2-227	Uncultured bacterium	99.94	99.00

* Multiple resistance genes

3.5 Discussion

In this study, we used quantitative culture and metagenomic techniques to understand the community composition and the level of antibiotic resistance genes in rural and urban surface water sites across Bangladesh. Selective plating showed that ESBL-producing coliforms were more prevalent in urban surface water than rural settings, consistent with reports of antibiotic-resistant faecal coliforms in rivers across Asia (Lamba *et al.*, 2018; Yu *et al.*, 2019). However, the predictive value of the abundance of ESBL-producing coliforms for the total abundance of antibiotic resistance genes was found to be limited, suggesting that ESBL-producing coliforms are not necessarily a valid proxy to determine the total load of antibiotic resistance genes in environmental ecosystems.

In addition to quantitative culture of ESBL-producing coliforms, a metagenomic shotgun sequencing approach was used to characterize the microbiota of each sample and quantify the abundance of antibiotic resistance genes in water and sediment samples. We found that the water and sediment samples grouped together by their type (water or sediment) rather than the location they were collected from. Sediment samples were dominated by bacteria belonging to the genera *Pseudomonas*, *Azoarcus*, and *Hydrogenophilacea*, which is in line with other studies which have shown that sediment is dominated by the phylum *Proteobacteria* (Nho *et al.*, 2018). Water samples were dominated by the cyanobacteria *Cyanobium* and *Microcystis* that cause harmful blooms in aquaculture ponds (Zhong *et al.*, 2011). *Microcystis* produces potent toxins which can kill fish but are also harmful to humans (Paerl and Tucker, 1995). The two river water samples and a public pond water sample collected in Dhaka clustered away from the other water samples and were defined by an increased abundance of bacteria associated with the human intestinal tract (Figure 3.6). The

presence of increased amounts of the faecal indicator bacteria *E. coli* strongly suggests that human waste is contaminating the urban surface water (Ouattara *et al.*, 2013).

Several different types of antibiotics were used in the rural aquaculture ponds which we surveyed. The antibiotics were either mixed with feed or added directly to the ponds for the treatment of disease. Fluoroquinolone antibiotics such as ciprofloxacin and levofloxacin were the most widely used antibiotics in the rural aquaculture ponds; however, high numbers of fluoroquinolone resistance genes were not observed in the rural sites with prior antibiotic use. Resistance to fluoroquinolone drugs is mainly mediated by chromosomal mutations in the *parC* and *gyrA* genes, so the absence of dedicated resistance genes in these ecosystems may be unsurprising (Hooper and Jacoby, 2015). However, we note that the multidrug efflux pump genes *mexV*, *mexF*, *adeI*, and *adeH* were exclusively found in the rural sites with prior antibiotic use, and these efflux systems are capable of exporting fluoroquinolones from the cell (Köhler *et al.*, 1999; Li *et al.*, 2003; Damier-Piolle *et al.*, 2008; Coyne *et al.*, 2010). The macrolide drug erythromycin was another antibiotic which was widely used in aquaculture ponds that were sampled in this study. However, levels of macrolide resistance genes were low in the rural aquaculture ponds but extremely high in a subset of the urban samples. Notably, the erythromycin resistance gene *msrA* (Reynolds, Ross and Cove, 2003) was only present in the aquaculture ponds with prior antibiotic use. This gene was previously found in the intestinal contents of farmed rainbow trout and may thus be more commonly associated with aquaculture (Muziasari *et al.*, 2017). We did not observe a difference in the total load of antibiotic resistance genes in rural ponds with and without a history of antibiotic use, suggesting that other factors than solely the historical use of antibiotics in fish farming are more important in shaping the resistome of the rural ponds. We postulate that heavy metals in the environment could be responsible for this observation. Previous work in

Bangladesh has shown high levels of heavy metals in surface water, sediments, and fish (Ahmad *et al.*, 2010; Sarker Md, 2016; Hossain *et al.*, 2021), and heavy metals can co-select for antibiotic resistance in aquaculture ponds (Seiler and Berendonk, 2012). Further research is needed to quantify the impact of different fish farming practices and environmental variables on the selection for antibiotic resistance in aquaculture.

Antibiotic resistance was the highest in urban areas, which suggests that human factors contribute to the accumulation of antibiotic-resistant bacteria in the environment. This was further corroborated by the correlation between the abundance of bacteria originating from the human gut and antibiotic resistance gene abundance observed in our study. The rivers and lakes of Dhaka are surrounded by slums with high population densities in which 13.7% of households report that human waste is directly released into lakes, ponds, or rivers (Arias Granada *et al.*, 2018). Our study thus extends on previous observations that link the introduction of human sewage into river and lake systems to high levels of antibiotic resistance genes. A study of over 500 metagenomes demonstrated a strong correlation between levels of faecal pollution and ARG abundance (Karkman, Pärnänen and Larsson, 2019). Several other studies have linked levels of antibiotic resistance genes with faecal contamination in both high and low-and-middle income countries demonstrating that this observation is not specific to a particular country or the level of development of that country (Agramont *et al.*, 2020; Reynolds *et al.*, 2020). Studies thus far have focused mainly on antibiotic resistance in the urban environment, but the use of antibiotics in agriculture and aquaculture necessitates the further study of antibiotic resistance in these rural environments. Studies that compare the urban and rural environment may also be required to help disentangle the different factors influencing the high levels of antibiotic resistance in the urban environment.

By creating a metagenomic assembly of our short-read sequencing data, we were able to identify contigs which contained plasmid replication initiation genes and found that IncP, IncQ, and various small plasmid types were most common. Only the IncQ1 plasmid pWD1 contained an antibiotic resistance gene (*sul2*). Using PlasFlow on the metagenomic assemblies, we were able to assign antibiotic resistance genes to 93 plasmid contigs, revealing that urban samples had a higher number of plasmids carrying antibiotic resistance genes. This suggests that, particularly in urban water bodies, there exists an increased potential of horizontal gene transfer of mobile genetic elements carrying antibiotic resistance genes.

The microbiotas of surface water and sediment samples across Bangladesh are diverse, but antibiotic resistance genes are highly abundant in urban samples and are more commonly associated with plasmids in this setting. While the abundance of antibiotic resistance genes was considerably lower in rural than in urban settings, we nonetheless observed evidence for the selection for fluoroquinolone resistance mechanisms in ponds used for fish farming. Policies to minimize the use of antibiotics in aquaculture should thus remain a priority to reduce selection for antibiotic resistance. The presence of human gut bacteria was associated with high levels of antibiotic resistance genes, suggesting that contamination by human waste is an important driver for the presence of antibiotic resistance genes in surface water. Interventions aimed at improving access to clean water, sanitation, and sewerage infrastructure may thus be important to reduce the risk of AMR dissemination in Bangladesh and other low- and middle-income countries.

3.5.1 Future Work

To confirm that the antibiotic resistant bacteria found in the urban surface water samples were from untreated sewage, it is necessary to sample the gut bacteria of people living within Dhaka. Comparing the bacteria in the environment and the bacteria in the human gut would identify any connectedness between them. However, it would be difficult to identify whether the residents had been colonised by the bacteria through their use of the surface water or whether they themselves contaminated these sites. Sampling persons that do not directly interact with these urban surface water sites would help to disentangle this connection. Long-read sequencing could also be used to provide novel insights into the hosts of the antibiotic genes present and would allow better reconstruction of MGEs such as plasmid and bacteriophage from the samples. As the majority of antibiotic resistance genes were found within the urban environment a study focussing solely on this environment with a larger samples size would provide important insights into ABR gene dynamics within a megacity.

3.6 Key Findings

- The microbiotas of surface water and sediment are distinct and are defined by their sample type (sediment or water) rather than their location or antibiotic use.
- Sediment samples were dominated by the phylum *Proteobacteria* while water samples were dominated by the phylum *Cyanobacteria*.
- There was a significantly higher human gut contribution in the urban water samples than the rural samples with no previous antibiotic use.
- The urban sediment samples were significantly more diverse than either of the rural sample types.
- Urban samples contained a greater number of antibiotic resistance genes than either of the rural sample types.
- The urban sediment and water samples contained significantly higher levels of antibiotic resistance genes than the rural sites with previous antibiotic use but not those that had no previous antibiotic use.
- There was a strong positive correlation between the contribution of human gut bacteria to a sample and the levels of antibiotic resistance genes in that sample.
- Metagenomic assembly uncovered plasmid pWD1 which is a possible precursor plasmid of the well-known plasmid RSF1010.
- The urban sediment samples contained significantly more plasmid contigs that carried antibiotic resistance genes than the either of the rural sediment sample types.

.

Chapter 4

Genomic epidemiology of *Enterococcus faecium* in a Birmingham hospital

4.1 Introduction

In the previous chapter I demonstrated that antibiotic resistance can be highly prevalent in the environment and that humans are a likely source of antibiotic resistant bacteria. To expand on these findings, I then focused on antibiotic resistance in a clinical setting, specifically at vancomycin resistant *Enterococcus faecium* (VRE) in a hospital in Birmingham (United Kingdom).

Highly similar *E. faecium* strains are found in hospitals across the UK and almost exclusively belong to the clinically associated Clade A1. However, some *E. faecium* strains cluster within referral networks which is likely caused by transmission and clonal expansion within the individual networks (Raven *et al.*, 2016). The phylogenetic distribution of *E. faecium* isolates in Birmingham hospitals has not been studied to date and this would potentially provide important insights into transmission and antibiotic resistance of *E. faecium* within the city.

Some *E. faecium* strains that have vancomycin resistance genes are not phenotypically resistant to vancomycin, these are known as vancomycin variable enterococci (VVE) (Kohler *et al.*, 2018). VVE strains have the ability to revert to a vancomycin-resistant phenotype when exposed to sub-MIC concentrations of vancomycin (Thaker *et al.*, 2015). Vancomycin variability has been linked to the movement of insertion sequences into and out of the *vanH* promoter region, but other mechanisms have been described (Thaker *et al.*, 2015; Sivertsen *et al.*, 2016; Wagner *et al.*, 2021).

In this chapter, I used short and long-read sequencing to investigate an outbreak of *Enterococcus faecium* in a Birmingham hospital and to investigate the retention of *E. faecium* strains over different time periods. I also used long-read sequencing to uncover the

mechanism behind the phenotype-genotype disparity in clinical *E. faecium* isolates and the mechanism behind their reversion to a resistant phenotype.

4.2 Hypothesis

It was hypothesised that the outbreak of *Enterococcus faecium* in a Birmingham hospital was caused by clonal expansion of a single *E. faecium* strain. It was also hypothesised that the same *E. faecium* strains would be maintained in patients over a long period of time.

4.3 Aims and objectives

1. To identify the *E. faecium* sequence types present in a Birmingham hospital.
2. To identify the antibiotic resistance profile of *E. faecium* strains present in Birmingham hospitals.
3. To identify which plasmids were present in these isolates.
4. To investigate the phenotype – genotype disparity found in some *E. faecium* isolates.
5. To investigate the mechanism by which vancomycin-variable isolates revert to a resistant phenotype.

4.4 Results

4.4.1 Description of isolates

Sixty *Enterococcus faecium* isolates (38 blood culture isolates, 21 rectal swab isolates and one urine isolate) were collected from 36 patients admitted to Heartlands hospital Birmingham (Table 4.1). Twenty-two of the blood culture isolates had been phenotypically typed as vancomycin resistant while 16 were typed as susceptible. All twenty-one of the rectal swab isolates were phenotypically vancomycin resistant. The single urine isolate was also characterised as being vancomycin resistant.

Table 4.1: Phenotypic characterisation of hospital *E. faecium* isolates.

	Vancomycin-resistant	Vancomycin-susceptible
Blood culture	22	16
Rectal swab	21	0
Urine	1	0
Total	44	16

4.4.2 Isolation date

The clinical *E. faecium* isolates were collected in two distinct time periods (Figure 4.1). The first set of isolates (n = 39) were collected during a period of increased vancomycin resistant *Enterococci* (VRE) linked bacteraemia in 2016 and 2017. The second set of isolates (n = 21) were routine screening isolates that were collected in 2019 and 2020. The shortest period of time between collection of isolates from the same patient was the same day (Patient 13) and the maximum period was 442 days (Patient 32).

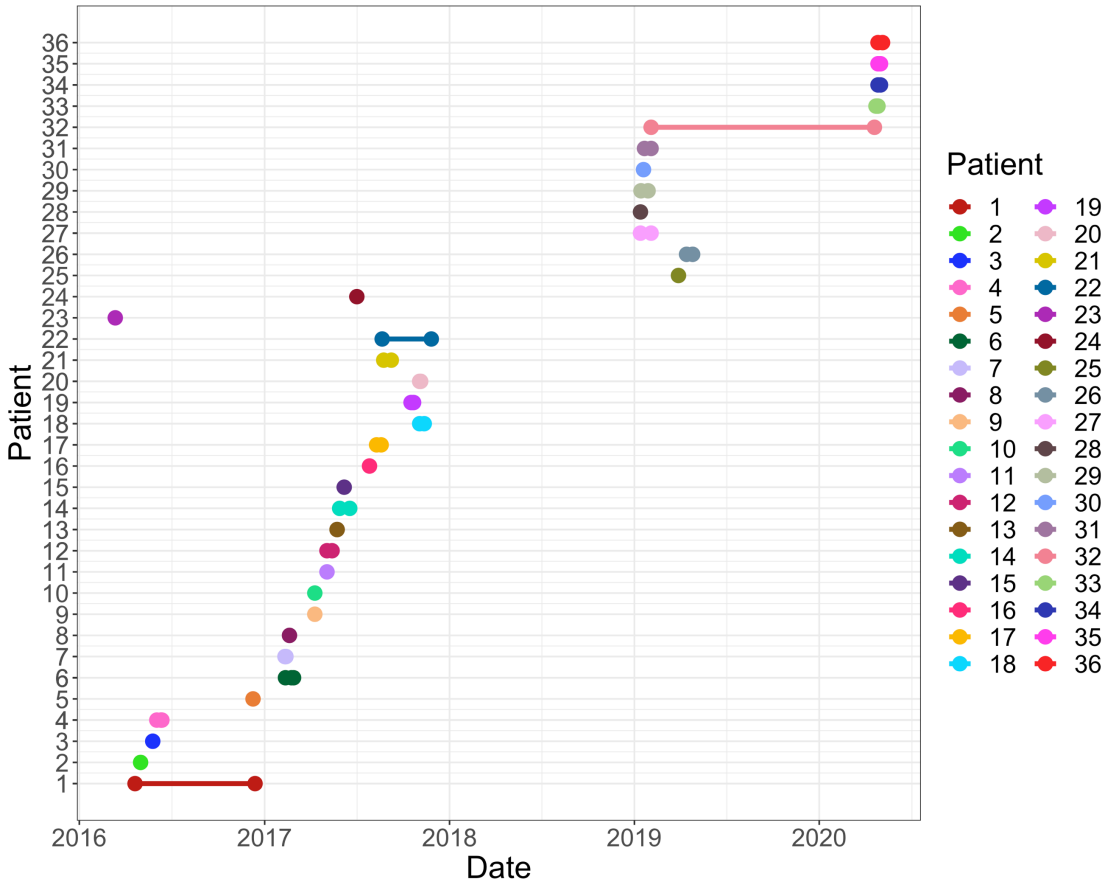

Figure 4.1: Isolation date of *E. faecium* isolates from each patient. Patients that had more than one isolate have had their isolation dates connected by a horizontal line.

4.4.3 Short-read assembly statistics

The short reads of the 60 isolates were assembled to produce genomic assemblies. These assemblies ranged in size from 2,821,205 bp to 3,191,929 bp with a mean length of 3,009,769.2 bp (Figure 4.2A). The assemblies had a mean %GC of 37.6% and only varied by a maximum of 0.45% across the assemblies (Figure 4.2B). The mean N50 of the assemblies was 41,235.2 bp with the lowest N50 being 32,792 bp and the highest N50 being 55,763 bp. All assemblies had an N50 greater than 15 kb which is an indication of good assembly quality (Ellington *et al.*, 2017)(Figure 4.2C). An average nucleotide identity (ANI) was also calculated against a clade A1 *E. faecium* E745 reference genome for each of the isolates. The assemblies had a mean ANI of 98.9% and all values were greater than 95% indicating that they were correctly assembled *E. faecium* genomes (Chan *et al.*, 2012) (Figure 4.2D).

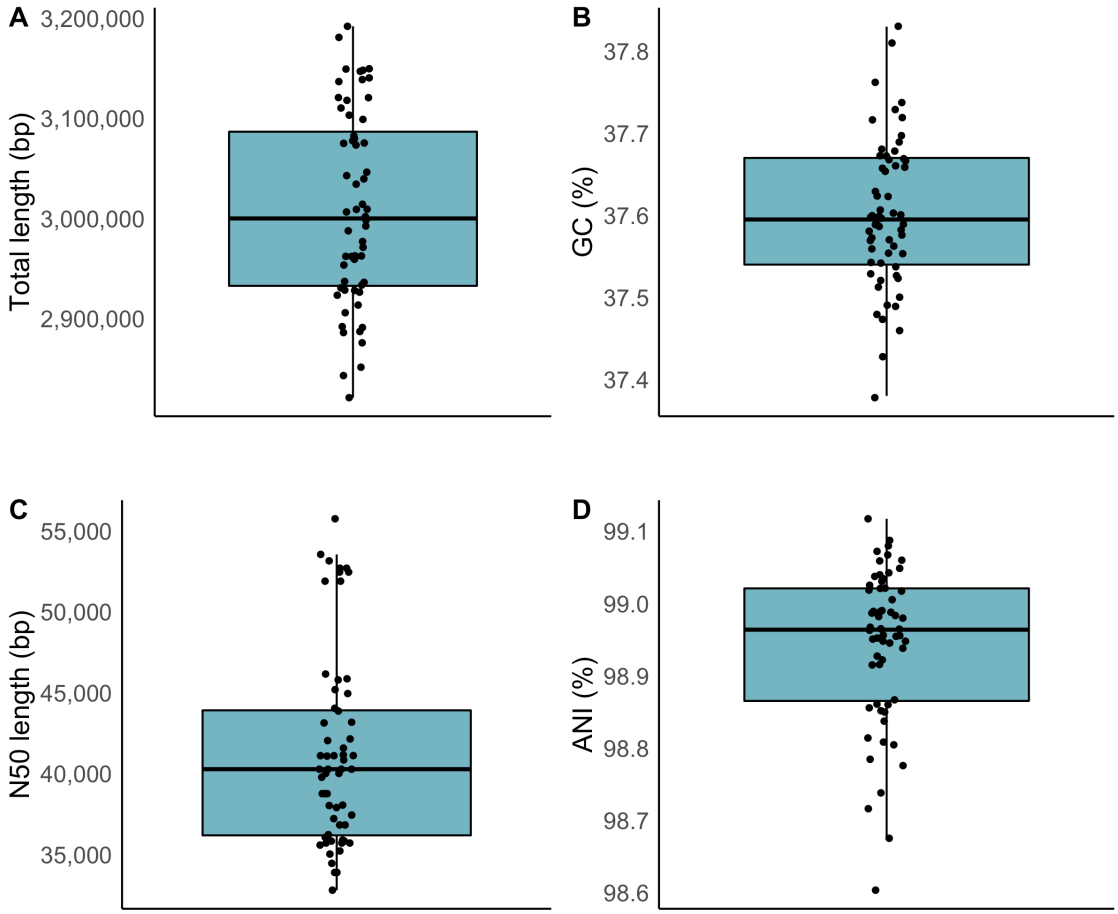

Figure 4.2: Short-read assembly statistics for 60 clinical vancomycin resistant *Enterococcus faecium* isolates. **A.** Total length of each assembly (bp). **B.** GC% of each assembly. **C.** N50 of each assembly (bp). **D.** Average Nucleotide Identity of each isolate to clinical *E. faecium* isolate E745 (%).

4.4.4 The hospital isolates belonged to multiple sequence types

The *E. faecium* isolates were multilocus sequence typed according to seven core genes (*atpA*, *ddl*, *gdh*, *purK*, *gyd*, *pstS* and *adk*) (Homan *et al.*, 2002). The isolates collected during the outbreak in 2016 and 2017 were composed of 8 different sequence types, the most prevalent of which was ST262 (*n* = 20) which accounted for 51% of all of the outbreak isolates (Figure 4.3). ST80 (*n* = 7), ST1478 (*n* = 5) and ST780 (*n* = 3) were the next most prevalent sequence types and collectively made up 38% of isolates. The remaining sequence types ST117, ST203, ST412 and ST787 were all represented by a single isolate. The isolates collected in 2019 and 2020 comprised of 4 different sequence types with ST80 (*n* = 11) being the most prevalent, followed by ST262 (*n* = 5), ST780 (*n* = 4) and a single isolate of ST787 (Figure 4.3). When combined, the hospital isolates represented 8 sequence types all of which were first seen in the isolates collected from the 2016 – 2017 outbreak (Figure 4.3). ST262 (*n* = 25) and ST80 (*n* = 18) were by far the most prevalent sequence types accounting for 72% of all isolates collected. Multiple isolates also belonged to sequence types ST780 (*n* = 7), ST1478 (*n* = 5) and ST787 (*n* = 2). Only a single isolate of sequence type ST117, ST203 and ST412 was present during the study period.

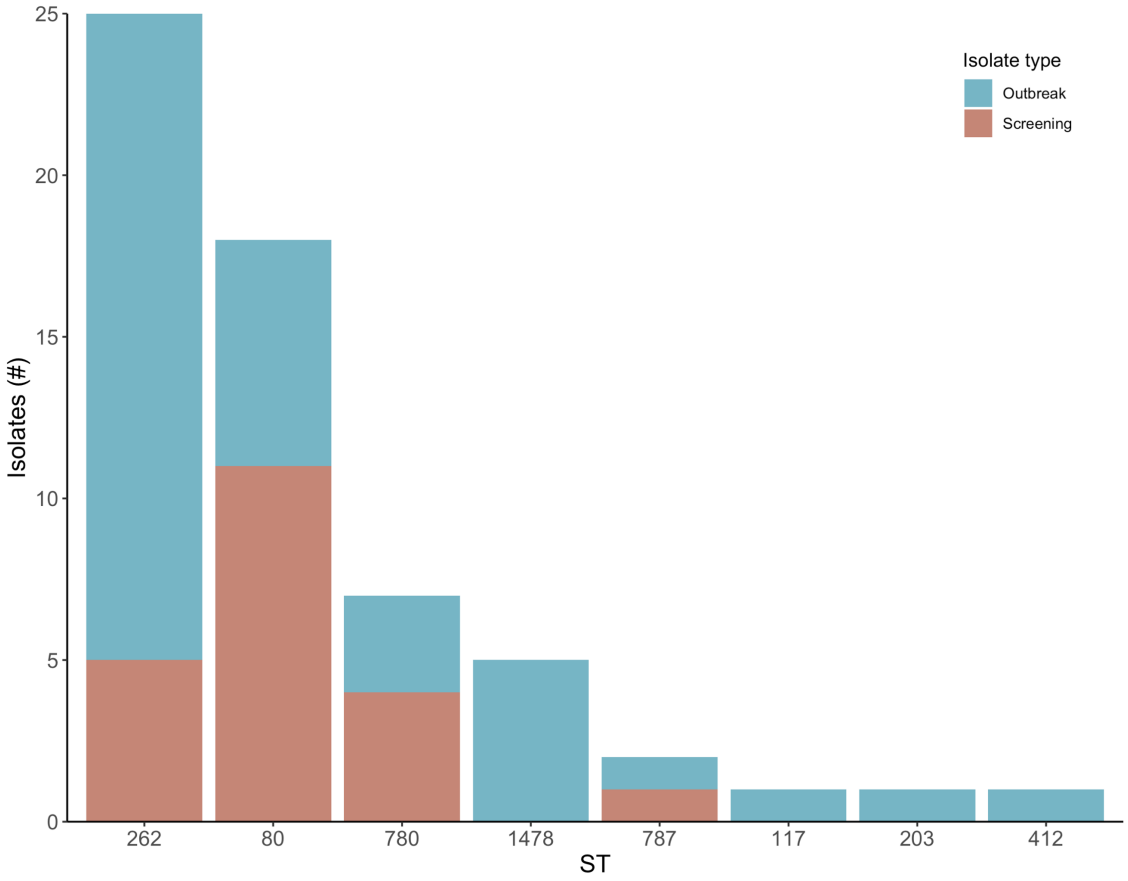

Figure 4.3: Multi locus sequence types of 60 clinical *E. faecium* isolates. Blue bars represent the number of outbreak isolates of each sequence type and the pink bars represent the number of screening isolates of each sequence type.

4.4.5 Phylogenetic structure of hospital isolates

The phylogenetic distribution of *E. faecium* in both the outbreak isolates and screening isolates was polyclonal in nature (Figure 4.4). However, there were cases of clonal expansion within the outbreak isolates. Seventeen of the outbreak isolates were closely related and belonged to ST262, demonstrating that although the outbreak was polyclonal, ST262 was the dominant clone. Another clonal group was formed of isolates OI9, OI11, OI12 and OI13 which belonged to ST1478. Although there were screening isolates that clustered apart from the outbreak isolates, many of the screening isolates clustered together with the outbreak isolates suggesting a high degree of relatedness.

There were eight patients (1, 7, 12, 13, 17, 18, 22 and 27) that were colonised by more than one *E. faecium* sequence type. Patient 27's isolates were collected from rectal swabs taken 3 weeks apart. The first isolate belonged to ST262 while the second isolate belonged to ST787. Patient 1 had two episodes of bacteraemia eight months apart, the strain isolated from the first episode belonged to ST262 and the strain from the second episode belonged to ST787. Patient 22's isolates were cultured during an episode of bacteriaemia and from a rectal swab taken 3 months later. The blood isolate belonged to ST80 while the rectal screening isolate belonged to ST262. Patients 12 and 17 had blood samples cultured 10 and 9 days apart respectively, patient 12 had an ST203 isolate in the first blood sample and an ST1478 in the second sample while patient 17 went from an ST262 isolate to an ST780 isolate. Patient 7 had blood cultures taken two days apart which showed an ST117 isolate in the first instance followed by an ST80 in the second instance. Patient 13 had two different sequence types, ST412 and ST1478, present in a single blood culture. A cluster of four isolates (OI9, OI11, OI12 and OI13) belonging to ST1478 were identical at the core genome

level and were found in four different patients (9, 11, 12 and 13). Patient 32 had rectal swabs taken 14 months apart which both contained ST262 *E. faecium* isolates.

Patients 18, 19, 20, 21 and 22 had both rectal screening swabs and blood cultures taken which meant that it was possible to compare isolates between the different body sites. Patients 18 and 22 both had ST80 isolates in their blood cultures and ST262 isolates in their rectal screening samples. Patients 19, 20 and 21 had strains belonging to the same sequence type in their rectal and blood samples. Patient 19 had strains belonging to ST262 in both samples that were identical, based on the absence of any core SNPs. Patient 20's strains belonged to ST80 and both samples were also identical at the core genome level. Patient 21 had strains belonging to ST262 which were 1 core SNP different between the isolates.

Comparison of the outbreak and screening isolates identified that many of them were unrelated between the two time periods. The screening isolates that were found to be unrelated to the outbreak isolates could either have been missed during the outbreak or they may be new acquisitions to the hospital. There were, however, several of the screening isolates that were highly related to those identified during the outbreak. The presence of highly related isolates could suggest that strains are being maintained in the hospital environment and acting as a reservoir for reinfection of patients.

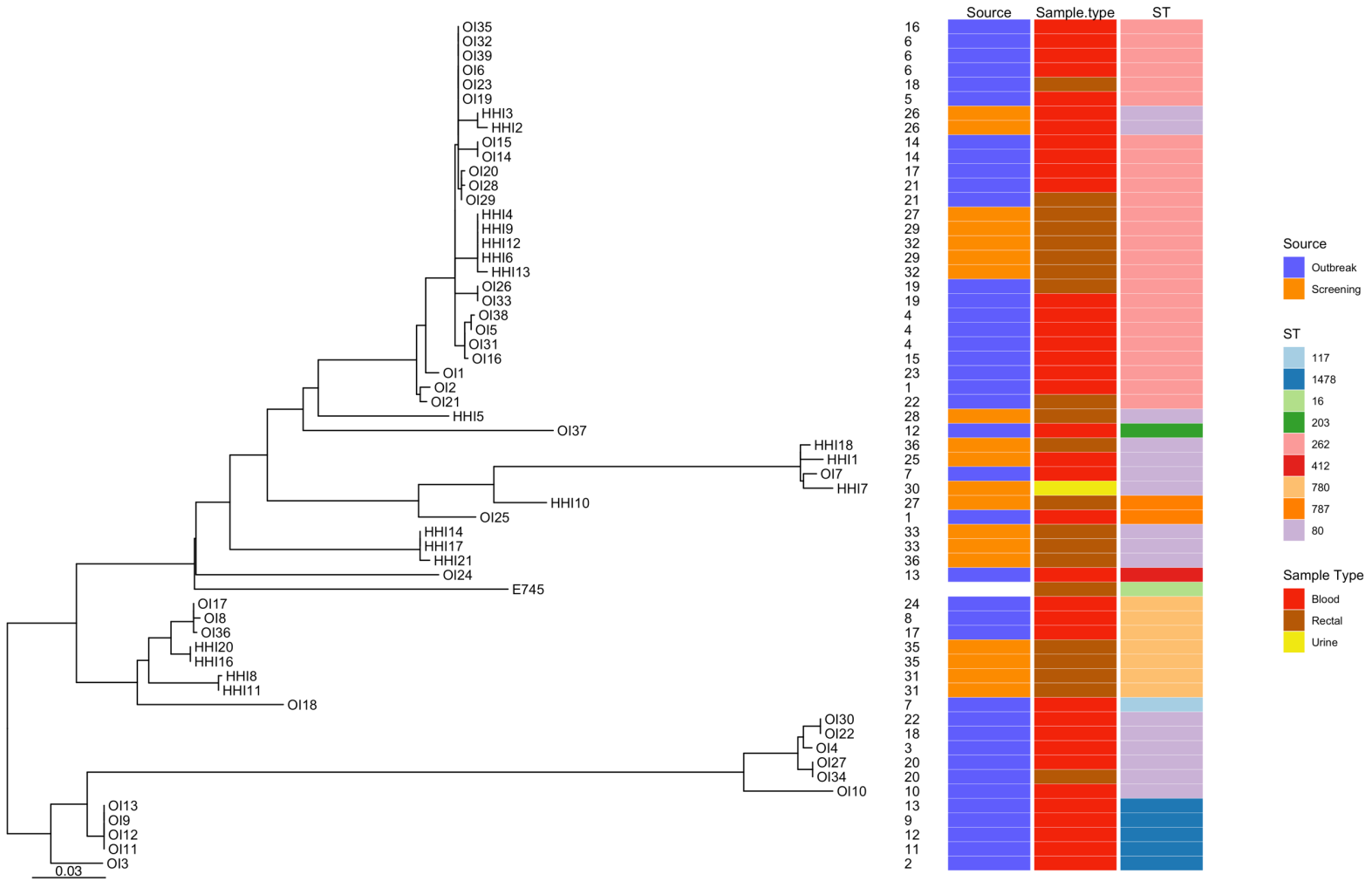

Figure 4.4: Maximum likelihood core genome phylogenetic tree of *E. faecium* isolates in a Birmingham hospital. Metadata includes the patient number, the source of the isolates (outbreak or routine screening), the sample type (blood, rectal swab or urine) and the sequence type (ST). *E. faecium* strain E745 was used as the reference strain. The scale bar indicates the average number of nucleotide substitutions per site.

4.4.6 Clinical *E. faecium* isolates contain a diverse population of plasmids

Seventeen plasmid replication (*rep*) genes were identified from the short-read assemblies of the hospital outbreak isolates (Figure 4.5). The number of plasmid *rep* genes varied greatly between the outbreak isolates. Isolates OI2 and OI22 had the most *rep* genes in a single isolate (n = 7) while isolates OI27 and OI34 had the least (n = 2). The mean number of *rep* genes per isolates was 5. *Rep2_1 _orf1*(pRE25) was the most abundant *rep* type as it was found in 37 of the 39 isolates. Five replication initiation genes (*rep14a_2_ORF1*(pKQ10), *rep14a_4_rep*(AUS0004p3), *rep14b_1_repA*(pJS42), *rep14b_2_EFAU085p6001* (AUS0085p6) and *rep29_2_EFAU085p4001*(Aus0085p4)) were only found in a single isolate. Twelve replication initiation genes were identified in the screening isolates (Figure 4.5). Isolates HHI14, HHI15, HHI17, HHI19 and HHI21 had the fewest *rep* genes at 2 and isolate HHI18 had the most with 6 *rep* genes. The mean number of *rep* genes per screening isolate was 4. The *rep11a_1_repA*(pB82) gene was the most widespread replication gene and was found in all of the screening isolates. Combining the outbreak and screening isolates, 18 plasmid *rep* genes were identified (Figure 4.5). *rep2_1_orf1*(pRE25) was the most widespread *rep* type being found in 52 of 60 isolates. The only strain in which a plasmid replication gene could be co-located with a *vanA* gene using the short-read assemblies was isolate OI38. The vancomycin resistance genes in this isolate were located on a pRUM-like plasmid (*rep17_1_CDS29*(pRUM)).

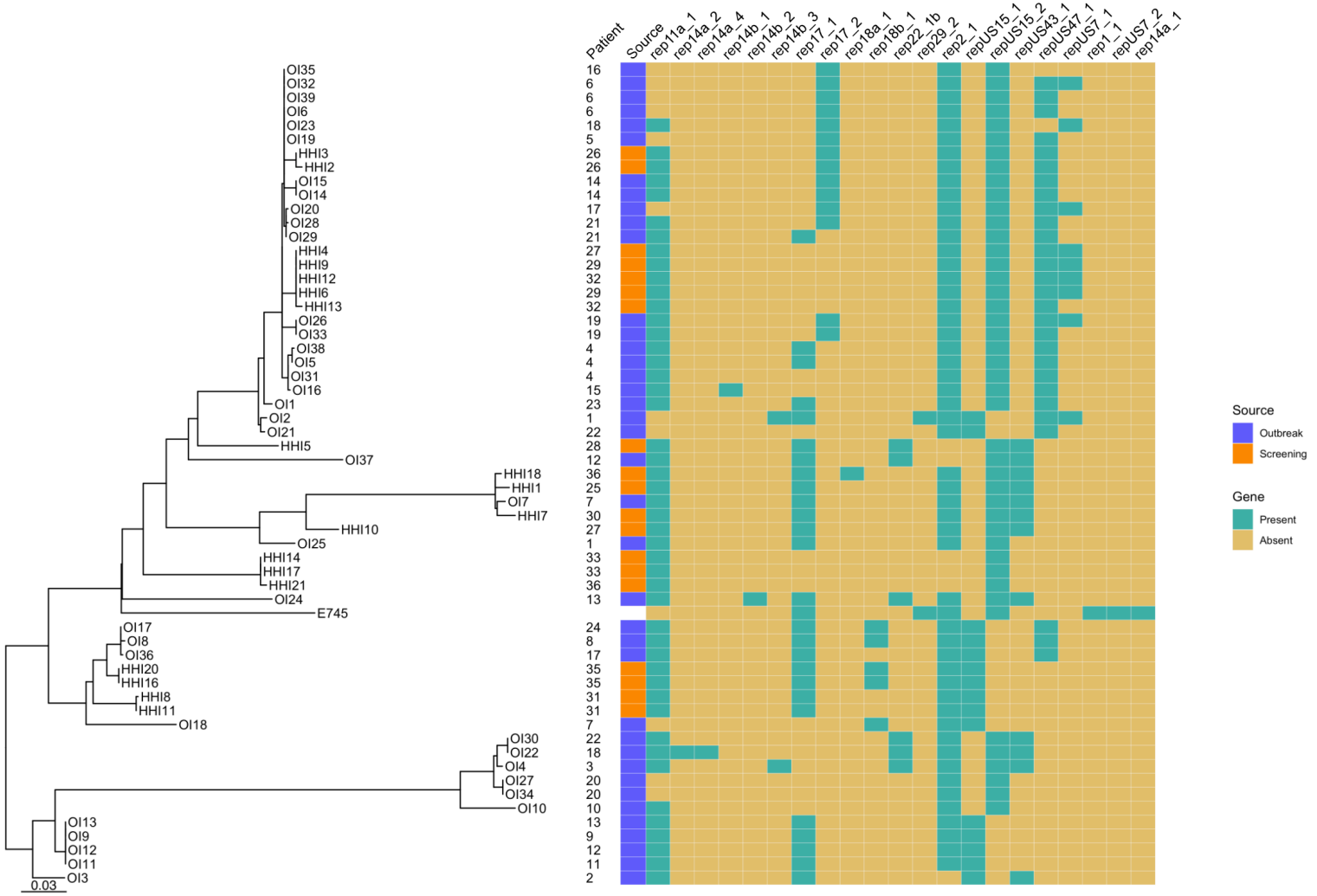

Figure 4.5: Maximum likelihood core genome phylogenetic tree of *E. faecium* isolates in a Birmingham hospital. Metadata includes the patient number, the source of the isolates (outbreak or routine screening) and the presence (green) or absence (gold) of plasmid replication genes in the isolates. *E. faecium* strain E745 was used as the reference strain. The scale bar indicates the average number of nucleotide substitutions per site.

4.4.7 Hybrid assembly statistics

Due to the fragmented nature of the short-read only assemblies, long- and short read data were used to construct hybrid assemblies with the aim to determine the genetic context of any vancomycin resistance genes. Fourteen phenotypically resistant and nine phenotypically sensitive isolates were sequenced and assembled. The hybrid assemblies ranged in size from 2,994,061 bp to 3,373,229 bp with a mean length of 3,174,496.9 bp (Figure 4.6A). The assemblies had a mean %GC of 37.7% and only varied by a maximum of 0.33% across the assemblies (Figure 4.6B). The mean N50 of the assemblies was 2,703,049.3 bp with the lowest N50 being 478,625 bp and the highest N50 being 2,958,528 bp (Figure 4.6C). An average nucleotide identity (ANI) was also calculated against an *E. faecium* E745 reference genome for each of the isolates (Figure 4.6D). The assemblies had a mean ANI of 99.1% and all values were greater than 95% indicating that they were correctly assembled *E. faecium* genomes.

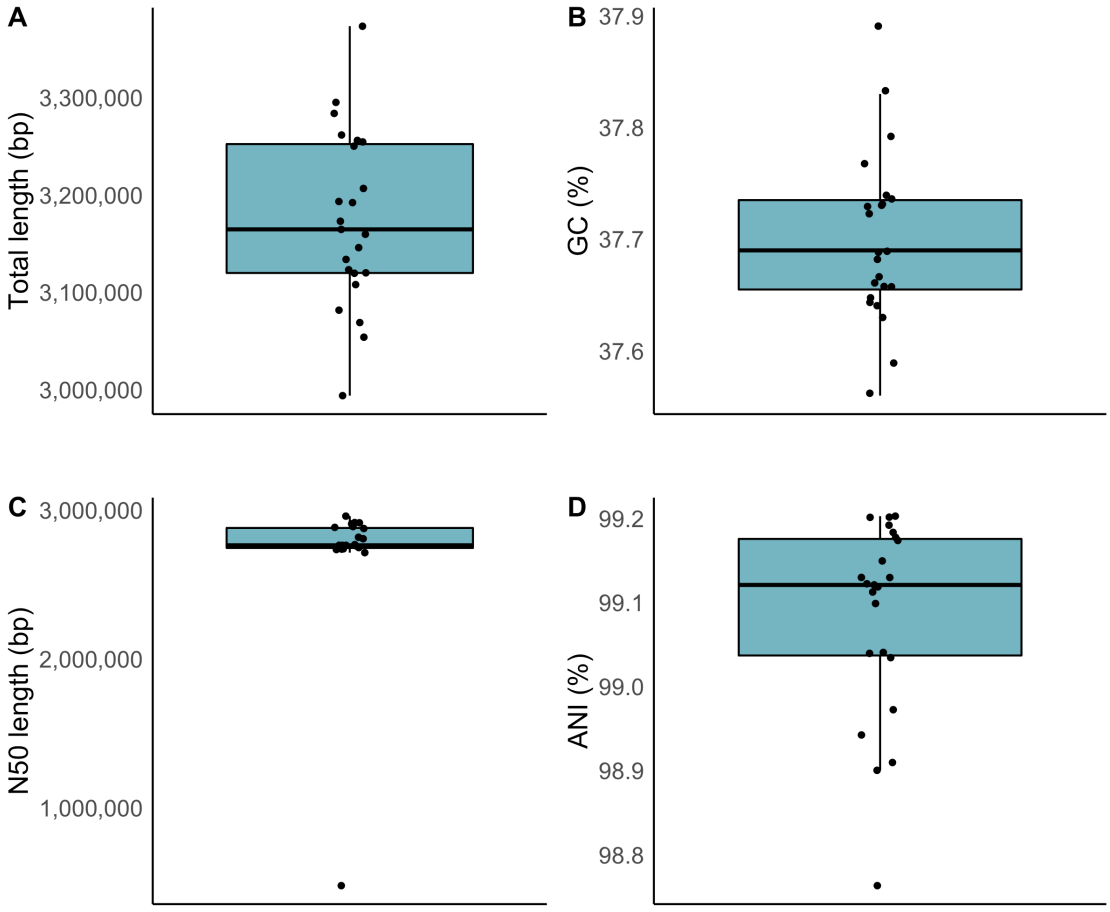

Figure 4.6: Long-read assembly statistics for 23 clinical *Enterococcus faecium* isolates. **A.** Total length of each assembly (bp). **B.** GC% of each assembly. **C.** N50 of each assembly (bp). **D.** Average Nucleotide Identity of each isolate to clinical *E. faecium* isolate E745 (%).

4.4.8 Hybrid assembly could not contextualise the majority of the vancomycin resistance genes

Hybrid assembly of 23 *E. faecium* isolates yielded an average of 5 circular replicons per isolate. The most replicons in a single isolate was 8 (isolate OI17) and the least was 2 (isolates OI19 and HHI13). Hybrid assembly of the long and short reads was only able to link the vancomycin resistance genes to a plasmid replication gene in four of the sixteen isolates that contained vancomycin resistance genes. The vancomycin resistance genes were co-located with three different *rep* types. As with isolate OI38, the vancomycin resistance genes in isolate OI5 were found on a plasmid containing *rep17_1_CDS29*(pRUM), the

vancomycin resistance genes in isolates HHI2 and HHI3 were associated with the *rep17_2_repA*(AUS0004p1) gene and the vancomycin resistance genes in isolate HHI18 were associated with a plasmid containing the *rep18a_1_repA*(p200B) gene. The completely assembled vancomycin resistance plasmids are discussed further below.

Plasmid pOI5 is a 25,483 bp plasmid that carries the *rep17_1_CDS29*(pRUM) replication initiation gene (Figure 4.7). Consistent with the prototypical pRUM plasmid (Genbank: NC_005000) it contains the Axe-Txe toxin-antitoxin system. The vancomycin resistance genes *vanRSHAXYZ* are the only antibiotic resistance genes carried by this plasmid. There are no conjugation or mobilisation genes on plasmid pOI5.

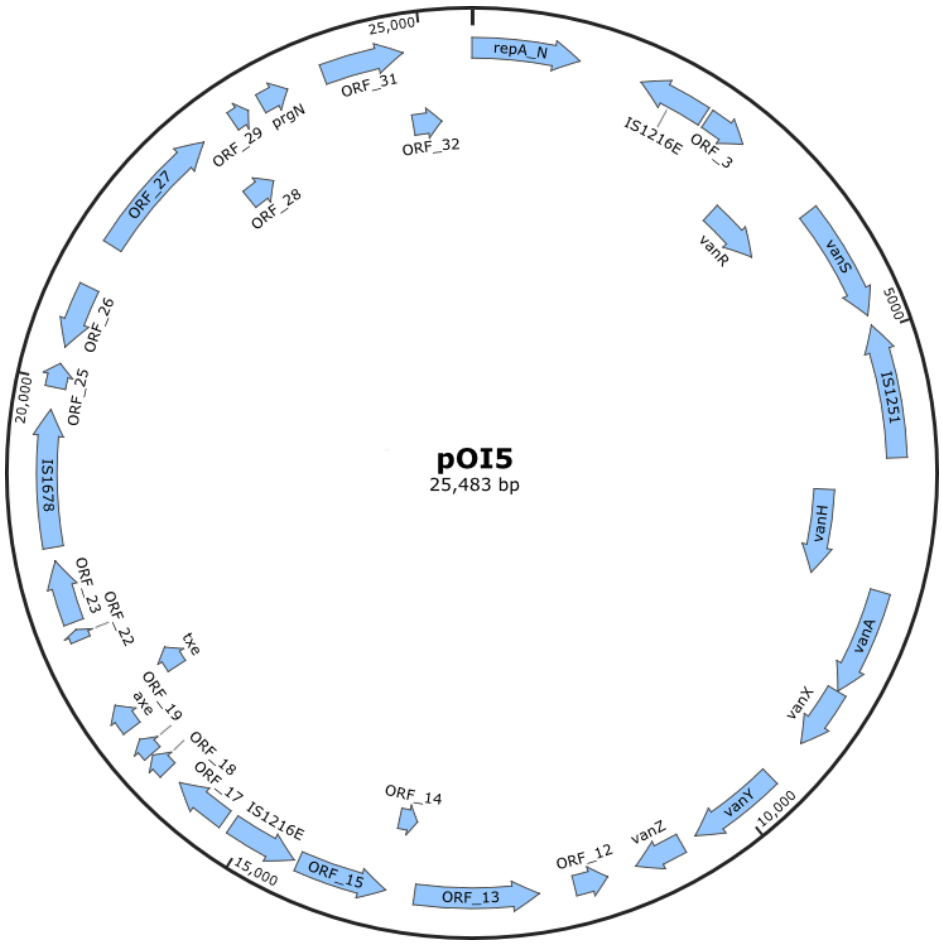

Figure 4.7: Map of plasmid pOI5. Blue arrows represent genes. ORF = Open Reading Frame.

Plasmid pHHI2 is a 71,762 bp plasmid that carries two replication initiation genes *rep17_2_repA*(AUS0004p1) and *repB* (Figure 4.8). The vancomycin resistance genes are the only antibiotic resistance genes found on plasmid pHHI2. However, the vancomycin resistance genes in this plasmid are thought to be non-functional based on the phenotype of the strain carrying this plasmid. Plasmid pHHI2 carries the *traE* and *traG* conjugation genes as well as the *mobL* mobilisation gene which suggests that it has the potential to be transferred horizontally to other *E. faecium* strains. The plasmid contains three toxin-antitoxin systems, an AbiE system, an Axe-Txe system and a RelE/ParE family system, making it highly likely that the plasmid is retained during cell division. Plasmid pHHI2 is identical to plasmid pHHI3 except for a 1332 bp insertion containing a copy of the IS*256* transposase.

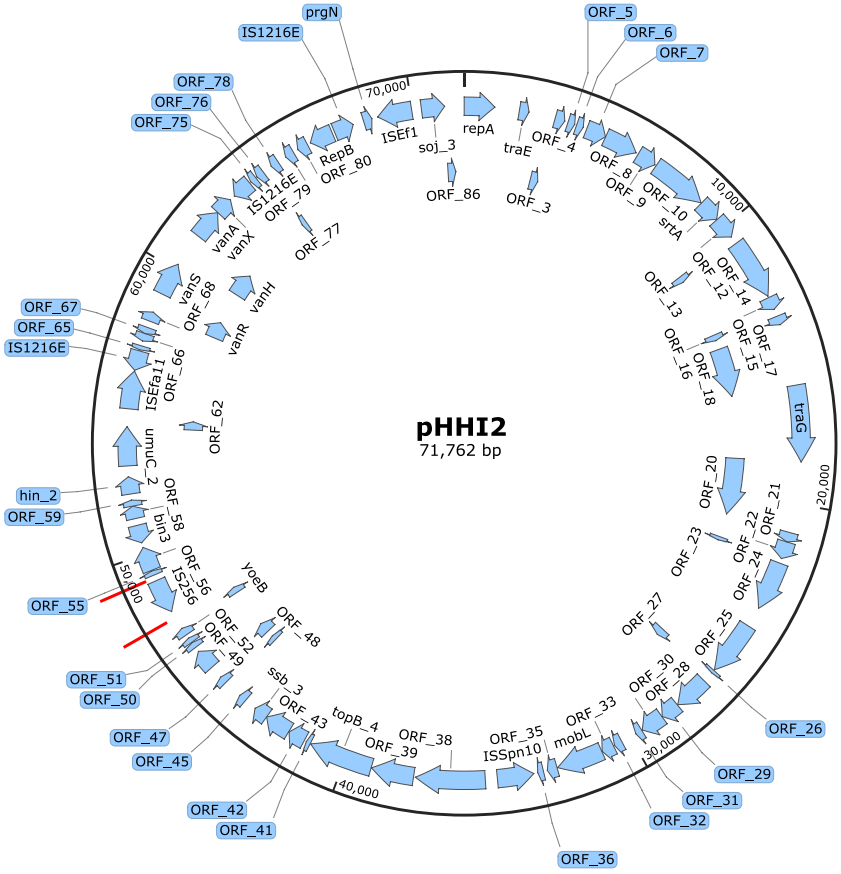

Figure 4.8: Map of plasmid pHHI2 and pHHI3. Red lines mark the section of DNA deleted in pHHI3. Blue arrows represent genes. ORF = Open Reading Frame.

4.4.9 Antibiotic resistance was ubiquitous in the hospital isolates

E. faecium strains isolated from the 2016 outbreak contained a total of 18 different antibiotic resistance genes (Figure 4.9). Of the 18 genes, the macrolide resistance gene *msrC_1* and the aminoglycoside resistance genes *aac(6')-li_1* and *aac(6')-aph(2")_1* were found in every isolate. The tetracycline resistance genes *tetM(8)* and *tetM(13)* were only found in single isolates. The mean number of antibiotic resistance genes per outbreak isolate was 10. The screening isolates contained 17 different antibiotic resistance genes (Figure 4.9). However, in this case *msrC_1*, *aac(6')-li_1* and *dfrG_1* were found in every genome and the erythromycin resistance gene *erm(T)_2* that was found in a single isolate. Combining both datasets, a total of 19 antibiotic resistance genes were identified (Figure 4.9). Genes conferring resistance to the aminoglycoside class of antibiotics were found in all of the hospital isolates. Each isolate had between two and five aminoglycoside resistance genes in their genome. Resistance to erythromycin was also widespread with every isolate in the dataset having at least one erythromycin resistance gene. Five different alleles of the *tet(M)* gene were identified. *tet(M)_13* was found in one isolate, *tet(M)_12* in two isolates and *tet(M)_8* in five isolates. *tet(M)_6* and *tet(M)_10* were the most widespread of the *tet(M)* genes and were found in two thirds and one third of the isolates respectively. Genes which conferred resistance to the glycopeptide antibiotic vancomycin were common with 45 of the 60 isolates carrying *vanA*-type resistance gene. None of the isolates carried any of the other vancomycin resistance types.

No transferable linezolid resistance genes (*optrA*, *poxtA* or *cfr*) were found in any of the isolates. However, linezolid resistance can also be conferred by specific mutations in the 23S rRNA gene. Isolates OI36, HHI16 and HHI20 had a G2576T mutation in several copies of their 23S rRNA gene which has been shown to confer clinically relevant levels of

resistance to linezolid. The linezolid resistant isolates were all closely related and belonged to ST780.

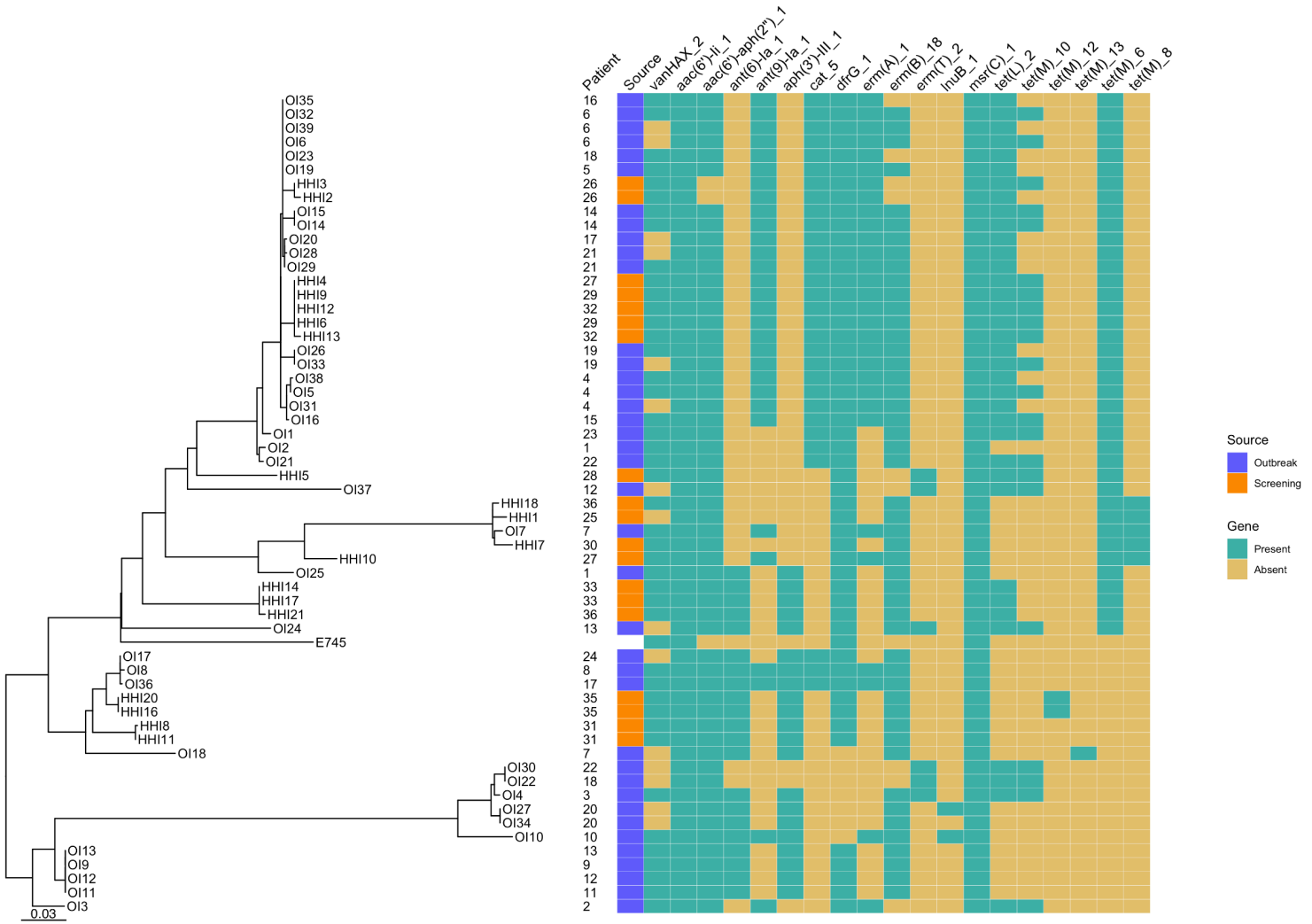

Figure 4.9: Maximum likelihood core genome phylogenetic tree of *E. faecium* isolates in a Birmingham hospital. Metadata includes the patient number, the source of the isolates (outbreak or routine screening) and the presence (green) or absence (gold) of antibiotic resistance genes in the isolates. *E. faecium* strain E745 was used as the reference strain. The scale bar indicates the average number of nucleotide substitutions per site.

4.4.10 Minimum inhibitory concentration assays confirmed susceptibility of isolates

Isolates HHI2, HHI3 and OI25 were identified as phenotypically susceptible to vancomycin by the VITEK 2 antibiogram system. However, when the isolates were short-read sequenced, *vanA*-type vancomycin resistance genes were present in the isolates. Minimum inhibitory concentration (MIC) assays against vancomycin were used to confirm the results of the antibiogram. Isolates HHI2 and OI25 had a vancomycin MIC of 1 μg/ml and isolate HHI3 had an MIC of 2 μg/ml, both of which fall below the EUCAST breakpoint of >4 μg/ml for *E. faecium*, confirming them as phenotypically susceptible.

4.4.11 Insertion of IS elements into the vancomycin resistance region led to a loss of vancomycin resistance in clinical VRE isolates

The short-read assemblies of HHI2, HHI3 and OI25 were heavily fragmented which obscured the structure of the vancomycin resistance operons. However, long-read sequencing was able to fully resolve the vancomycin resistance plasmids of isolates HHI2 and HHI3. The vancomycin resistance plasmid of isolate OI25 was not fully resolved, but it was possible to assemble the region surrounding the vancomycin resistance genes (Figure 4.10). Compared to the prototypical *vanA* transposon Tn*1546* (GenBank: M97297.1), isolate OI25 had a number of insertion sequences disrupting the vancomycin resistance region. Upstream of the *vanR* gene, there was an insertion of an IS*6* family transposase which disrupted both the previously characterised promoter region and the first 50 bp of the *vanR* gene. An IS*L3* family transposase was also present which disrupted the region upstream of the *vanH* gene. The vancomycin resistance region of isolates HHI2 and HHI3 were identical and had undergone extensive rearrangement upstream of *vanR* compared to Tn*1546*. Isolates HHI2 and HHI3 had an AAA-family ATPase gene which overlapped the *vanR* gene by 4 bp,

this did not disrupt the gene itself, but instead interrupted the previously characterised promoter region upstream of *vanR* (Holman *et al.*, 1994).

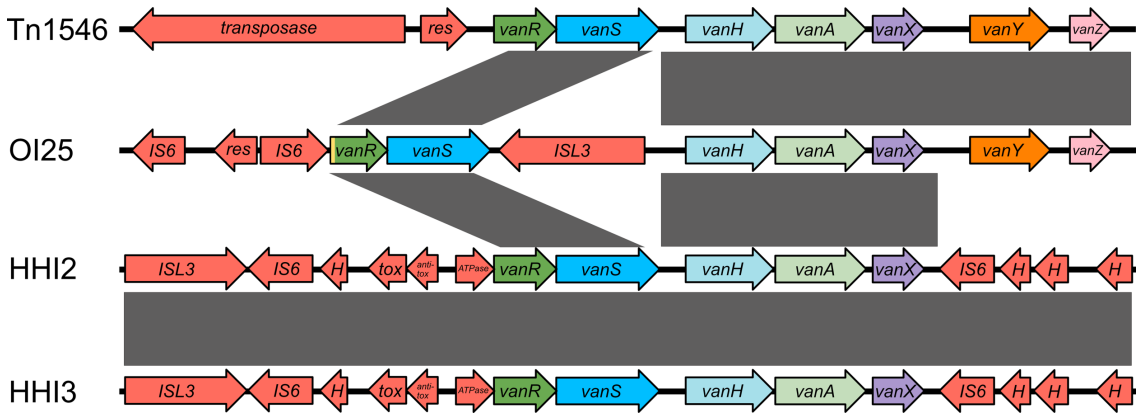

Figure 4.10: Alignment of the vancomycin resistance region of *E. faecium* isolates OI25, HHI2 and HHI3 against the vancomycin resistance region of the prototypical Tn*1546* transposon. Grey boxes represent regions which are identical between isolates. H = Hypothetical gene.

4.4.12 Disruption of *vanR* impairs the ability of the VVE isolates to respond to vancomycin

To understand the effect of the *vanR* gene disruptions on the vancomycin variable enterococci's ability to respond to vancomycin, RT-qPCR analysis was performed on the VVE isolates pre and post exposure to 8 µg/ml vancomycin (Figure 4.11). Expression of the *vanHAX* genes in E8202 (a clade A1, high-level vancomycin-resistant strain) increased 310-fold and the *vanRS* gene expression increased by 52-fold in the presence of vancomycin. In contrast, expression of the *vanHAX* genes in the VVE isolates only increased by a maximum of 26-fold and the maximum increase in gene expression for the *vanRS* genes was 5-fold.

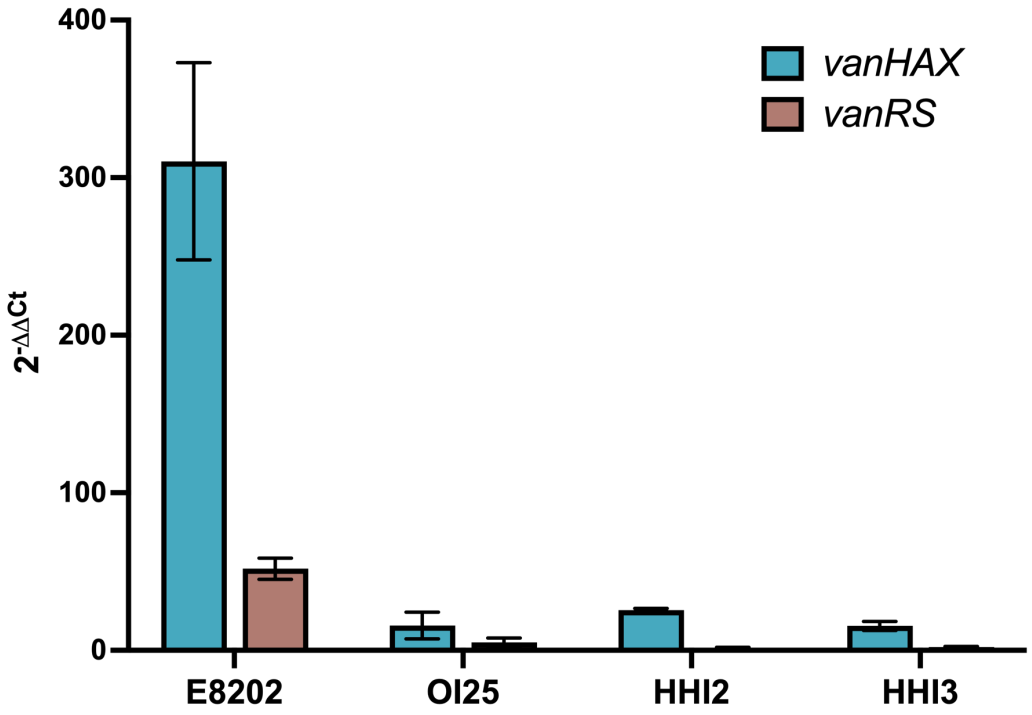

Figure 4.11: RT-qPCR analysis of the vancomycin resistance gene operons *vanHAX* and *vanRS* of the VVE isolates and E8202, before and after exposure to 8 µg/ml vancomycin. Expression data was normalised to the internal control gene *tufA*. Experiments were carried out with biological triplicates and technical duplicates. Error bars represent standard deviation.

4.4.13 Phenotypically susceptible isolates reverted to a vancomycin resistant phenotype under vancomycin selection

To investigate the potential for susceptible isolates to revert to a resistant phenotype, vancomycin variable isolates HHI2, HHI3 and OI25 were cultured in the presence of 8 μg/ml vancomycin. Growth was observed after 24 hours for HHI2 and HHI3 and after 48 hours for OI25. Two isolates (from this point forward referred to as revertants) were selected from each parent strain and were analysed in MIC assays against vancomycin. All six of the revertant isolates had regained high-level vancomycin resistance several fold higher than the parent strains (Table 4.2).

Table 4.2: MIC of parental VVE Isolates and their revertants.

E. faecium strain	MIC of parental strain (μg/ml)*	MIC of revertant (μg/ml)*
HHI2rev1	1	128
HHI2rev2	1	256
HHI3rev1	2	512
HHI3rev2	2	128
OI25rev1	1	512
OI25rev2	1	512

* Mode of 3 biological replicates

4.4.14 Like the wildtype VVE isolates, the revertant isolates have an impaired ability to sense vancomycin.

To test whether the response to vancomycin had been restored in the revertant isolates, RT-qPCR analysis was performed on the VVE revertant isolates pre- and post-exposure to 8 μg/ml vancomycin (Figure 4.12). The greatest change in expression was the *vanHAX* genes in HHI2rev1 which had an expression 8-fold higher when exposed to vancomycin. All of the other revertant isolates had expression changes lower than HHI2rev1, which suggested that vancomycin sensing was still impaired in the revertant isolates.

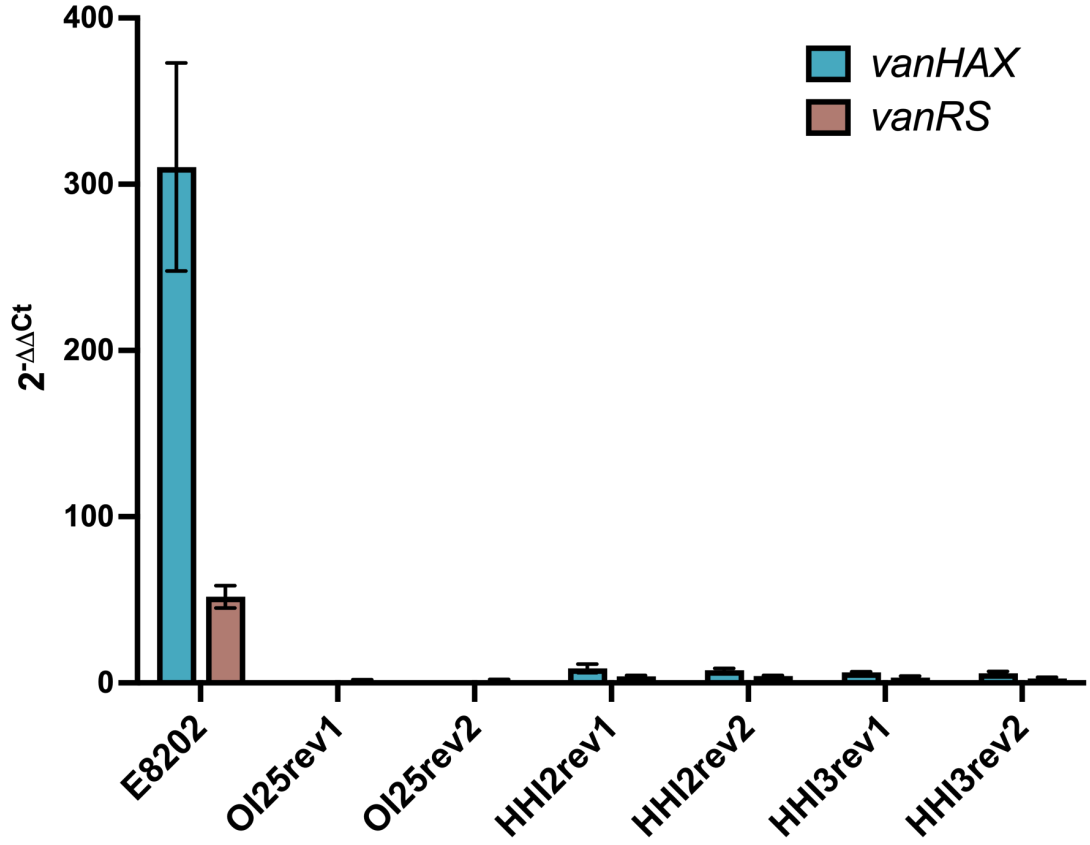

Figure 4.12: RT-qPCR analysis of the vancomycin resistance gene operons *vanHAX* and *vanRS,* of the revertant isolates and E8202 before and after exposure to 8 μg/ml vancomycin. Expression data was normalised to the internal control gene *tufA.* Experiments were carried out with biological triplicates and technical duplicates. Error bars represent standard deviation.

4.4.15 Expression of the *vanHAX* resistance genes in the OI25 revertant isolates is far greater than the wildtype

Although expression of the vancomycin resistance genes was still impaired in the presence of vancomycin, it was hypothesised that the genes were being constitutively expressed thus providing continuous, non-inducible resistance. RT-qPCR analysis was performed on the *vanRS* and *vanHAX* genes in the revertant isolates *versus* the wildtype isolates (Figure 4.13). Expression of the *vanHAX* genes was $2.4x10^4$-fold (OI25rev1) and $3.3x10^4$-fold (OI25rev2) greater in the OI25 revertant isolates compared to the wildtype isolate. Expression of the *vanRS* genes increased on average 27-fold in the revertant isolates with the exception of HHI2rev2 and HHI3rev1, where the expression increased by 19-fold and 9-fold respectively. The expression of the *vanHAX* genes in the HHI2 revertant isolates was 125-fold greater than the wildtype and 50-fold greater than the wildtype in the HHI3 revertant isolates.

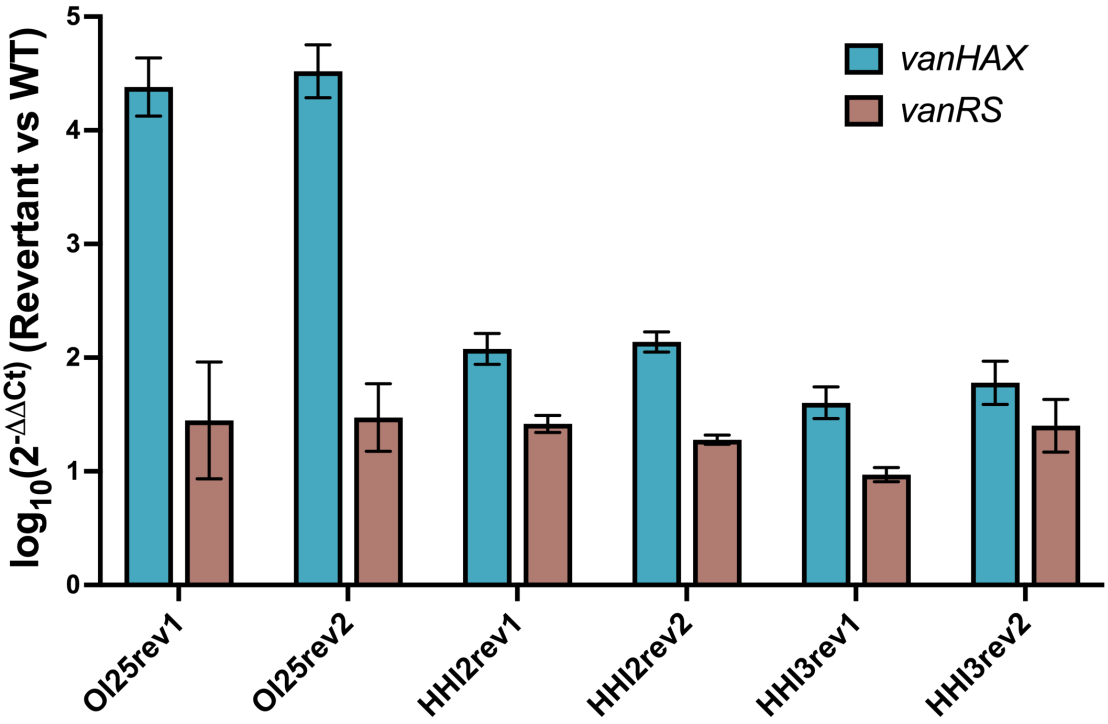

Figure 4.13: RT-qPCR analysis of the vancomycin resistance gene operons *vanHAX* and *vanRS* in the revertant isolates compared to the wildtype (WT) isolates. Expression data was normalised to the internal control gene *tufA* and values were log_{10} transformed. Experiments were carried out in biological triplicate and technical duplicate. Error bars represent standard deviation.

4.4.16 No variants were found in the revertant isolates that could explain the reversion to resistance

The revertant isolates were compared to their parent strain to identify any variants that could have caused the restoration of the vancomycin-resistant phenotype (Table 4.3). HHI2rev1 and HHI2rev2 both had a non-synonymous H279D mutation in the *gmuD_2* gene which encodes a 6-phospho-beta-glucosidase. HHI2rev1 had an additional SNP in an intergenic region. Both of the HHI3 revertant isolates had two synonymous mutations in a single copy of the IS*1216E* gene (I196I and G192G). In addition, HHI3rev2 had a SNP in an intergenic region. OI25rev1 and OI25rev2 had a P69L non-synonymous mutation in a copy of IS*1062*, they also had the insertion of an "AT" dinucleotide in the same intergenic position. Isolate OI25rev1 had a further non-synonymous mutation (T416K) in a hypothetical gene and insertion of a 23 bp fragment of DNA into an intergenic region. None of the variants are within or in proximity to genes known to be involved in high-level vancomycin resistance so are unlikely to explain the vancomycin-resistant phenotype of the revertants.

Table 4.3: Genomic variants in VVE isolates compared to the wildtype isolates.

Isolate	Variant type	Contig	Position	Gene	Base change	Effect
HHI2rev1	SNP	contig00006	8123	*gmuD_2*	C > G	H279D
	SNP	contig00048	9420	Intergenic	C > G	NA
	SNP	contig00115	23	Intergenic	T > C	NA
HHI2rev2	SNP	contig00006	8123	*gmuD_2*	C > G	H279D
	SNP	contig00115	23	Intergenic	T > C	NA
HHI3rev1	SNP	contig00094	4018	IS*1216E*	A > G	I196I
	SNP	contig00094	4030	IS*1216E*	G > A	G192G
HHI3rev2	SNP	contig00094	4018	IS*1216E*	A > G	I196I
	SNP	contig00094	4030	IS*1216E*	G > A	G192G
	SNP	contig00117	3410	Intergenic	T > G	NA
OI25rev1	SNP	contig00087	1888	IS*1062*	C > T	P69L
	Ins	contig00102	121	Intergenic	A > AT	NA
OI25rev2	SNP	contig00038	23874	HG	C > A	T416K
	SNP	contig00087	1888	IS*1062*	C > T	P69L

	Ins	contig00097	80	Intergenic	T > TTTTATCTACATCGTTTTGTCTG	NA
	Ins	contig00102	121	Intergenic	A > AT	NA

HG = Hypothetical gene
SNP = Single nucleotide polymorphism
Ins = Insertion

4.4.17 Insertion of the vancomycin resistance genes into the chromosome as an explanation for restoration of resistance

Long-read sequencing of the revertant isolates revealed a mechanism by which isolates could have reverted to vancomycin resistance. The vancomycin resistance plasmid of isolate OI25rev2 was inserted into the chromosome 27 bp downstream of a copy of the 5S rRNA gene (372,841 bp:372,951 bp) (Figure 4.14A). At each end of the inserted plasmid sequence, was an 8 bp target site duplication (ACTAGAAA). Similarly, isolate OI25rev1 had an insertion in the same site on the chromosome (Figure 4.14B). However, in this case it was not the complete plasmid but instead only a 15,299 bp fragment of DNA containing the vancomycin resistance genes that had inserted. The other 21,107 bp region that had not inserted into the chromosome, formed a complete plasmid sequence containing a plasmid replication gene. It is unclear whether the whole plasmid was inserted and then this latter fragment was excised, leaving behind the vancomycin resistance genes, or whether the stretch of DNA was excised from the plasmid, formed an intermediate, and then was inserted into the chromosome leaving behind the replicating plasmid.

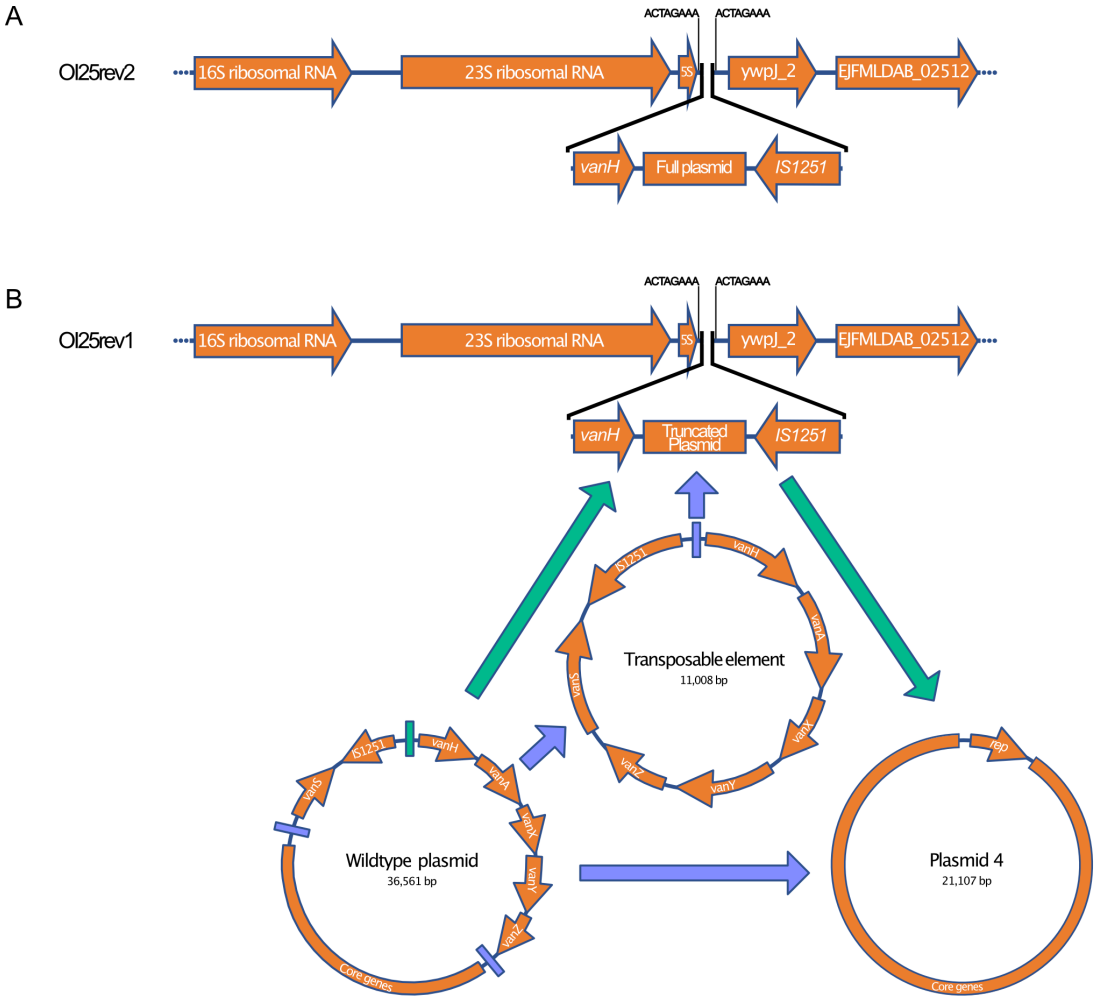

Figure 4.14: Mechanisms of VVE reversion to vancomycin resistance. **A.** Insertion of the vancomycin resistance plasmid into the chromosome of OI25rev2. **B.** Insertion of the vancomycin resistance genes into the chromosome of OI25rev1 and the possible intermediate stages in the insertion.

Unlike OI25, the revertant isolates of HHI2 and HHI3 had no changes in their vancomycin resistance regions. The vancomycin resistance regions of HHI2rev2 and HHI3rev1 were found on the same plasmids as in the parent strains. However, plasmid 5 of HHI2rev1 had inserted into the vancomycin resistance plasmid, but this did not affect the vancomycin resistance region (Figure 4.15A). The vancomycin resistance region of HHI3rev2 had transferred to plasmid 2, but again there were no changes to the region directly surrounding the vancomycin resistance genes (Figure 4.15B).

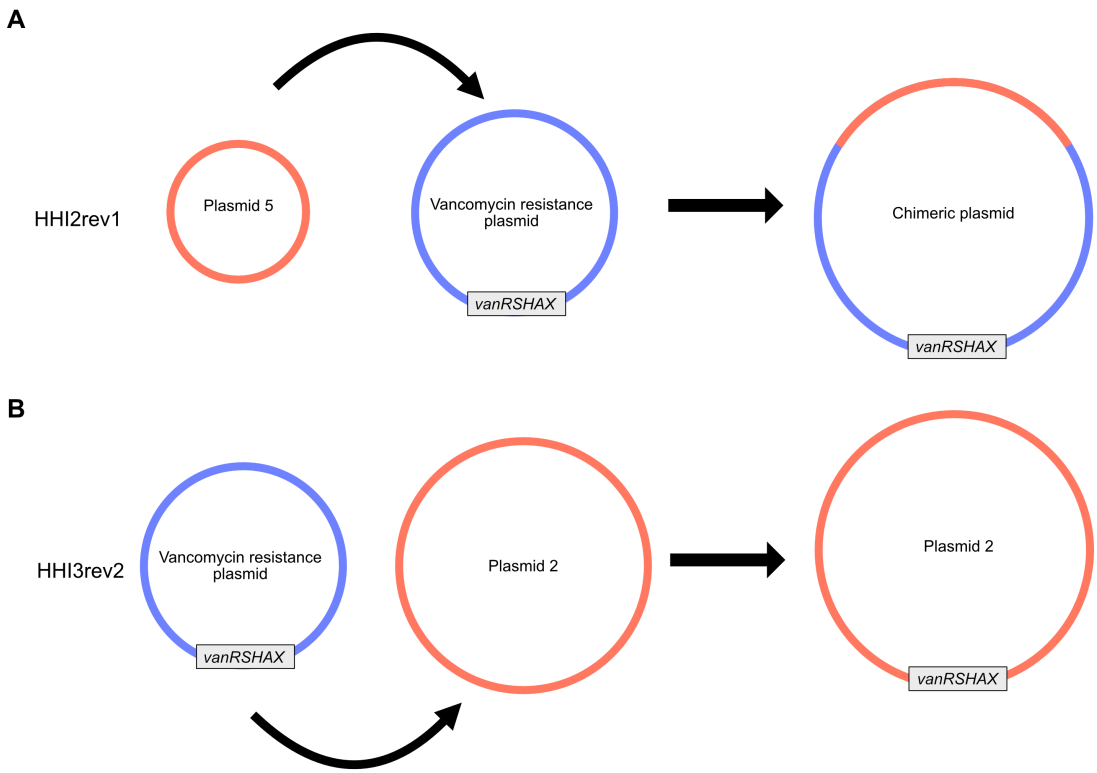

Figure 4.15: Genomic rearrangements in the revertant isolates. **A.** Insertion of plasmid 5 into the vancomycin resistance plasmid in isolate HHI2rev1. **B.** Insertion of the vancomycin resistance operons into plasmid 2 in isolate HHI3rev2.

4.4.18 Insertion of the vancomycin resistance genes into the chromosome may disrupt the 5S rRNA gene terminator

As the insertion of the vancomycin resistance genes in strains OI25rev1 and OI25rev2 was 27 bp from the predicted end of the 5S rRNA gene, I hypothesised that this insertion may disrupt the gene terminator. In the OI25-wildtype isolate, an A-tail, stem-loop and T-tail could be found within the first 50 bp of DNA downstream of the 5S rRNA gene, all of which were characteristic of a Rho-independent terminator (Kingsford, Ayanbule and Salzberg, 2007)(Figure 4.16). However, in mutants OI25rev1 and OI25rev2, the vancomycin resistance genes were inserted approximately halfway through the stem-loop structure. This insertion completely disrupted the structure of the terminator downstream of the 5S rRNA gene.

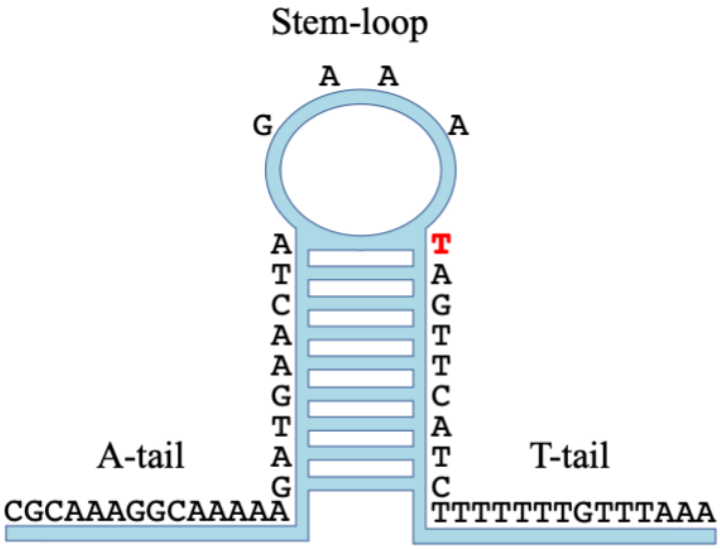

Figure 4.16: The predicted Rho-independent terminator of the 5S rRNA gene. The terminator was found within 50 bp of the predicted 5S rRNA gene end. The insertion site of the vancomycin resistance plasmid is highlighted in red.

4.4.19 PCR across the rRNA-*vanHAX* junction suggests that the vancomycin resistance genes are co-transcribed with the rRNA operon.

To test whether the vancomycin resistance genes were co-transcribed with the ribosomal RNA genes, reverse transcriptase-polymerase chain reaction (RT-PCR) was performed on cDNA extracted from the OI25 revertant isolates. Three PCR reactions were designed which bridged from the upstream 23S rRNA gene into *vanH*, *vanA* and *vanX* genes (Figure 4.17A). Amplicons were present for *vanH* (376 bp), *vanA* (1348 bp) and *vanX* (2366 bp) which suggested that the vancomycin resistance genes were transcribed with the rRNA genes (Figure 4.17B).

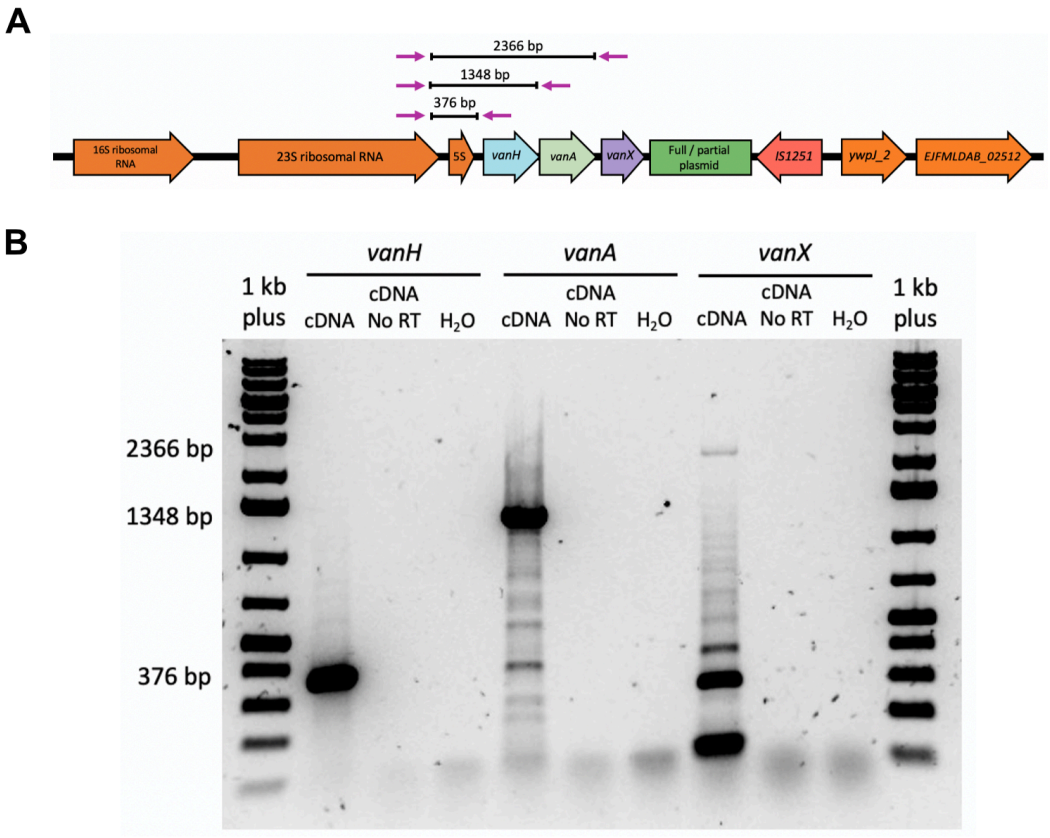

Figure 4.17: RT-PCR on the rRNA-*vanHAX* junction. **A.** Schematic showing the expected amplicon sizes. **B.** Agarose gel showing the amplicons with the expected products sizes from panel **A** indicated for the RT-PCR reactions between the 23S rRNA gene and *vanH* (376 bp), *vanA* (1348 bp) and *vanX* (2366 bp).

4.4.20 Vancomycin resistance could not be transferred to a susceptible host

To understand whether horizontal or vertical transmission would play a larger role in the onward dissemination of the newly acquired vancomycin resistance phenotype in the revertant isolates, conjugation assays were performed between isolate OI25rev1, OI25rev2, HHI2rev1, HHI2rev2, HHI3rev1, HHI3rev2 and the recipient strain *E. faecium* 64/3. No transconjugants were observed for any of the donor strains which suggested that either the vancomycin resistance genes were in a context that did not allow them to be transferred or that the rate of transfer was below the limit of detection for these assays.

4.5 Discussion

Enterococcus faecium is an opportunistic human pathogen that is a major cause of bacterial bloodstream and urinary tract infections (Zhou *et al.*, 2020). *E. faecium* is particularly well-suited to survive in the hospital environment due to its intrinsic and acquired resistance to antibiotic drugs, alcohols and its ability to survive in nutrient-limited conditions (de Maat *et al.*, 2020). This chapter uncovered the diversity of *E. faecium* in a hospital population and provided evidence for a novel mechanism by which a phenotypically susceptible *E. faecium* strain becomes resistant to vancomycin.

It was hypothesised that an outbreak of vancomycin resistant *E. faecium* in a Birmingham hospital was caused by the expansion of a single clone. However, eight different sequence types of *E. faecium* were identified that belonged to the hospital-associated clonal complex 17 (CC-17) (Willems *et al.*, 2005). The two most abundant sequence types identified in this study were ST262 and ST80. ST262 has been identified in hospitals within the UK (Lemonidis *et al.*, 2019) and Europe (Falgenhauer *et al.*, 2019; Egan *et al.*, 2022) but has not been linked to hospital outbreaks. In contrast, ST80 has been linked to VRE outbreaks in Ireland and Sweden and is now the predominant lineage in Stockholm (Egan *et al.*, 2020; Fang *et al.*, 2021). ST262 and ST80 remained the most prevalent sequence types in the screening isolates collected in 2019 and 2020 suggesting that the same *E. faecium* strains were being maintained within the hospital. A recent study which investigated *E. faecium* in a number of UK hospitals found that ST18 was the major sequence type followed closely by ST17 and ST203 (Gouliouris *et al.*, 2018). The major sequence types in this study, ST262 and ST80, were found at much lower levels with only a single isolate of ST262 identified in other UK hospitals. The lack of ST262 isolates in other UK hospitals suggests that ST262 may be more prevalent in Birmingham than elsewhere in the United Kingdom. The rise of

ST80 in Sweden and its prevalence in this study may suggest that it is an emerging major sequence type.

Several patients were colonised by more than one *E. faecium* sequence type over both long and short periods of time. In Patient 7, the *E. faecium* strain isolated from blood cultures switched from ST117 to ST80 over a period of two days and patient 13 had two different sequence types (ST412 and ST1478) present in their blood at the same time. Both of these cases suggest that *E. faecium* bacteraemia may in some cases be caused by multiple strains. Mixed infections with two *E. faecium* strains have been noted previously in UK hospitals, it has also been observed that mixed infections are often caused by a vancomycin-resistant strain and a vancomycin-sensitive strain (Raven *et al.*, 2018) which is also the case in the current study.

It is well established that due to its persistence in the environment *E. faecium* can be transmitted between patients in hospital wards (van Hal *et al.*, 2016; Gouliouris *et al.*, 2021). Patients 9, 11, 12 and 13 all carried ST1478 isolates that were identical at the core genome level. All of these patients were present on the same hospital ward between April and May 2017 which raises the possibility that they acquired the same *E. faecium* strain during their hospital stay. The possible acquisition of vancomycin resistant *E. faecium* strains within the hospital setting highlights the challenges of containing the spread of *E. faeecium* within hospitals. To prevent the transmission of *E. faecium* strains between patients, it is essential that the bed and surrounding area of a patient with a VRE infection is thoroughly and regularly disinfected. Hand-hygiene is also extremely important in the prevention of cross-contamination between patients. Re-training of staff in regard to hand-hygiene and the

monitoring of hand-hygiene compliance could also help to prevent the spread of VRE between patients in the clinical setting (Reyes, Bardossy and Zervos, 2016).

The high rate of horizontal gene transfer in *Enterococcus faecium* is mediated by a large repertoire of plasmids (Palmer, Kos and Gilmore, 2010). These plasmids contain *rep* genes that can be used to group the plasmids, allowing the number and diversity of the plasmids present to be explored. However, plasmids can contain multiple *rep* genes meaning that the number of *rep* genes is not necessarily the same as the number of plasmids. The clinical *E. faecium* isolates in this study carried on average four plasmid replication genes. The most common plasmid replication initiation gene (*rep*) type was *rep2_1_orf1*(pRE25). This plasmid is a theta-replicating plasmid that belongs to the Inc18 incompatibility group (Kohler, Vaishampayan and Grohmann, 2018). These plasmids typically have a broad host range and have been linked with the transfer of vancomycin resistance genes from enterococci to *Staphylococcus aureus* (Zhu *et al.*, 2010). The second most abundant *rep* type was *rep11a_1_repA*(pB82) which belongs to the Rep_3 family that are also theta-replicating plasmids. The Rep_3 plasmids are typically small plasmids that have a narrow host range, almost always being found in *E. faecium* rather than *E. faecalis*. Despite their narrow host range, these plasmids are potential carriers of tetracycline and aminoglycoside resistance genes between *E. faecium* strains (Clewell *et al.*, 2002; Wardal *et al.*, 2017). The final *rep* type that was found in >75% of isolates was *repUS15_2_repA*(pNB2354p1). This *rep* type belongs to the RepA_N family, which also has a very narrow host range. However, these plasmids have extremely high conjugation rates and carry antibiotic resistance genes such as *vanA* making them important for the transfer of antibiotic resistance within *E. faecium* (Weaver *et al.*, 2009). A recent study which investigated plasmids in clade A *E. faecium*

strains also found that the Rep_3 and RepA_N plasmids were amongst the most prevalent (Arredondo-Alonso *et al.*, 2020).

Patients which had matched vancomycin-resistant and -susceptible isolates allowed the identification of possible cases of horizontal gene transfer. Upon inspection of the plasmid profile of patient 21's isolates, it was found that their vancomycin resistant rectal swab isolate contained a *rep17_1_CDS29*(pRUM) plasmid that their bloodstream isolate did not. The presence of the *rep17_1_CDS29*(pRUM) plasmid and the vancomycin resistance genes in the same isolate suggests that either horizontal gene transfer had occurred within the patient or the plasmid was lost. A similar situation occurred with patient 19, where the antibiotic resistance and plasmid profiles indicated the presence of a *repUS7_1_rep*(pHTbeta) plasmid carrying the vancomycin resistance genes in the rectal swab isolate but not the bloodstream isolate. The prototypical plasmid of this *rep* gene type, plasmid pHTβ, is a highly conjugative plasmid that carries the Tn*1546* vancomycin resistance transposon (Tomita *et al.*, 2003). Two isolates cultured from patient 32 14 months apart were near identical (3 core SNPs different) at the core genome level but had lost a *repUS7_1_rep*(pHTbeta) plasmid, demonstrating either that the same strain was maintained over a long period of time, and had a dynamic accessory genome or that the patient had been recolonised with the same strain that had evolved outside of the body.

Vancomycin resistance genes are typically found on plasmids larger than 20 kbp and generally cannot be resolved via short-read sequencing alone (Arredondo-Alonso *et al.*, 2017). To overcome this issue, a selection of the vancomycin resistant *E. faecium* isolates were long-read sequenced. Of the 16 isolates that contained vancomycin resistance genes, the genetic context could only be determined in four isolates. The vancomycin resistance

plasmids belonged to three different *rep* types which included *rep17_1_CDS29*(pRUM), *rep17_2_repA*(AUS0004p1) and *rep18a_1_repA*(p200B). Plasmids belonging to *rep17_1_CDS29*(pRUM) and *rep17_2_repA*(AUS0004p1) belong to the wider RepA_N family of theta-replicating plasmids. The *rep18a_1_repA*(p200B) plasmids are also theta-replicating plasmids but belong to the Rep_3 family. As discussed earlier, both the RepA_N and Rep_3 plasmids have previously been linked to vancomycin gene carriage. Consistent with other pRUM-like plasmids, pOI5 did not contain conjugation machinery that would allow it to horizontally transfer. However, unlike the prototypical pRUM plasmid (Genbank: NC_005000.1), pOI5 lacks a mobilisation gene which further restricts it ability to transfer between strains. It is likely that vancomycin resistance is transferred vertically to daughter cells in this case rather than horizontally. Plasmids pHHI2 and pHHI3 have two plasmid replication genes suggesting that they have the ability to replicate in a broad range of hosts. These plasmids also possess conjugation and mobilisation genes which may confer the ability to transfer horizontally to other *E. faecium* strains within the gut or within the hospital environment.

The short-read sequencing data highlighted cases of potential in-patient horizontal gene transfer and identified plasmid *rep* types that are highly conjugative and often associated with vancomycin resistance genes. I was also able to fully resolve two vancomycin resistance plasmids, pHHI2 and pHHI3, that contained putative conjugation machinery and a mobilisation gene. The in-patient acquisition of vancomycin resistance plasmids is a cause for concern as previously treatable *E. faecium* infections could become resistant to treatment. Drugs are currently in development that can prevent plasmid conjugation or remove plasmids altogether, these could be used in the future to prevent the acquisition of resistance plasmids by susceptible strains (Buckner *et al.*, 2020; Palencia-Gándara *et al.*, 2021).

Two of the main factors that make *E. faecium* a successful nosocomial pathogen is its intrinsic resistance to antibiotic drugs and its genome plasticity (Guzman Prieto *et al.*, 2016). Aminoglycoside resistance is widespread in *Enterococcus faecium* and this is particularly true for determinants which confer resistance to gentamicin. The aminoglycoside resistance gene *aac(6')-Ie-aph(2")-Ia* is the most widespread of the aminoglycoside resistance genes in *E. faecium*, conferring resistance to all clinically available aminoglycosides with the exception of streptomycin (Sparo, Delpech and García Allende, 2018; Chen *et al.*, 2021). *aac(6')-Ie-aph(2")-Ia* was found in 58 of the 60 isolates in this study. *AAC(6')-Ii* was the most widespread aminoglycoside resistance gene in my dataset being found in all 60 isolate. This chromosomal gene is specific to *E. faecium* and confers resistance to amikacin and certain forms of gentamicin (Costa *et al.*, 1993). All of the isolates contained more than one aminoglycoside resistance gene and several isolates contained five. The carriage of multiple aminoglycoside resistance genes is a hallmark of clinical *E. faecium* strains (Chow, 2000). As well as the aminoglycoside resistance genes, tetracycline resistance genes were also highly prevalent in the isolates. Five different alleles of *tet(M)* which encodes a ribosomal protection protein were found across the isolates with the most prevalent allele being *tet(M)_6* (Connell *et al.*, 2003). A tetracycline efflux gene, *tet(L)* was also found in around two thirds of isolates (McMurry *et al.*, 1987). Vancomycin resistance is conferred by at least nine different resistance operons designated *vanA*, *vanB*, *vanC*, *vanD*, *vanE*, *vanG*, *vanL*, *vanM* and *vanN* (Ahmed and Baptiste, 2018). All 45 vancomycin resistance genes identified in this study belonged to *vanA*-type resistance which confers high-level resistance to vancomycin with a MIC typically greater than 64 μg/ml (Cetinkaya, Falk and Mayhall, 2000). Three linezolid resistant isolates were identified in this study that have the same G2576T mutation identified in the first linezolid resistant UK isolates (Auckland *et al.*,

2002). A study that took place in the Queen Elizabeth University Hospital (Birmingham, United Kingdom) in 2015 also exclusively identified the G2576T mutation among linezolid-resistant *E. faecium* collected from their Liver Unit (Niebel *et al.*, 2016).

The widespread resistance to gentamicin and vancomycin in this study highlights the importance of performing antibiotic susceptibility testing on all clinical *E. faecium* isolates. Good antibiotic stewardship, including the use of the correct antibiotics and frequent reviews of treatment will help to reduce the development and spread of antibiotic resistance in the clinical setting (Ha, Haste and Gluckstein, 2017). The use of last-line antibiotics such as linezolid will also become increasingly important for the treatment of *E. faecium* infections as clinical strains become increasingly resistant to vancomycin (Narayanan *et al.*, 2019). It should be noted, however, that resistance to linezolid is also increasing in *E. faecium* strains in UK and Irish hospitals (Inkster *et al.*, 2017; Egan *et al.*, 2020). The development of new antibiotics that can treat infections caused by *E. faecium* resistant to last-line antibiotics will be critical for the management of *E. faecium* infections in the future.

There were three isolates that possessed vancomycin resistance genes but were phenotypically susceptible to vancomycin (OI25, HHI2 and HHI3). The vancomycin resistance regions of these isolates contained a number of insertions and deletions compared to the prototypical Tn*1546* transposon (Arthur *et al.*, 1993). The wildtype OI25 isolate had an insertion of an IS*L3* family element between the *vanS* and *vanH* genes. However, the *vanH* promoter, including the two VanR binding sites (TTTTAGGAAAAT and TCTTAGGAAATT), was intact. *E. faecium* strain V1836 plasmid pHVH-V1836-2 has an identical insertion to isolate OI25 but in the case of strain V1836 it is phenotypically resistant to vancomycin (Rubin *et al.*, 2020). This suggests that insertion of the IS*L3* family element

between *vanS* and *vanH* did not make the strain phenotypically susceptible to vancomycin. In addition to the IS*L3* insertion, the promoter region and the first 50 bp of the *vanR* gene were disrupted by the insertion of an IS*1216* family element. A similar truncation of the *vanR* promoter/*vanR* gene is seen in the vancomycin variable *E. faecium* strain VVE*Swe,* where it is thought to inactivate the vancomycin resistance operon (Wagner *et al.*, 2021). The disruption of the response regulator meant that the isolate could no longer respond to vancomycin and in turn the *vanHAX* genes which confer resistance were not transcribed in the presence of vancomycin, thus leading to a susceptible phenotype.

Upon exposure to vancomycin, the plasmid carrying the vancomycin resistance genes in isolate OI25 was inserted into the chromosome and the strain reverted to a vancomycin resistant phenotype. An 8-bp target site duplication was present at each end of the inserted plasmid that is consistent with the action of IS*1251* which is found at the 3' end of the inserted DNA (Handwerger *et al.*, 1995). This would be the first evidence of IS*1251* inserting a plasmid into the chromosome through a currently uncharacterised mechanism. Insertion of the plasmid into the chromosome of OI25 did not restore the sequence of the *vanR* gene or its promoter and also removed the VanR-P binding regions from the *vanH* promoter. This suggested that vancomycin sensing was not restored in the revertant isolates. RT-qPCR confirmed this hypothesis and showed that the increase in expression of the *vanRS* and *vanHAX* genes when exposed to vancomycin was actually less than the wildtype OI25 strain. In RT-qPCR experiments that compared the expression of the *vanHAX* genes in wildtype OI25 isolate and the OI25 revertant isolates, I found that the expression was $3.3x10^4$-fold greater in the revertant isolates. These data suggest that the *vanHAX* genes in the revertant isolates were being expressed constitutively and at extremely high levels. Insertion of the vancomycin resistance plasmid into the chromosome of the revertant isolates

occurred 27 bp downstream of the predicted 5S rRNA gene stop codon and disrupted the putative Rho-independent terminator. PCR reactions which bridged between the 23S gene and *vanHAX* genes on cDNA from the OI25 revertant isolates revealed that the vancomycin resistance genes are likely co-transcribed with the ribosomal RNA genes. However, it was not possible to rule out that the insertion of the vancomycin resistance plasmid into the chromosome had instead created a new constitutive promoter upstream of the *vanHAX* genes. No precedent can be found for this mechanism of reversion in the literature.

HHI2 and HHI3 also possessed vancomycin resistance genes but were phenotypically susceptible to vancomycin. In this case, an AAA-family ATPase gene overlapped the *vanR* gene by 4 bp disrupting the promoter region. Comparison of this region against the NCBI Genbank database revealed no other sequences with this configuration. As with isolate OI25, it was hypothesised that disruption of the *vanR* promoter impaired the ability of the isolates to sense vancomycin which was confirmed by RT-qPCR. Exposure of HHI2 and HHI3 to vancomycin also caused a reversion to high-level vancomycin resistance and like isolate OI25, RT-qPCR confirmed that the isolates did not regain their ability to sense vancomycin. In this case, short-read and long-read sequencing were unable to identify a possible mechanism behind the reversion to resistance. Revertant isolates were not cultured in the presence of vancomycin prior to sequencing which raised the possibility that they reverted back to the susceptible state prior to genome sequencing. This would be in line with previous work on VVE isolates, which lost resistance when not cultured in the presence of vancomycin (Wagner *et al.*, 2021).

Vancomycin-variable *Enterococcus faecium* pose a new challenge for the treatment of infections caused by *E. faecium*. Standard antibiotic susceptibility testing does not identify

VVE isolates due to their phenotype-genotype disagreement. Treatment of VVE-infected patients with vancomycin likely provides the selection pressure needed for the isolates to revert to a high-level vancomycin-resistant phenotype within the patient. We have shown that this reversion can occur rapidly, highlighting that treatment of VVE infections with vancomycin is likely to fail in clinical practice. VVE isolates currently only make up a small proportion of clinical *E. faecium* isolates but this may increase as these isolates escape treatment with vancomycin (Viswanath *et al.*, 2022). Should VVE isolates make up a significant proportion of clinical *E. faecium* isolates in the future, it may be necessary to implement sequencing approaches to screen for VVE isolates in order to provide the correct treatment.

4.5.1 Future work

To gain a full understanding of the spread of *E. faecium* within Heartlands hospital a further study could be performed which focuses on the carriage of *E. faecium* by staff and the hospital environment. Rectal swabs of inpatients and staff and swabs from the hospital environment could be taken over a defined time period to track the movement and horizontal gene transfer of *E. faecium* isolates within the hospital. Sequencing of multiple isolates per swab would also allow the investigation of the within-person and environmental diversity of *E. faecium*. The data in this thesis suggested that the *vanHAX* genes in the OI25 VVE revertant isolates were being co-transcribed with a set of ribosomal RNA genes. Primer extension assays or 5' RACE could be used to identify the transcription start site of the *vanHAX* genes and provide further evidence that they are co-transcribed with the rRNA genes. Further work is needed to determine the mechanism behind the reversion of isolates HHI2 and HHI3 to a highly vancomycin-resistant phenotype. These revertant isolates should be cultured in the presence of vancomycin prior to long-read sequencing to ensure that the

isolates do not revert back to a vancomycin-susceptible phenotype. Work is also needed to understand why the majority of vancomycin resistance plasmids cannot be completely assembled and associated with plasmid *rep* genes. The vancomycin resistance plasmids could be transferred to an *E. faecium* strain such as *E. faecium* 64/3 that does not contain plasmids. This would determine whether IS elements or other regions of homology on plasmids in the parental strain prevent the vancomycin resistance plasmids from being fully assembled.

4.6 Key Findings

- Several sequence types were circulating in the hospital during the period of the outbreak demonstrating that it was not clonal expansion that caused the outbreak.
- Antibiotic resistance was highly prevalent in these clinical *E. faecium* isolates, particularly aminoglycoside and erythromycin resistance which was found in all isolates. Vancomycin resistance was also common with 45 of the 60 isolates carrying *vanA*-type resistance genes.
- Plasmid replication initiation genes were identified for all of the major *rep* gene families of *E. faecium*. *rep2_1_orf1*(pRE25) was the most prevalent *rep* type and was identified in 52 of the 60 isolates. Vancomycin resistance genes were linked to a plasmid replication gene in five isolates. The genes were linked to the *rep17_1_CDS29*(pRUM) gene in two isolates, the *rep17_2_repA*(AUS0004p1) gene in two isolates and the *rep18a_1_repA*(p200B) gene in a single isolate. Hybrid assembly was not able to link vancomycin resistance genes and plasmid replication genes in the majority of isolates.
- Three of the isolates were phenotypically susceptible to vancomycin but genotypically resistant. Upon exposure to vancomycin these isolates reverted to a high-level vancomycin resistant phenotype within 48 hours.
- The mechanism behind the phenotypic reversion could only be determined in the OI25 revertant isolates. The vancomycin resistance genes were inserted into the chromosome downstream of a ribosomal RNA operon. A high constitutive expression level and mRNA transcripts across the ribosomal RNA – *vanHAX* gene junction suggested that the vancomycin resistance genes were co-transcribed with the ribosomal RNA genes.

Chapter 5
Identification of genes linked to vancomycin resistance in *Enterococcus faecium*

5.1 Introduction

Vancomycin is a glycopeptide drug that binds to the terminal D-alanyl D-alanine motif of the stem peptide in the peptidoglycan precursor lipid II. Binding of vancomycin prevents the crosslinking of peptidoglycan by transglycosylases leading to the disruption of cell wall integrity (Stogios and Savchenko, 2020). As discussed in Chapter 1, the genes that confer resistance to vancomycin in *E. faecium* are well-characterised. A two-component system encoded by genes *vanR* and *vanS* senses the presence of vancomycin and then activates the *vanHAX* operon (Guffey and Loll, 2021). These genes encode a D-specific alpha-keto acid dehydrogenase, D-alanyl D-lactate ligase and D-alanyl D-alanine dipeptidase, that replace the terminal D-Ala-D-Ala motif of the peptidoglycan precursor with a D-Ala-D-Lac motif reducing the binding affinity of vancomycin 1000-fold (Stogios and Savchenko, 2020). An additional gene, *vanY*, is a D-alanyl D-alanine carboxypeptidase that targets the D-Ala-D-Ala motif on lipid II thus reducing the amount of native lipid II that can be incorporated into the cell wall (Wright *et al.*, 1992). It has been shown in *E. faecium* that the response to the cell wall targeting antibiotic ampicillin involves a number of chromosomal genes not typically associated with ampicillin resistance (Xinglin Zhang *et al.*, 2012). This suggests that the response to vancomycin may also involve hitherto unidentified chromosomal genes that are associated with the cell wall and cell envelope in *Enterococcus faecium.*

In this chapter, I used functional genomics (Tn-seq) and transcriptomics (RNA-seq) to uncover genes involved in the response to a sub-lethal concentration of vancomycin in vancomycin resistant *Enterococcus faecium*. RNA-seq was used to identify genes that were differentially expressed in the presence of vancomycin and Tn-seq was used to identify genes which contributed to fitness in the presence of vancomycin. Both RNA-seq and Tn-seq were performed on two clinical *vanA*-carrying strains (E745 and E8202) to identify

similarities and differences in the transcriptional response and genes contributing to fitness during vancomycin challenge between different *vanA*-carrying strains. A third *E. faecium* strain (Aus0004) carrying the *vanB* vancomycin resistance genes was also included in the Tn-seq experiment to investigate whether the same accessory genes conferred a fitness advantage in the presence of vancomycin as in the *vanA*-carrying strains.

5.2 Hypothesis

Although the genes that confer resistance to vancomycin are well known, there are likely other chromosomal genes which are also involved in the response to vancomycin and may contribute to high-level resistance. Genes associated with the bacterial cell wall are likely to be involved in the resistance to vancomycin.

5.3 Aims and objectives

1. To use RNA-seq to identify genes that are differentially expressed in vancomycin resistant *E. faecium* strains exposed to ½ MIC vancomycin.
2. To use transposon insertion sequencing (Tn-seq) to identify genes that contribute to fitness in *E. faecium* when exposed to ½ MIC vancomycin.
3. To describe the phylogenetic distribution of genes identified in the Tn-seq analysis.
4. To validate the Tn-seq data by deleting the identified genes in *E. faecium* E745.

5.4 Results

RNA-seq Analysis

RNA-seq was performed on two *vanA*-type vancomycin resistant *E. faecium* strains, E745 and E8202, to identify genes which were significantly differentially expressed when exposed to vancomycin at ½ MIC in three independent experiments. Identification of genes whose expression is highly upregulated or highly downregulated will give novel insights into the response to vancomycin in *E. faecium* strains that carry the *vanA*-type resistance genes.

5.4.1 RNA-seq reads were successfully mapped to the parent genome

The sequenced RNA-seq libraries were demultiplexed to give an average of 35,934,512 paired-end reads per sample after quality trimming (Table 5.1). To count the number of reads mapping to a particular transcript, the reads were pseudo-aligned to the genomes of E745 and E8202 with an average of 88% of reads successfully aligned per sample (Table 5.1).

Table 5.1: Quality metrics of the RNA-seq data.

Sample[a]	Total reads	Pseudo-aligned reads	Reads aligned (%)
E745 1	33887101	29207902	86.2
E745 2	34039886	29733870	87.4
E745 3	29815216	26092101	87.5
E745 ½ MIC 1	38289274	33378981	87.2
E745 ½ MIC 2	36901461	32243753	87.4
E745 ½ MIC 3	36728793	33057319	90.0
E8202 1	41245763	36524184	88.6
E8202 2	31415673	27627231	88.0
E8202 3	40755851	36898736	90.5
E8202 ½ MIC 1	36252389	32085432	88.5
E8202 ½ MIC 2	38324307	34015738	88.8
E8202 ½ MIC 3	33558428	29613155	88.2

[a] Samples were collected from cultures at OD_{600} = 1 in the presence and absence of ½ MIC vancomycin (128 ug/ml for E745 and 64 ug/ml for E8202).

5.4.2 Replicate RNA-seq samples cluster together

Principal component analysis identified that the transcriptional profiles of the vancomycin treated samples were very different to the control samples for both E745 and E8202 (Figure 5.1). All three of the E8202 vancomycin-treated replicates were tightly clustered together and two of the control samples were tightly clustered with the third spaced further away (Figure 5.1B). The E745 replicates were more loosely clustered suggesting more variance in the transcriptome of E745 compared to E8202 (Figure 5.1A).

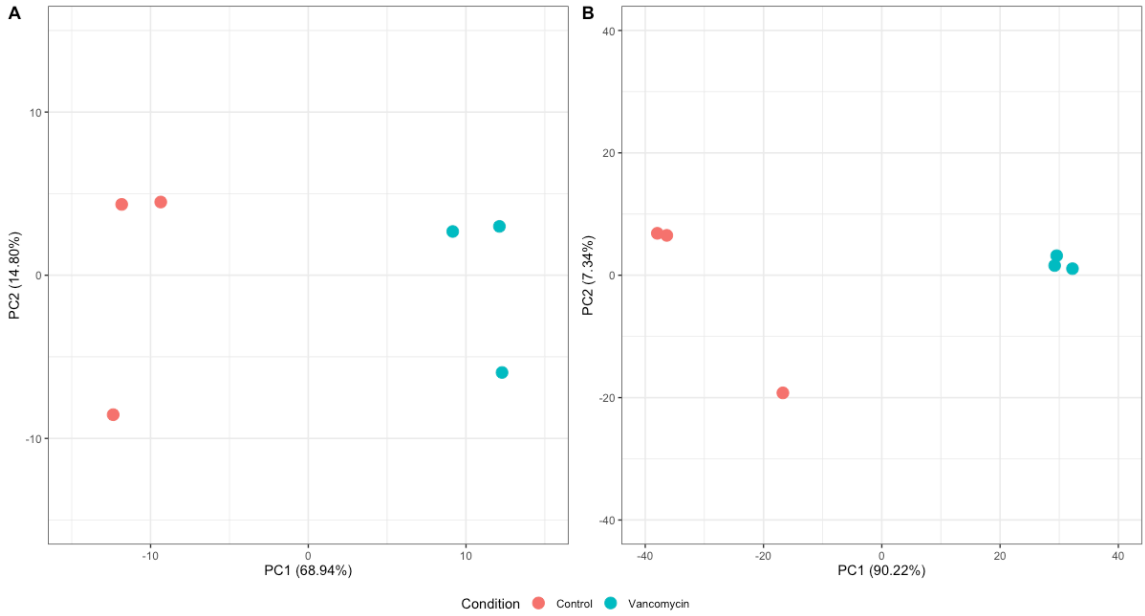

Figure 5.1: Principal component analysis of RNA-seq samples. **A.** E745 samples. **B.** E8202 samples. Control samples are represented by red circles and vancomycin treated samples are represented by blue circles.

5.4.3 Exposure to vancomycin led to widespread transcriptional changes in *E. faecium* E745 and E8202

Genes with a $\log_2$ fold-change ≤ -1 or ≥ 1 compared to the control and an FDR (Benjamini-Hochberg) < 0.05 were considered to be differentially expressed. Of the 3182 genes found in *E. faecium* E745, 315 genes were differentially expressed when exposed to ½ MIC vancomycin. 229 of these genes were upregulated and 86 were downregulated (Figure 5.2A). *E. faecium* E8202 had 962 genes that were differentially expressed out of 3271 when exposed to vancomycin (Figure 5.2B). 368 of the differentially expressed genes were upregulated and 594 of the genes were downregulated.

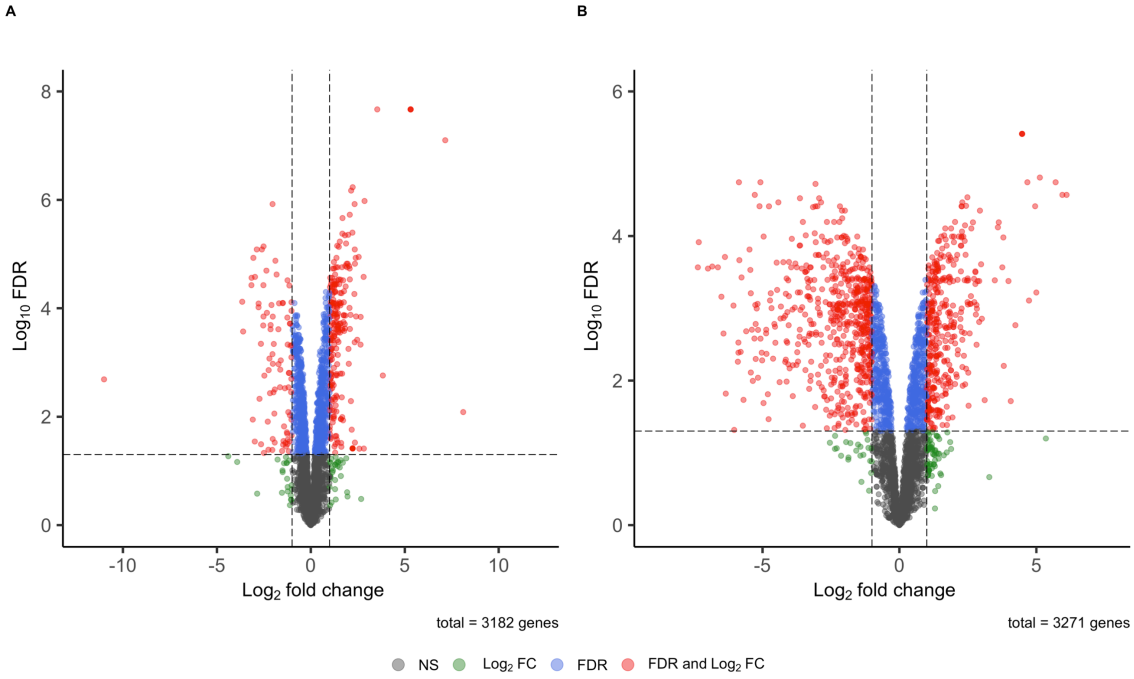

Figure 5.2: Volcano plots of RNA-seq expression data. **A.** *E. faecium* E745 samples. **B.** *E. faecium* E8202 samples. Genes that are not significant (NS) (FDR (Benjamini-Hochberg) > 0.05) and have a $\log_2$ fold-change (FC) between -1 and 1 are represented by grey circles. Genes with a $\log_2$ fold-change (FC) < -1 or > 1 but are not significant (FDR > 0.05) are represented by green circles. Genes that are significant (FDR < 0.05) but have a $\log_2$ fold-change between -1 and 1 are represented by blue circles. Significant genes (FDR < 0.05) with a $\log_2$ fold-change (FC) < -1 or > 1 are represented by red circles.

5.4.4 Differentially expressed genes are shared between *E. faecium* E745 and E8202

The differentially expressed genes were compared between *E. faecium* E745 and E8202 to identify genes involved in the response to vancomycin across different *E. faecium* strains. Of the genes that were upregulated in both E745 and E8202, 59 genes were shared (Figure 5.3A). An additional 20 upregulated genes were shared but these were all identical copies of the IS*1216* transposase. E745 shared 35% of its upregulated genes while E8202 shared 21% of its upregulated genes. The number of downregulated genes shared between E745 and E8202 was fewer with only 20 genes shared between the strains (Figure 5.3B). E745 shared 23% of its downregulated genes and E8202 shared far fewer downregulated genes with only 3% of its downregulated genes shared.

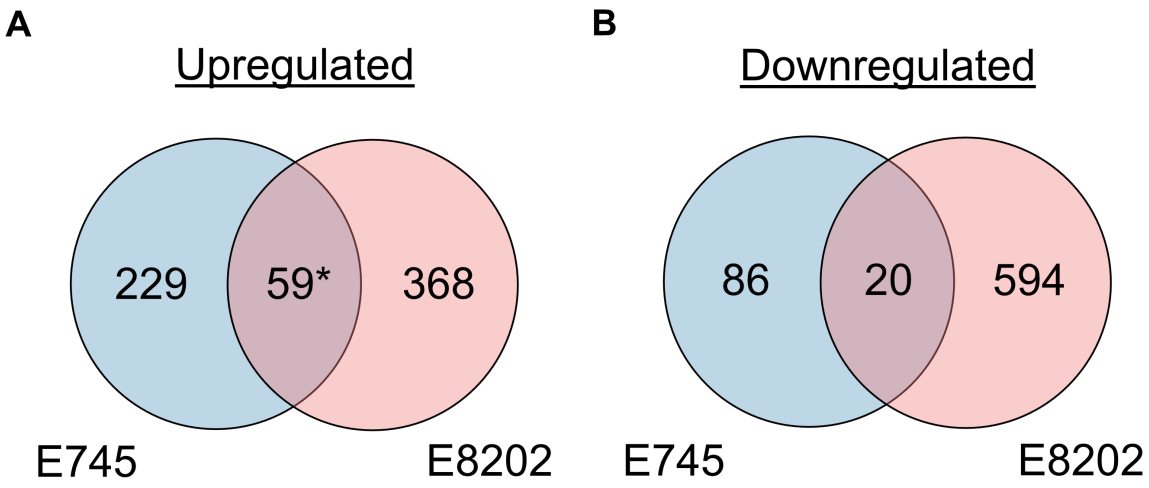

Figure 5.3: Differentially expressed genes shared between *E. faecium* E745 and E8202. **A.** Number of upregulated genes which are shared between E745 and E8202. **B.** Number of downregulated genes which are shared between E745 and E8202. *20 copies of IS*1216* were shared between E745 and E8202, these were not included as mapping was not accurate with identical genes.

5.4.5 Genes with upregulated expression that were shared between E745 and E8202

The vancomycin resistance genes *vanRSHAX* were among the genes with the largest fold-change relative to the control for both E745 and E8202 (Table 5.2). Expression of the *vanHAX* genes, which are essential for vancomycin resistance (Arthur *et al.*, 1998), was increased on average 131.5-fold in E745 and 61.0-fold in E8202 when compared to the relevant control. The vancomycin response regulator genes *vanRS* were not as highly expressed as the other vancomycin resistance genes, with expression on average 3.9-fold higher than the control in E745 and 28.2-fold higher in E8202. Both E745 and E8202 lack the *vanY* and *vanZ* genes found in the prototypical Tn*1546* transposon (Arthur *et al.*, 1993). However, E745 has a *vanZ*-like gene on its vancomycin resistance plasmid that had highly upregulated expression (65.8-fold) and E8202 had a *vanZ*-like gene on the chromosome that also had upregulated expression, but to a lesser extent (4.0-fold).

Table 5.2: Expression of the vancomycin resistance genes in E745 and E8202.

Gene	Product	Fold Change E745	Fold Change E8202
vanR	Regulatory protein	4.2	31.0
vanS	Sensor kinase	3.5	25.5
vanH	D-specific alpha-keto acid dehydrogenase	128.1	68.9
vanA	D-alanyl-D-lactate ligase	123.4	61.9
vanX	D-alanyl-D-alanine dipeptidase	142.9	52.1
vanY	D-alanyl-D-alanine carboxypeptidase	NP[†]	NP[†]
vanZ[*]	Unknown	65.8	4.0

* a *vanZ*-like gene is found on the same plasmid as the other vancomycin resistance genes in E745. A *vanZ*-like gene is found on the chromosome in E8202.
[†] NP: not present.

A cluster of 12 upregulated genes was shared between E745 and E8202 (Table 5.3). The first gene in the gene cluster encoded *tRNA-Met*, this single gene did not belong to an operon. The following two genes in the gene cluster formed an operon encoding two hydrolase genes. The next 5 genes encoded a general stress protein and genes involved in ion transport and

metabolism. The remaining genes all encoded stress response proteins suggesting that the gene cluster is important in the response to exogenous stressors.

Table 5.3: Gene cluster 1 in E745 and E8202 that exhibited upregulated expression in the presence of vancomycin.

Gene*	Product	Fold Change E745	Fold Change E8202
EfmE745_01661	tRNA-Met	3.1	4.2
EfmE745_01662	Cysteine hydrolase	4.0	3.6
EfmE745_01662	HAD superfamily hydrolase	4.2	3.9
EfmE745_01663	General stress protein A	5.1	8.1
EfmE745_01664	Glycosyl transferase family protein	7.0	6.7
EfmE745_01665	Iron-sulfur cluster biosynthesis family protein	5.3	6.8
EfmE745_01666	Mg2+ cation transporter CorA family protein	5.2	7.5
EfmE745_01667	Short chain dehydrogenase/reductase family oxidoreductase	4.6	4.9
EfmE745_01668 (amaP)	Alkaline shock response membrane anchor protein	2.3	3.6
EfmE745_01669	Small integral membrane protein	2.1	3.7
EfmE745_01670	Asp23/Gls24 family envelope stress response protein	2.9	3.6
EfmE745_01671	GlsB/YeaQ/YmgE family stress response membrane protein	2.6	3.6

*Gene/locus tag in *E. faecium* E745

The expression of a small three gene operon was upregulated in both E745 and E8202 (Table 5.4). It contained a gene which encoded a PadR family transcriptional regulator, a DUF1700 domain-containing protein and a hypothetical protein. In E745, the DUF1700 protein encoding gene had the largest increase in expression in the operon (2.9-fold) but in E8202 it was the transcriptional regulator (3.1-fold).

Table 5.4: Gene cluster 2 in E745 and E8202 that exhibited upregulated expression in the presence of vancomycin.

Gene*	Product	Fold Change E745	Fold Change E8202
EfmE745_00371	PadR family transcriptional regulator	2.6	3.1
EfmE745_00372	DUF1700 domain-containing protein	2.9	2.4
EfmE745_00373	Hypothetical protein	2.4	2.8

*Gene/locus tag in *E. faecium* E745

Another small cluster of genes that were found in a degraded prophage region were upregulated in both E745 and E8202 (Table 5.5). The first gene (*liaX*) in the cluster, encoded a daptomycin-sensing surface protein that was the most upregulated gene in both the E745 and E8202 operons at 3.2 and 7.9-fold respectively. The following two overlapping genes encoded a PspC domain containing protein and a phage holin family protein.

Table 5.5: Gene cluster 3 in E745 and E8202 that exhibited upregulated expression in the presence of vancomycin.

Gene*	Product	Fold change E745	Fold change E8202
EfmE745_01138 (liaX)	Daptomycin-sensing surface protein	3.2	7.9
EfmE745_01139	PspC domain-containing protein	2.4	4.1
EfmE745_01140	Phage holin family protein	2.1	4.7

*Gene/locus tag in *E. faecium* E745

5.4.6 Genes with downregulated expression that were shared between E745 and E8202

E745 and E8202 shared only 20 genes with downregulated expression. An operon consisting of four genes was highly downregulated in both E745 and E8202 (Table 5.6). The most downregulated gene of the operon in both strains was a gene encoding the accessory regulator AgrB, upon exposure to vancomycin the expression was downregulated 5.6-fold in strain E745 and 43.5-fold in strain E8202. The other genes in the operon encoded a LytTR family DNA-binding domain-containing protein and a sensor histidine kinase that were both

downregulated 4.1-fold in E745 and 8.6 and 10.7-fold in E8202. The final gene in the operon encoded a hypothetical protein that was downregulated 2.9-fold in E745 and 29-fold in E8202.

Table 5.6: Genes linked to quorum sensing that exhibited downregulated expression in the presence of vancomycin.

Gene*	Product	Fold Change E745	Fold Change E8202
EfmE745_00391	LytTR family DNA-binding domain-containing protein	-4.1	-8.6
EfmE745_00392	Sensor histidine kinase	-4.1	-10.7
EfmE745_00393	Hypothetical protein	-2.9	-29.2
EfmE745_00394	Accessory regulator AgrB	-5.6	-43.5

*Gene/locus tag in *E. faecium* E745

The alpha-galactosidase and beta-galactosidase large subunit genes were downregulated in both E745 and E8202 (Table 5.7). They were downregulated 2.4-fold and 2.5-fold in E745 but more so in E8202 at 99.0-fold and 90.5-fold respectively. The beta-galactosidase small subunit gene was also downregulated 66.7-fold in E8202 but was not downregulated in E745.

Table 5.7: Genes linked to lactose metabolism that exhibited downregulated expression in the presence of vancomycin.

Gene*	Product	Fold change E745	Fold change E8202
EfmE745_01560	Alpha-galactosidase	-2.4	-99.0
EfmE745_01561	Beta-galactosidase large subunit	-2.5	-90.5
EfmE745_01562	Beta-galactosidase small subunit	NS	-66.7

*Gene/locus tag in *E. faecium* E745
NS = Not significantly differentially expressed.

5.4.7 Upregulated genes found in a single strain

As well as the highly upregulated genes that were shared between strains, there were highly upregulated genes that were found in only one of the strains. Gene *EfmE745_01745* was the most upregulated chromosomal gene in strain E745, when exposed to vancomycin, with expression increasing by 7.2-fold. This gene encoded a protein that was most closely related to stress response protein *ysnF*. Another highly upregulated (7.0-fold) gene was *EfmE745_01664,* which encoded a glycosyl transferase family protein.

The genes in strain E8202 were typically upregulated to higher levels than E745. The expression of gene *EQB38_RS12340* that encoded a small rhodanese-like domain-containing protein was upregulated 35-fold. A small operon of two overlapping genes (*EQB38_RS12885* and *EQB38_RS12880*) was upregulated 32-fold, these genes encoded a glycine-betaine/L-proline ABC transporter ATP-binding protein. Another highly upregulated gene in E8202, *EQB38_RS06545,* encoded a LysM domain-containing protein that was upregulated 19-fold. Two other LysM domain-containing proteins were also upregulated in E8202 suggesting that they are important in the response to vancomycin.

5.4.8 Downregulated genes found in a single strain

As with the upregulated genes, there were highly downregulated genes that were found in only one of the strains. The most downregulated chromosomal gene in E745 was the *clpP* protease gene, the expression of this gene was downregulated 13-fold in the presence of vancomycin. Many of the other genes that were highly downregulated in E745 were prophage genes and these will be discussed in more detail in the next section. The most downregulated gene in E8202 was *EQB38_RS02405* which encoded a galactokinase, it was downregulated 164-fold in the presence of vancomycin. The gene directly upstream

EQB38_RS02410 encoded a galactose mutarotase that had similarly decreased expression levels. Genes belonging to the glycerol metabolism pathway were also highly downregulated in E8202, with genes *glpO* and *glpF* being downregulated 128-fold in the presence of vancomycin. Another gene in the glycerol metabolism pathway *glpK* was downregulated 58-fold. Many of the genes that were highly downregulated in strain E8202 belonged to phosphotransferase systems and this will be discussed in the following sections.

5.4.9 Prophage genes were downregulated in both E745 and E8202

Analysis of the chromosomes of *E. faecium* E745 and E8202 using PHASTER (Arndt *et al.*, 2016) revealed that they contained 3 and 4 intact prophage respectively. In E745, a prophage that spanned the region 1378331 bp – 1421682 bp had many genes that were differentially expressed. Genes involved in lysogeny, replication and regulation of the prophage were upregulated 2 to 4-fold while genes involved in head morphogenesis, tail morphogenesis and lysis were downregulated 2 to 10-fold. *E. faecium* E8202 had a prophage which spanned the region 2425435 bp – 2475952 bp. Unlike in E745, almost all of the genes in this prophage were downregulated including genes involved in lysogeny, replication, regulation, head morphogenesis, tail morphogenesis and lysis. The majority of the genes were downregulated between 4 and 5-fold.

5.4.10 Nutrient and ion transport were strongly downregulated in E8202

Nutrient and ion transport systems were drastically downregulated in *E. faecium* E8202 when exposed to vancomycin. Genes related to phosphotransferase systems (PTS) and ATP-binding cassette (ABC) transporters made up 18% of all downregulated genes in E8202. The expression of 71 PTS genes was downregulated on average 5-fold with the PTS genes for mannose-fructose-sorbose transport downregulated the most at 64-fold. The expression

of 33 ATP-binding cassette (ABC) transporter related proteins was also heavily downregulated. The ABC transporter related genes were downregulated on average 4-fold, however the maltodextrin ABC transporter permease gene *mdxG* was downregulated 45-fold.

5.4.11 The functional response to vancomycin varied considerably between strains

To gain a broader understanding into the effects of vancomycin exposure, the differentially expressed genes were grouped into functional Cluster of Orthologous Group (COG) categories. As was seen when comparing individual differentially expressed genes between E745 and E8202, the broader transcriptional response to vancomycin differed between the two strains (Figure 5.4). The COG group with genes involved in post-translational modification/protein turnover/chaperones was the most upregulated in E745 with 21% of the genes belonging to this category being upregulated. Genes belonging to the COG categories amino acid transport and metabolism (8%), signal transduction mechanisms (8%), cell cycle control/cell division/chromosome partioning (7%), inorganic ion transport & metabolism (6%) and coenzyme transport & metabolism (6%) were also upregulated when exposed to vancomycin. Intracellular trafficking/secretion and vesicular transport and carbohydrate transport and metabolism were the only COG categories in E745 where more genes in the category were downregulated than upregulated. E8202 differed considerably compared to E745 in the respect that the majority of the COG categories had a higher percentage of genes which were downregulated rather than upregulated. Carboydrate transport and metabolism was the COG category with the largest percentage of downregulated genes with 65% of genes in this category being downregulated. The COG categories energy production & conversion (30%), signal transduction mechanisms (23%),

amino acid transport & metabolism (22%) and transcription (21%) also had a large percentage of their genes downregulated. Genes involved in cell cycle control/cell division/chromosome partioning (15%) and transport and metabolism had a large percentage of upregulated genes particularly the categories; lipid transport & metabolism (24%), inorganic ion transport and metabolism (20%) and nucleotide transport and metabolism (17%). Comparing E745 with E8202, both strains had a large percentage of upregulated genes in the cell cycle control/cell division/chromosome partioning and inorganic ion transport and metabolism categories. Carbohydrate transport and metabolism, and intracellular trafficking/secretion and vesicular transport were the only two categories shared between the strains where the percentage of genes which were downregulated exceeded the percentage that were upregulated.

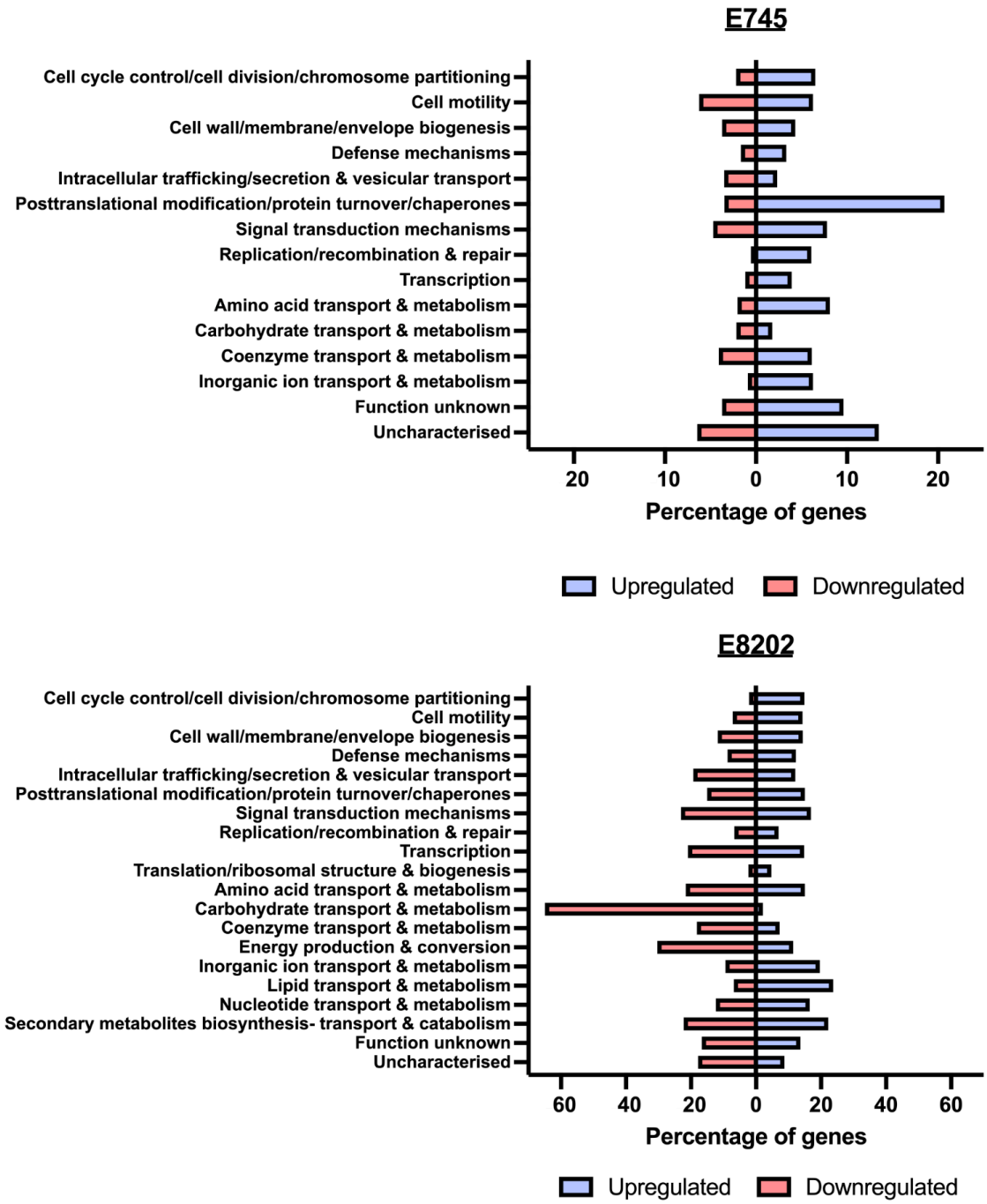

Figure 5.4: Functional analysis of differentially expressed genes. Genes are grouped into Cluster of Orthologous Group categories. Plots represent the percentage of genes in each category which are differentially expressed in the vancomycin treated samples compared to the control samples. Upregulated genes are shown in blue and downregulated genes are shown in red.

Tn-seq Analysis

RNA-seq was used to investigate the transcriptional response of the vancomycin resistant *E. faecium* strains E745 and E8202 to a sub-MIC concentration of vancomycin. However, this did not assess the fitness contribution of the genes in these strains. In this section, transposon sequencing (Tn-seq) was used to identify genes which contributed to fitness in the two *vanA*-carrying *E. faecium* strains E745 and E8202 and a *vanB*-carrying strain Aus0004 when exposed to ½ MIC vancomycin (E745: 128 μg/ml, E8202: 64 μg/ml and Aus0004: 8 μg/ml).

5.4.12 The vancomycin MIC was consistent in Iso-Sensitest broth and BHI broth for strains E745 and Aus0004 but not for strain E8202

The vancomycin minimum inhibitory concentration (MIC) was measured for *E. faecium* strains E745, Aus0004 and E8202 to determine a suitable concentration to treat the transposon libraries. The original transposon mutant libraries were constructed in BHI broth (by Dr Vincent de Maat, University Medical Centre Utrecht, The Netherlands (de Maat, 2022)) and to avoid selecting for mutants that had a fitness advantage in other media, BHI broth was used throughout the experiment. The MIC determinations were performed in Iso-Sensitest broth which is the standard media designed for antibiotic susceptibility testing and BHI broth which is widely used for the cultivation of *E. faecium*. The MIC of vancomycin was measured in both media to ensure that the correct concentration of vancomycin was used when performing the vancomycin selection in BHI broth (Table 5.8). The MIC values for strains E745 and Aus0004 were 256 and 16 μg/ml respectively in both types of media whereas strain E8202 had a lower MIC in BHI (128 μg/ml) compared to 256 μg/ml in Iso-Sensitest broth.

Table 5.8: MIC of vancomycin for *E. faecium* strains grown in BHI broth and Iso-Sensitest broth.

E. faecium strain	MIC in Iso-Sensitest broth (μg/ml)	MIC in BHI broth (μg/ml)
E745	256	256
Aus0004	16	16
E8202	256	128

BHI = Brain Heart Infusion

5.4.13 The transposon sequencing libraries were successfully prepared

The transposon sequencing libraries were run on a 2.5% agarose gel to confirm that they had been constructed successfully. All 24 of the libraries produced a band at the correct size of 130 bp (Figure 5.5AC). The libraries were pooled in equimolar concentrations to produce two pooled libraries consisting of the first 12 libraries and the second 12 libraries. The pooled libraries were also checked on a 1% agarose gel, both pooled libraries gave a single band at the correct size of 130 bp (Figure 5.5BD). These pooled libraries were then taken forward for single-end 50 bp nucleotide sequencing.

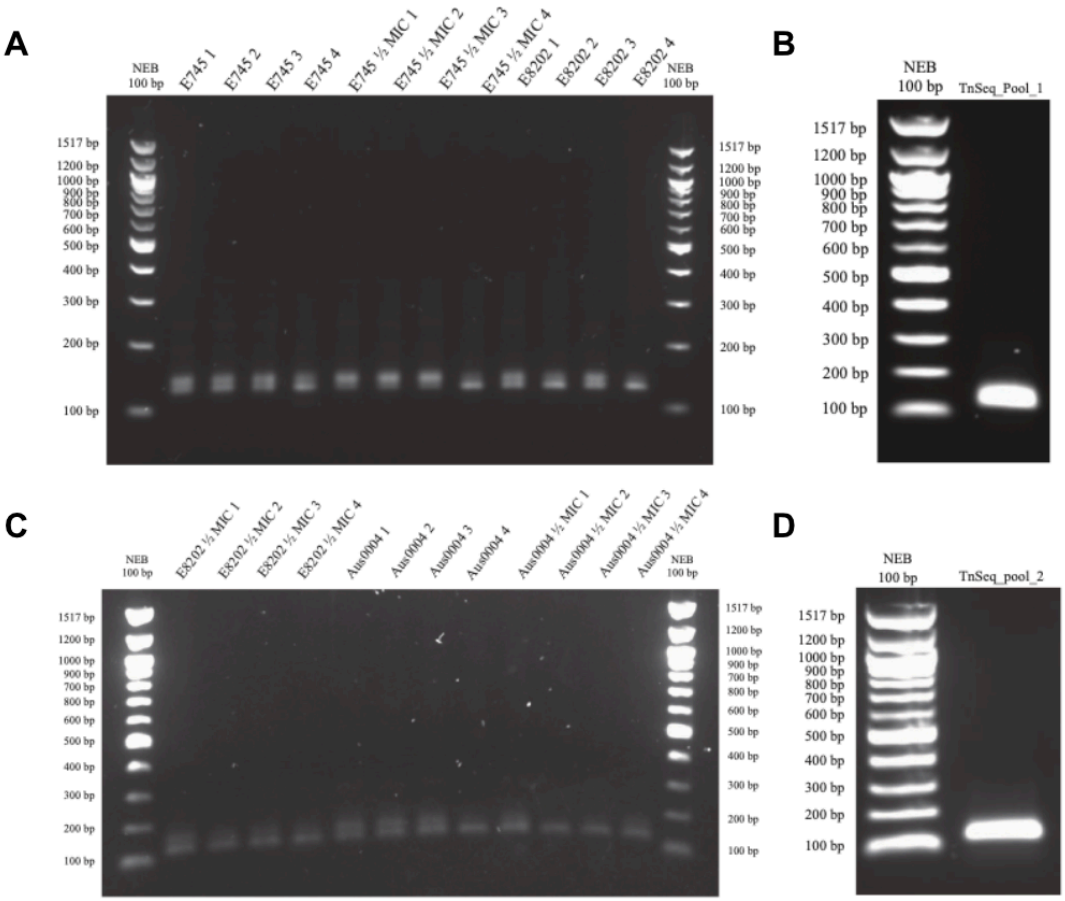

Figure 5.5: Tn-seq library preparation agarose gels. **A.** Individual Tn-seq libraries for the E745 and E8202 control samples and the E745 vancomycin treated samples. **B.** Pooled library of samples from A. **C.** Individual Tn-seq libraries for the Aus0004 control samples and the E8202 and Aus0004 vancomycin treated samples. **D.** Pooled library of samples from C.

5.4.14 Raw read demultiplexing and mapping

Raw sequencing reads were demultiplexed to give an average of 30,402,068 reads per sample (Table 5.9). The sequencing reads were then mapped to the genomic assemblies of the respective parent strains (E745, E8202 and Aus0004). The majority of reads could be mapped to their respective genomes with an average of 92% of reads mapped per sample.

Table 5.9: Mapping of Tn-seq reads to the *E. faecium* parent strain assemblies.

Sample	Total reads	Reads mapped	Reads mapped (%)
E745 1	33383129	29572911	88.59
E745 2	31939542	27989752	87.63
E745 3	32747318	29403833	89.79
E745 4	27558016	23174257	84.09
E745 ½ MIC 1	33768324	29664610	87.85
E745 ½ MIC 2	33546883	30162696	89.91
E745 ½ MIC 3	30290827	26997566	89.13
E745 ½ MIC 4	24934193	21598317	86.62
E8202 1	31336449	28595286	91.25
E8202 2	29236425	25608534	87.59
E8202 3	30832500	27968025	90.71
E8202 4	24559935	22133018	90.12
E8202 ½ MIC 1	32911318	30077647	91.39
E8202 ½ MIC 2	29026007	26737539	92.12
E8202 ½ MIC 3	32573394	30119175	92.47
E8202 ½ MIC 4	28869175	26085774	90.36
Aus0004 1	30025772	28781430	95.86
Aus0004 2	30893350	29770618	96.37
Aus0004 3	32018405	31022543	96.89
Aus0004 4	25632194	24822607	96.84
Aus0004 ½ MIC 1	31577287	30166051	95.53
Aus0004 ½ MIC 2	35064182	33877744	96.62
Aus0004 ½ MIC 3	28808525	27330263	94.87
Aus0004 ½ MIC 4	28116480	27138621	96.52

MIC = Minimum inhibitory concentration

5.4.15 Transposon sequencing is reproducible

Principal component analysis was used to investigate whether the gene insertion profiles of the replicate samples were similar and whether the control samples and vancomycin treated samples were different. The insertion profiles of the vancomycin treated samples clustered away from the control samples for E745, E8202 and Aus0004 (Figure 5.6ABC). The control samples tended to cluster more closely together than the vancomycin treated samples, especially for strains E8202 and Aus0004, which suggested that there was some variation in the insertion profiles of the replicate samples. However, the majority of the variance was explained by the X-axis meaning the variation between the samples was small.

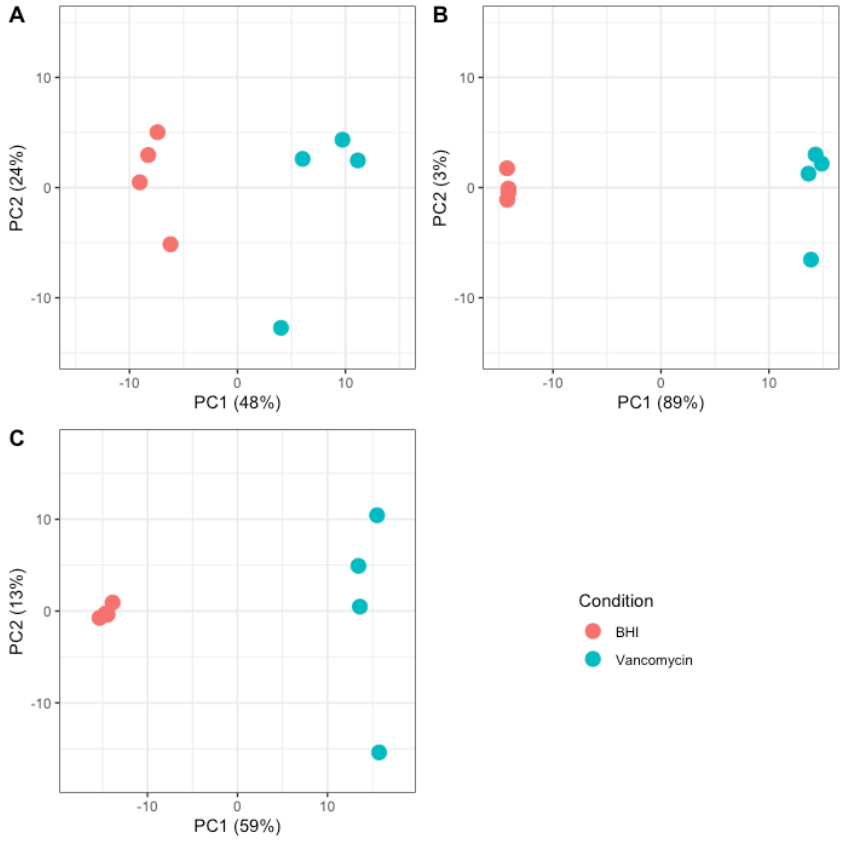

Figure 5.6: PCA plots showing the variation between Tn-seq samples. **A.** E745 samples. **B.** E8202 samples. **C.** Aus0004 samples. Insertion counts underwent variance stabilising transformation prior to plotting. Red circles represent the control samples and blue circles represent samples treated with ½ MIC vancomycin.

5.4.16 Genes in the *vanA* operon had the greatest effect size in E745, but chromosomal genes were also identified as contributing to fitness in the presence of vancomycin.

There were 30 genes in *E. faecium* E745 identified as having a $\log_2$ fold-change < -1 and a Benjamini-Hochberg adjusted *P* value < 0.05, these genes were designated as potentially contributing to increased fitness in the presence of ½ MIC vancomycin (Table 8.2, Figure 5.7). The top four genes with the largest effect size were *vanX*, *vanA*, *vanR* and *vanS* which are part of the well characterised *vanA*-type vancomycin resistance operons. *vanH*, the first gene in the *vanHAX* operon was also among the genes with the largest fold-change. In addition to the *vanA* operon, chromosomally encoded genes were also identified as significantly contributing to fitness. The genes *mgs, npr_1, avrA, rarA, gmuF, tig* and *ispE* all had a greater than 3-fold change between the control and vancomycin treated libraries.

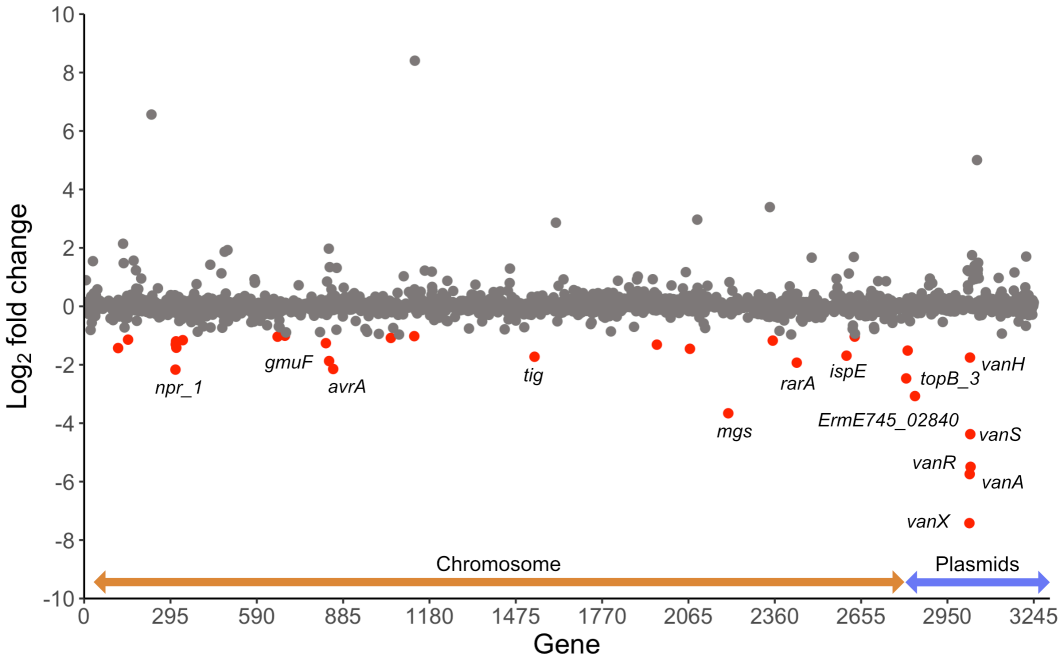

Figure 5.7: Tn-seq analysis to identify genes that contribute to survival of *E. faecium* E745 in the presence of vancomycin. The X-axis represents each gene in the genome and the Y-axis represents the $\log_2$ fold-change in reads between the control and vancomycin treated condition. Circles with a negative value represents genes contributing to increased fitness in the presence of vancomycin. Circles with a positive value represent genes that contribute to increased fitness in the control condition but not in the vancomycin treated condition. Genes that are red had a $\log_2$ fold-change < -1 and a Benjamini-Hochberg adjusted *P* value < 0.05. Red circles which have been labelled have a greater than 3-fold change from the control condition. The orange arrow highlights genes found in the chromosome while the blue arrow highlights genes found on plasmids.

5.4.17 Genes in the *vanA* operon also had the greatest effect size in E8202, and again chromosomal genes were identified as contributing to fitness in the presence of vancomycin.

61 genes were identified in E8202 that had a $\log_2$ fold-change < -1 and a Benjamini-Hochberg adjusted *P* value < 0.05 (Table 8.3, Figure 5.8). As with E745 the gene with the largest effect size was *vanX* and the other genes of the *vanHAX* operon were among the genes with the largest effect size. E8202 also had a number of chromosomal genes with a large effect size, these included *asnB*, *gshAB* and *glpO*.

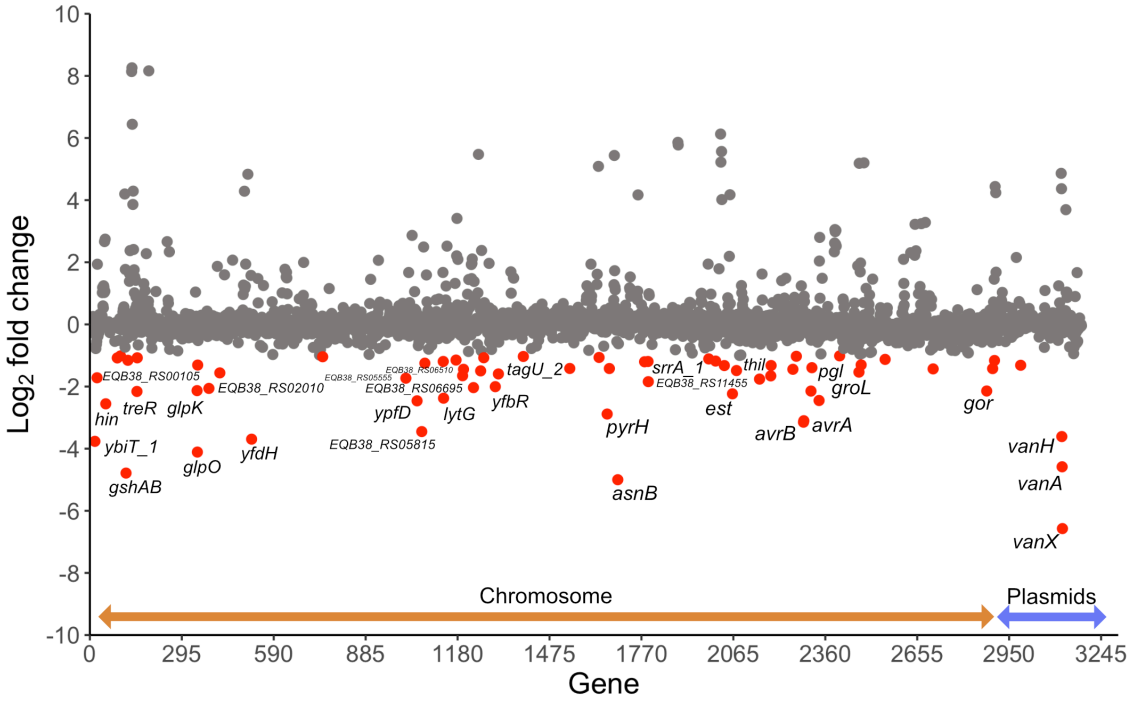

Figure 5.8: Tn-seq analysis to identify genes that contribute to survival of *E. faecium* E8202 in the presence of vancomycin. The X-axis represents each gene in the genome and the Y-axis represents the $\log_2$ fold-change in reads between the control and vancomycin treated condition. Circles with a negative value represents genes contributing to increased fitness in the presence of vancomycin. Circles with a positive value represent genes that contribute to increased fitness in the control condition but not in the vancomycin treated condition. Genes that are red had a $\log_2$ fold-change < -1 and a Benjamini-Hochberg *P* value < 0.05. Red circles which have been labelled have a greater than 3-fold change from the control condition. The orange arrow highlights genes found in the chromosome while the blue arrow highlights genes found on plasmids.

5.4.18 Many chromosomal genes were identified as contributing to fitness in the presence of vancomycin in Aus0004, but the *vanB*-type resistance genes were not identified

29 genes were identified in Aus0004 that had a $\log_2$ fold-change < -1 and a Benjamini-Hochberg adjusted *P* value < 0.05 (Table 8.4, Figure 5.9). Unlike E745 and E8202, none of the known vancomycin resistance genes were identified as significant. However, a number of chromosomal genes had large effect sizes such as *mdxG*, *argS* and *groL*. Compared to E745 and E8202, Aus0004 had more genes with a large fold-change compared to the controls. Gene *EFAU004_00372* had a fold-change of 650-fold in Aus0004 which was much greater than the largest fold-change of 170-fold and 95-fold in E745 and E8202, respectively. Aus0004 had two regions that no reads mapped to in either the control or vancomycin-treated conditions. Both of the regions corresponded to intact prophages, raising the possibility that the regions were lost during the construction of the library. The *vanB-type* vancomycin resistance genes were not found in either of these regions.

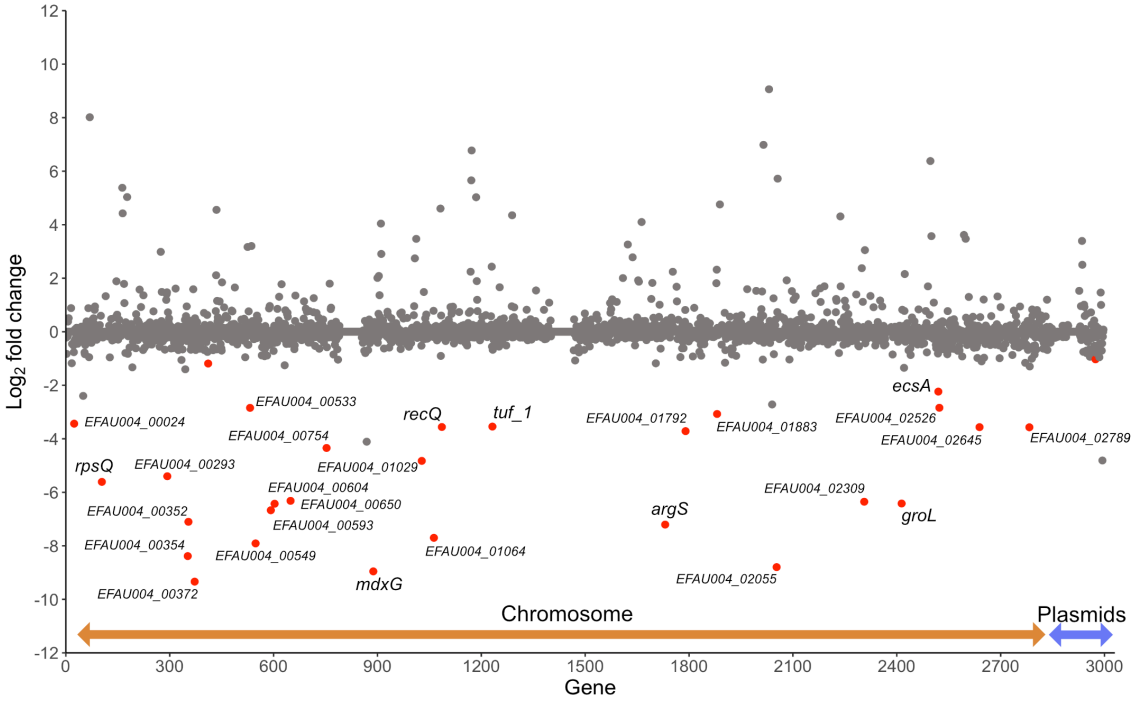

Figure 5.9: Tn-seq analysis to identify genes that contribute to survival of *E. faecium* Aus0004 in the presence of vancomycin. The X-axis represents each gene in the genome and the Y-axis represents the $\log_2$ fold-change in reads between the control and vancomycin treated condition. Circles with a negative value represents genes contributing to increased fitness in the presence of vancomycin. Circles with a positive value represent genes that contribute to increased fitness in the control condition but not in the vancomycin treated condition. Genes that are red had a $\log_2$ fold-change < -1 and a Benjamini-Hochberg adjusted *P* value < 0.05. Red circles which have been labelled have a greater than 3-fold change from the control condition. The orange arrow highlights genes found in the chromosome while the blue arrow highlights genes found on plasmids.

5.4.19 Genes identified by transposon sequencing had varied functions

Genes identified as significant from the Tn-seq analysis were annotated with eggNOG-mapper to identify groups of genes with similar functions (Figure 5.10). E745, E8202 and Aus0004 had genes which belonged to the three major categories: cellular processes and signalling, information storage and processing, and metabolism. The number of genes in each category varied considerably between strains, but certain gene functions were enriched. The replication, recombination and repair category had the most genes in strains E745 and Aus0004 and the third most in strain E8202. Genes relating to transcription were also prevalent across the strains. In the metabolism category, genes relating to carbohydrate transport and metabolism were frequently identified in all strains. In the cellular processes and signalling category cell wall/membrane/envelope biogenesis genes were common among all three strains, particularly in E8202 where these genes made up 17% of all the genes that were identified as contributing to fitness in the presence of vancomycin.

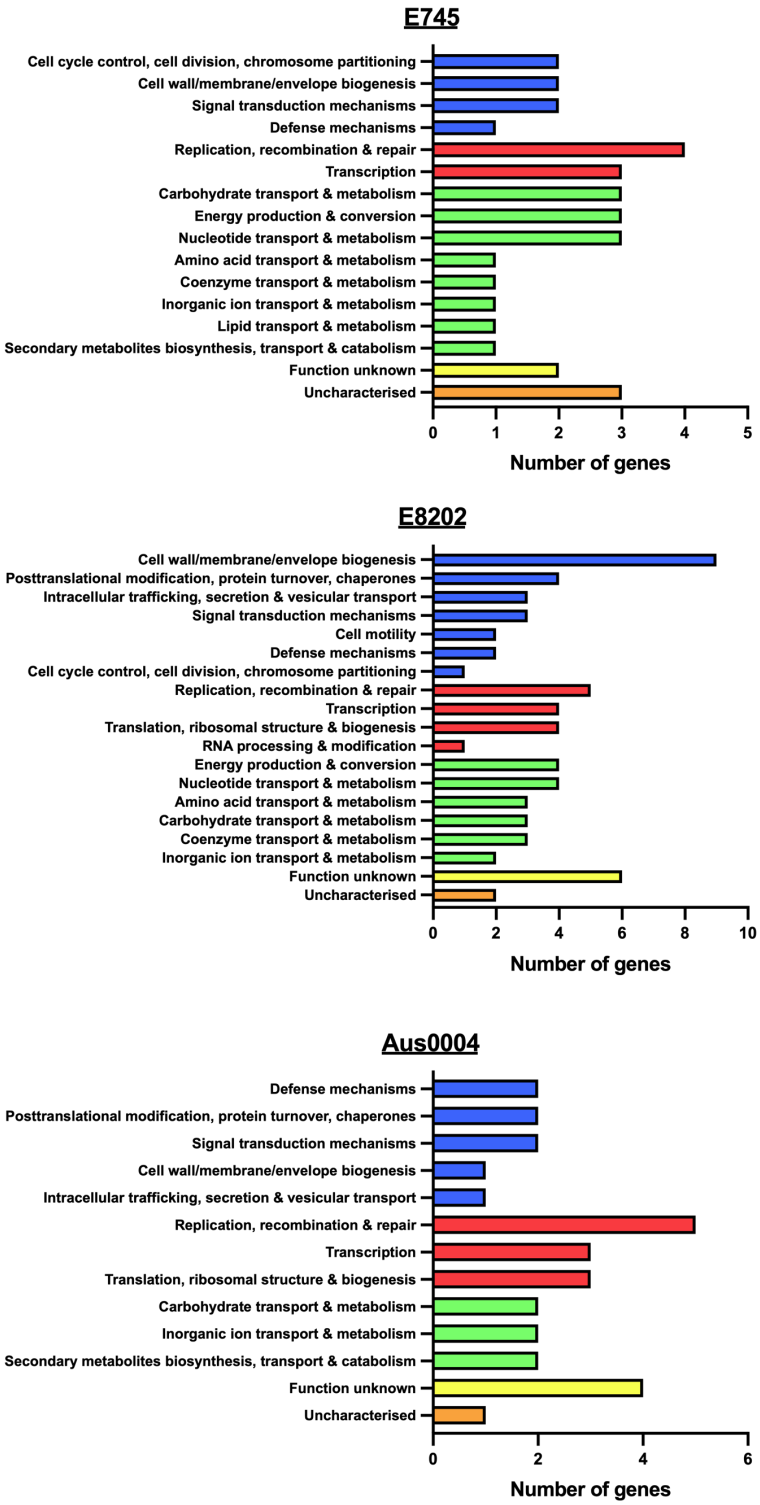

Figure 5.10: Functional annotation of significant Tn-seq gene hits. Genes were grouped into different functional categories and summed for each category. Bars were coloured to designate higher level function; blue bars represent Cellular processes & signalling, red bars represent Information storage & processing, green bars represent Metabolism, yellow bars represent poorly characterised genes and orange bars represent uncharacterised genes.

5.4.20 Identification of operons contributing to vancomycin resistance

Comparing strains E745 and E8202, seven potential operons were identified that contained more than one significant gene in the Tn-seq analysis. The *vanHAX* operon, which is essential for vancomycin resistance, was identified in both E745 and E8202 (Figure 5.11A). Additionally, the *vanRS* genes that encode the two-component regulatory system that senses the presence of vancomycin were also identified in E745. Another operon that was identified in both E745 and E8202 was involved in the metabolism of glycerol (Figure 5.11B). *EfmE745_00313, glpK* and *glpO* were identified as having a large effect size in both strains. However, the first gene in the operon *npr_1*, was only identified in strain E745 and the last gene in the operon *glpF* was only identified in strain E8202. An operon was also identified that contained genes involved in peptidoglycan remodelling (Figure 5.11C). A D-alanyl-D-alanine carboxypeptidase gene *vanYB* and a peptidoglycan hydrolase gene *lytG* were identified in E8202 and a gene encoding a hypothetical protein at the start of the operon was identified in strain E745. A two-component system was also identified in E8202 that belongs to the PhoP family (Figure 5.11D). A small operon of two genes was identified in E8202 that was involved in the sulphur modification of tRNAs (Figure 5.11E). The final operon identified in E745 and E8202 consisted of two overlapping genes (*avrAB*). The function of these genes is unknown (Figure 5.11F).

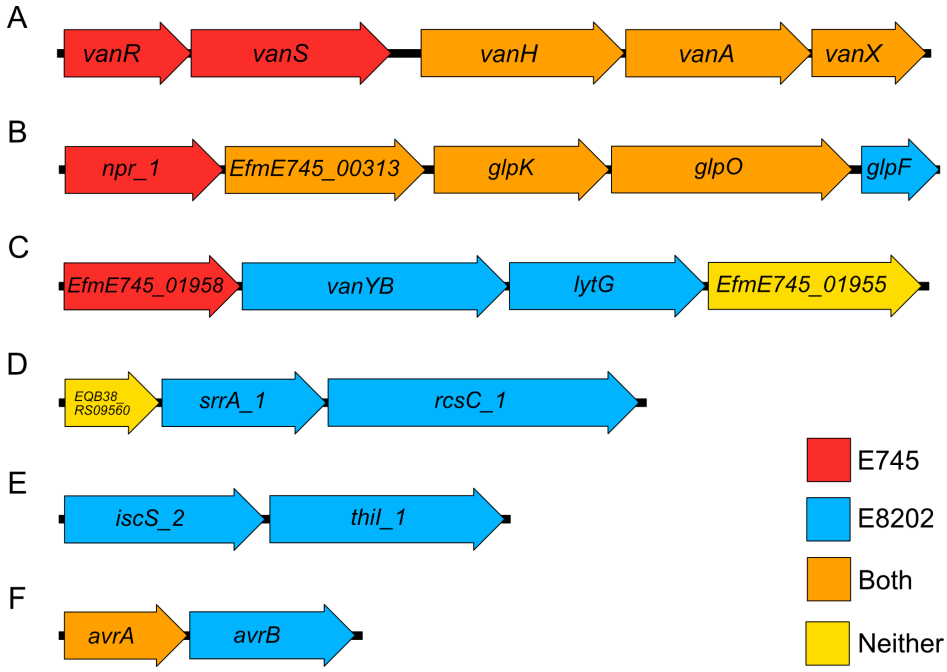

Figure 5.11: Operons containing genes identified as significant in the Tn-seq analysis. Red genes were found in E745, blue genes were found in E8202, orange genes were found in both E745 and E8202 and yellow genes were found in neither strain.

5.4.21 There was overlap between the genes identified for the *vanA* carrying strains but not the *vanB* carrying strain

To identify genes that provided a fitness advantage in presence of vancomycin across different strains of *E. faecium* the significant genes were compared between strains E745, E8202 and Aus0004. 90 significant genes were found in total between strains E745, E8202 and Aus0004. 8 of the significant genes were identified in both E745 and E8202, none of the genes identified in E745 or E8202 were identified in Aus0004 (Figure 5.12). The shared genes were composed of the vancomycin resistance genes *vanH*, *vanA*, *vanX* and genes *avrA*, *gshAB*, *glpO*, *glpK* and *treR*.

Figure 5.12: The number of genes that had a greater than 2-fold change in insertion frequency between the control group and the vancomycin treated group. The circles are proportional to the number of genes and show the number of unique genes, the intersection of the circles is the number of genes shared between E745 and E8202.

5.4.22 Comparison of RNA-seq and Tn-seq genes

Genes that were found to be significant in the Tn-seq analysis were compared with those identified by RNA-seq to determine whether these genes were differentially expressed when the strain was exposed to vancomycin. *E. faecium* E745 had seven genes that were significant in the Tn-seq analysis and were upregulated in the presence of vancomycin (Figure 5.13A). Five of the seven upregulated genes belonged to the vancomycin resistance operons. Two further genes that were identified by Tn-seq, *yvgN* and *dinB_2*, were also upregulated when exposed to vancomycin. *E. faecium* E8202 had 18 genes that were identified by Tn-seq and were also upregulated when exposed to vancomycin (Figure 5.13B). Like E745, the *vanHAX* genes of the vancomycin resistance operon were the most upregulated. Genes *avrA*, *avrB, yfbR, EQB38_RS00105, tagU_2, , EQB38_RS00630* and , *EQB38_RS13625* were also upregulated at least 3-fold (Figure 5.13B). Unlike E745, nine of the genes that were found to be significant in the Tn-seq analysis were significantly downregulated when exposed to vancomycin. Genes *glpO*, *glpK* and *glpF* of the glycerol

metabolism pathway were strongly downregulated as well as the trehalose hydrolase gene *treA*.

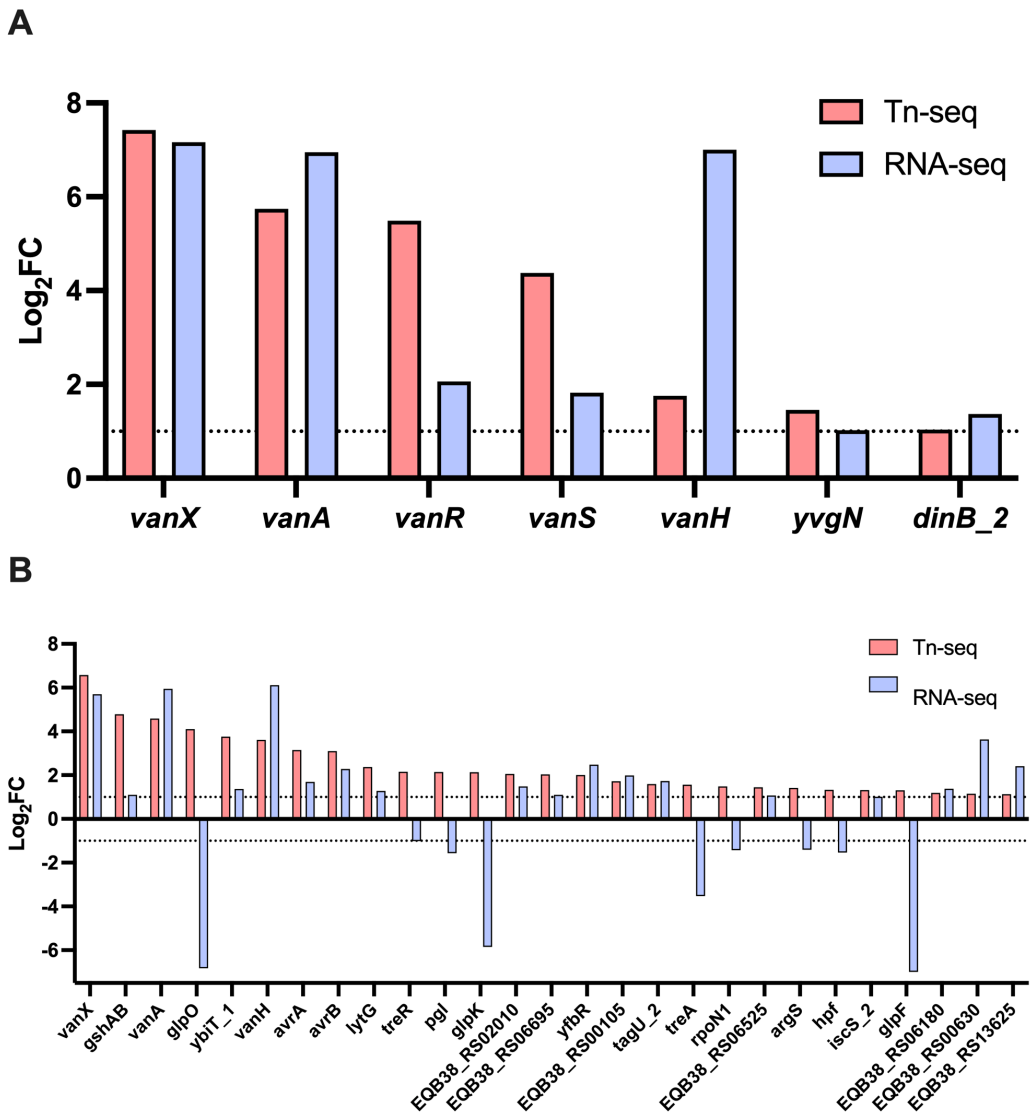

Figure 5.13: Genes identified by both RNA-seq and Tn-seq. **A.** Significant genes identified in *E. faecium* E745. **B.** Significant genes identified in *E. faecium* E8202. Values are the $\log_2$ fold-change between the vancomycin treated samples and the control samples. Only genes with an FDR < 0.05 and $\log_2$ fold-change < -1 or > 1 were included. Dotted lines at $\log_2$ fold change -1 and 1 represent the $\log_2$ fold change cut-offs.

5.4.23 The majority of genes were found in only the most closely related species to *E. faecium*

The genomes of 24 species of *Enterococcus*, 3 strains of closely related genera (*Melissococcus plutonius*, *Tetragenococus halophilus* and *Vagococcus lutrae*) and 2 outgroup species (*Carnobacterium maltaromaticum* and *Lactococcus garvieae*) were queried for the genes identified in the previous analysis (Figure 5.14). The *mgs* gene was found in all *Enterococcus* genomes but was absent in the two outgroup species and *Vagococcus lutrae*. Genes *asnB*, *avrB*, *glpO* and *gshAB* were all present in *E. faecium* and in the closely related species *Enterococcus durans*, *Enterococcus villorum*, *Enterococcus mundtii*, *Enterococcus hirae* and the more remotely related species *Enterococcus pallens* and *Enterococcus gilvus*. The *vanX* gene was only found in *E. faecium* and *E. faecalis*. The *avrA* gene was only found in the *E. faecium* genomes.

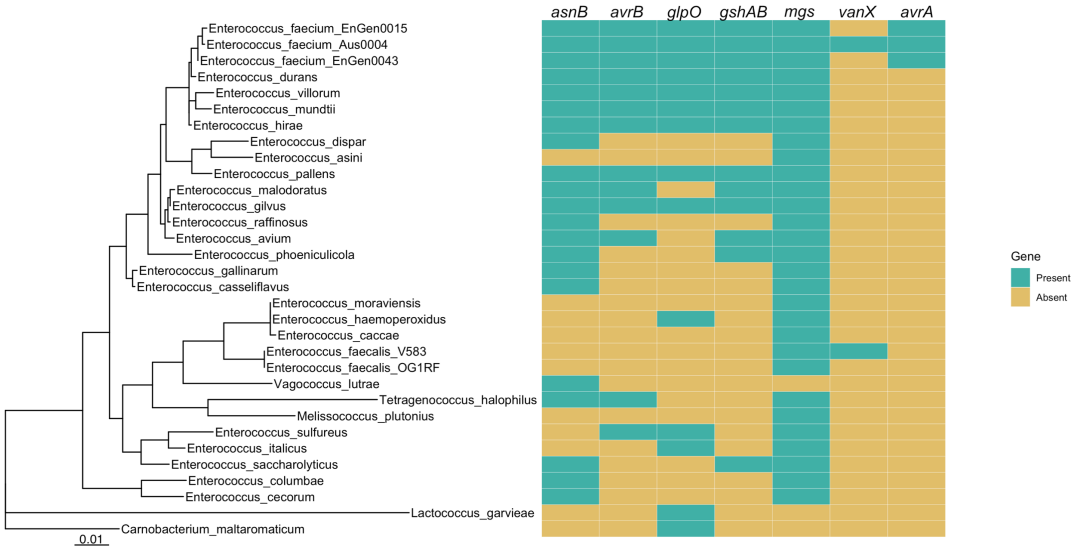

Figure 5.14: Phylogenetic distribution of the identified genes in the genus *Enterococcus*. Genes present in the genome are represented by turquoise squares and genes which are absent from the genome are represented by gold squares. Tree was constructed using MEGA form the alignment of the 16S genes. The scale bar indicates the average number of nucleotide substitutions per site.

5.4.24 The identified genes were conserved throughout the species *E. faecium*

To understand how well the identified genes were conserved within the species *Enterococcus faecium*, the genes were queried against the genomes of 1644 clade A *E. faecium* isolates (Figure 5.15). Gene *asnB* was found in all 1644 isolates, *avrB* and *mgs* were found in all but one isolate and *avrA* and *gshAB* were found in all but two isolates. Gene *glpO* was found in 1631 of the isolates and *vanX* was found in around half of the isolates. The distribution of *vanX* can be explained by its genetic context, with *vanX* being typically found on extra-chromosomal elements whereas the other genes are chromosomally encoded.

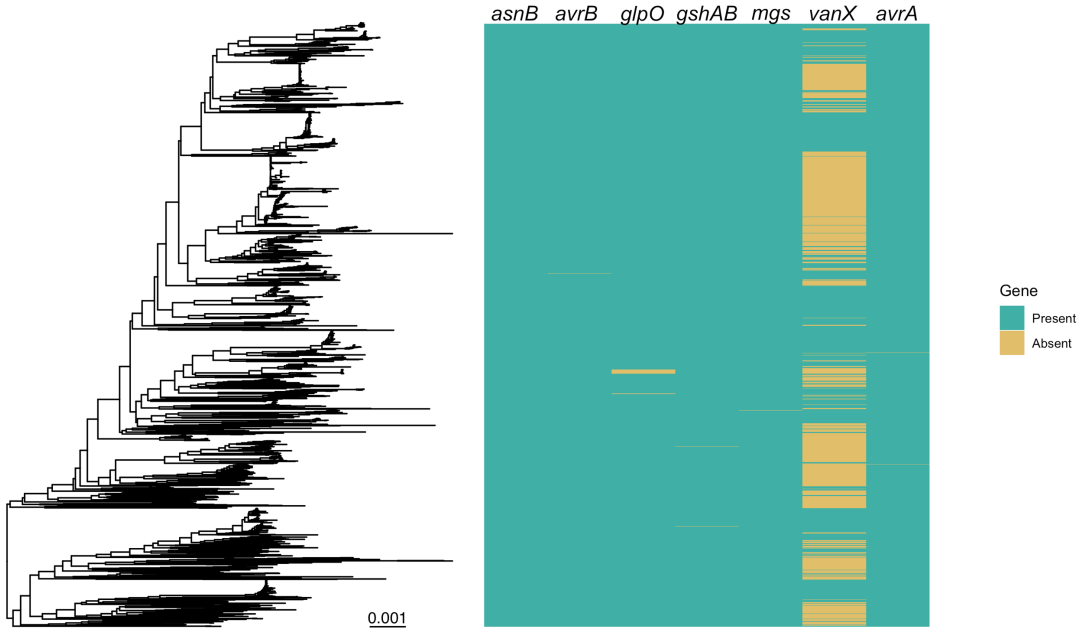

Figure 5.15: Phylogenetic distribution of the identified genes in the species *Enterococcus faecium*. Genes present in the genome are represented by turquoise squares and genes which are absent from the genome are represented by gold squares. Tree constructed from mash distances. The scale bar represents the mash distance.

5.4.25 AvrA and AvrB may be cell surface proteins

Genes *asnB*, *glpO*, *gshAB* and *mgs* all have predicted functions but the genes *avrA* and *avrB* encoded hypothetical proteins. The amino acid sequences of both genes were queried using the HMMER web server (Potter *et al.*, 2018), no functional domains were predicted for *avrA* however it did predict an N-terminal signal sequence. SignalP 6.0 (Teufel *et al.*, 2022) detected that it was a Sec/SPI signal sequence which suggested that the protein is secreted using the Sec pathway. The HMMER webserver identified that *avrB* had a DUF6681 domain and the Phobius webserver (Käll, Krogh and Sonnhammer, 2007) was able to identify two transmembrane domains (amino acids 24 – 44 and 56 – 79) which suggested that it may be a cell surface-associated protein.

5.4.26 Generation of deletions in genes identified by Tn-seq

To confirm the results of the Tn-seq experiment six genes (*avrA*, *avrB*, *asnB*, *glpO*, *gshAB* and *mgs*) that were predicted to increase fitness in the presence of vancomycin in *E. faecium* E745 and/or E8202 were selected for the generation of deletion mutants in *E. faecium* E745. Five of the six genes were successfully deleted using a CRISPR-Cas9 based system (Figure 5.16). Despite several attempts, it was not possible to delete the *mgs* gene. While it was possible to delete five of the six original genes, plasmid incompatibility issues during generation of the mutants meant that the plasmid carrying the vancomycin resistance genes was lost in strains with deletions in *avrA, asnB* and *gshAB*. In order to investigate the effect of deleting genes *avrA, asnB* and *gshAB*, a strategy was devised to conjugate the *vanA* carrying plasmid back into these mutants.

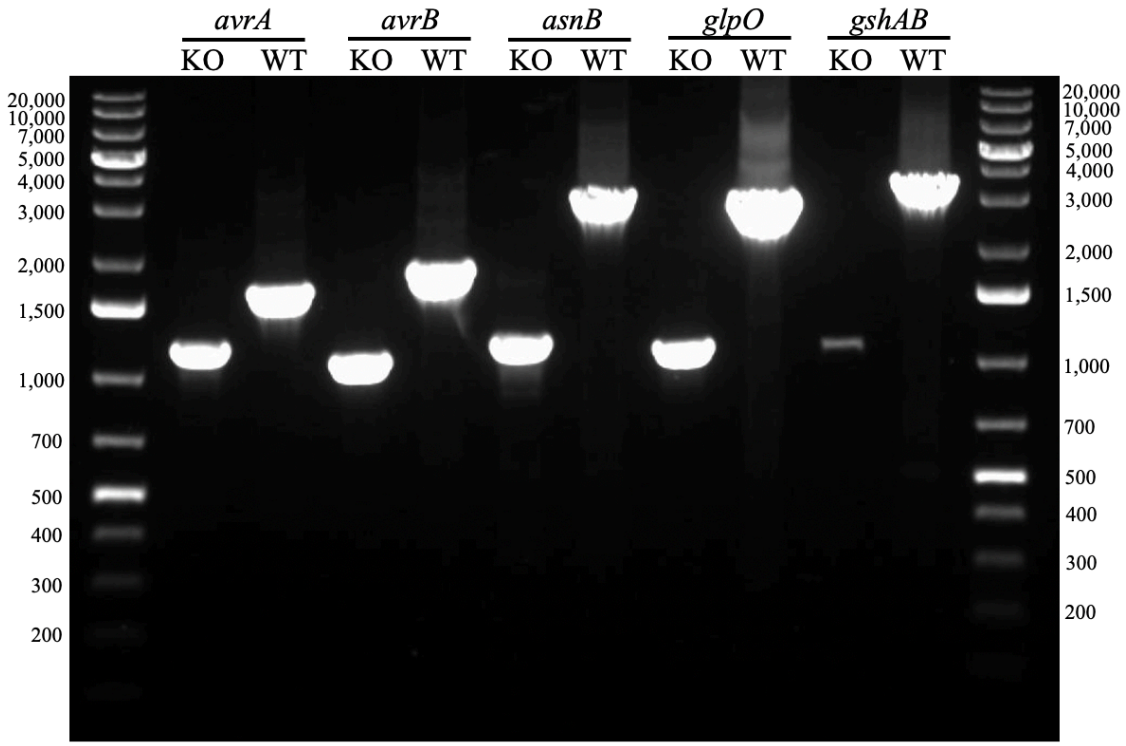

Figure 5.16: PCR analysis of successful gene deletions in *E. faecium* E745. KO = Deletion mutant. WT = Wildtype E745.

To conjugate the vancomycin resistance plasmid back into the deletion strains, a strategy had to be devised that did not use vancomycin for selection as this would have selected for mutations that could have obscured the downstream analysis. To overcome this issue, conjugation was performed between the E745 transposon library and the intermediate strain *E. faecium* 64/3 using the gentamicin resistance gene of the transposon as a selection marker (Figure 5.18). Ten transconjugants were picked and the position of the transposon insertion site in the vancomycin resistance plasmid was mapped to ensure it did not disrupt any essential genes. Inverse PCR with outward facing primers specific for the gentamicin transposon were used to map the insertion site. Sanger sequencing of the inverse PCR products determined that the transposon had inserted into plasmid 5 of E745 in all 10 transconjugants (Figure 5.17). Nine of the transconjugants had insertions in the Beta-L-arabinobiosidase precursor gene *hypBA2* and a single transconjugant had an insertion in gene

EfmE745_03105 that encoded a hypothetical protein. Plasmid 5 was not the plasmid carrying the vancomycin resistance genes suggesting that both the vancomycin resistance plasmid and plasmid 5 were co-transferred in the conjugation.

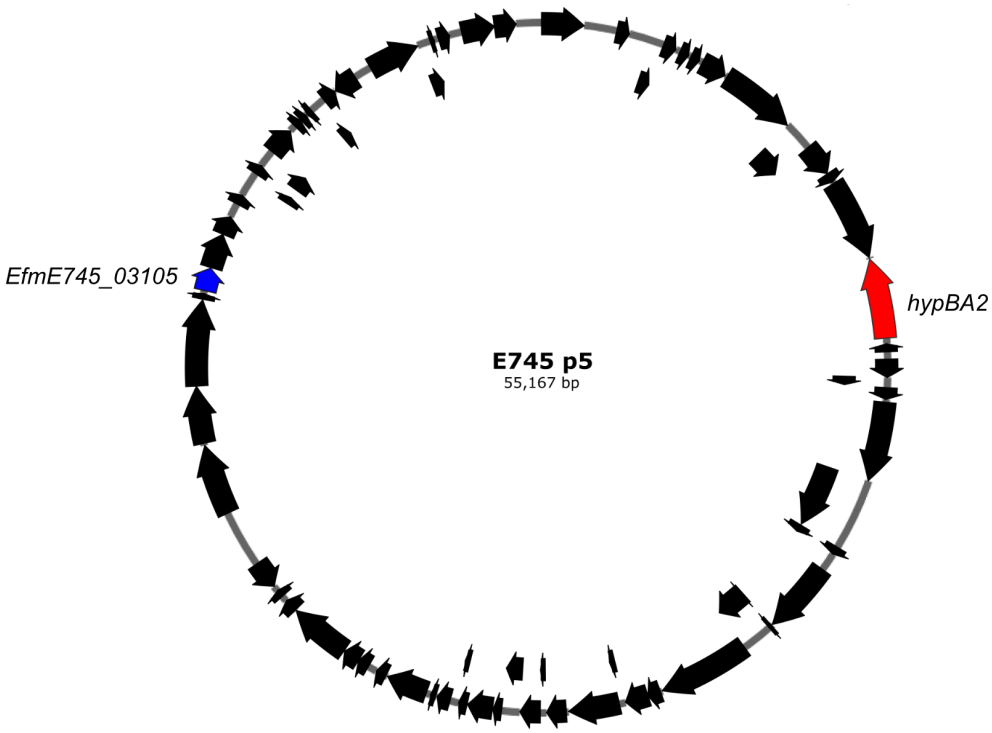

Figure 5.17: Transposon insertion sites. Plasmid map of E745 p5 showing insertion sites of the mariner transposon. Highlighted genes in blue and red contain insertions of the mariner transposon.

A further conjugation was then performed from the 64/3 transconjugant to each of the deletion strains and transconjugants were identified that carried the vancomycin resistance plasmid (Figure 5.18). As the insertions occurred in a separate plasmid to the vancomycin resistance genes, it was believed that the insertion should not affect the vancomycin resistance genes or the replication of the vancomycin resistance plasmid.

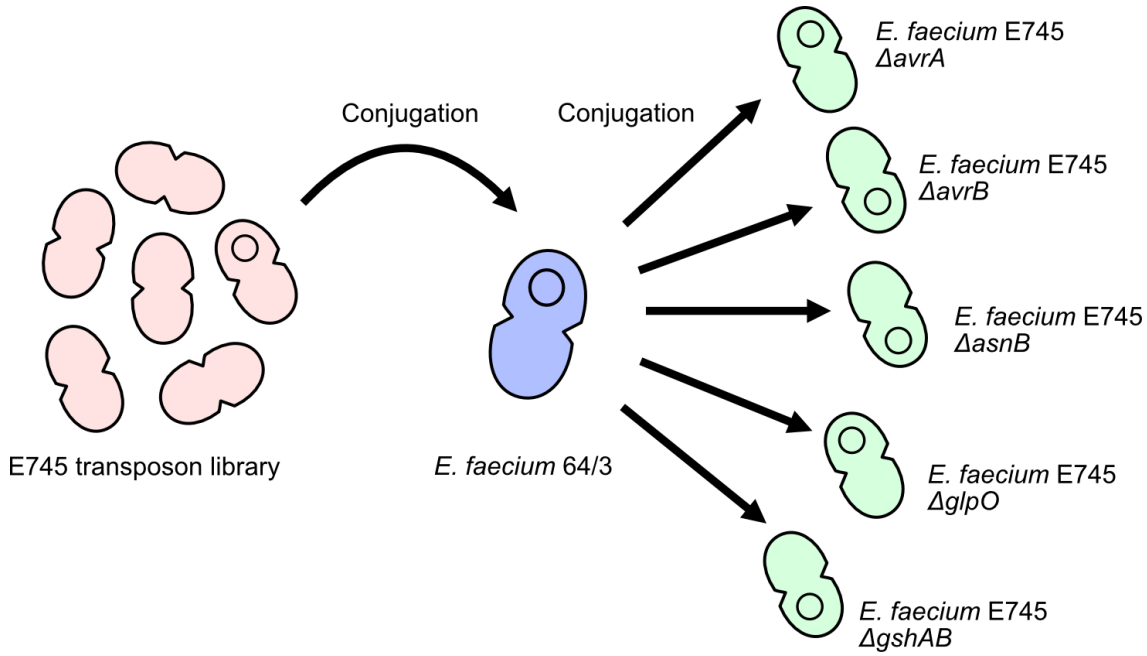

Figure 5.18: A schematic of the strategy used to restore the vancomycin resistance plasmid to the E745 deletion mutants. The vancomycin resistance plasmid encoding a gentamicin resistance transposon was conjugated from the E745 transposon library to *E. faecium* 64/3. The plasmid was then further conjugated from *E. faecium* 64/3 to each of the deletion mutants.

5.4.27 Deletion of genes identified by Tn-seq had no effect on vancomycin MIC

Minimum inhibitory concentration assays were carried out on the deletion mutants to determine whether the deleted genes contributed to vancomycin resistance. The wildtype isolate E745 and strains with deletions in *avrB* and *glpO* had a vancomycin MIC of 256 μg/ml. Strains with deletions in genes *avrA*, *asnB* and *gshAB* all had an MIC of 1 μg/ml. However, an E745 control strain that had no deletions and in which the vancomycin plasmid had been conjugated back in, also had an MIC of 1 μg/ml suggesting that conjugation of the vancomycin resistance plasmid back into the E745 strains did not restore vancomycin resistance. An attempt was also made to overexpress genes *avrA*, *avrB*, *asnB*, *glpO* and *gshAB* however due to further plasmid incompatibility issues it was not possible to overexpress the genes in the current project.

5.4.28 Growth deficiencies were identified in the deletion mutants

Although the deletion of genes *avrB* and *glpO* did not have an effect on the vancomycin MIC in *E. faecium* E745 it was hypothesised that they may have an effect on growth in the presence of vancomycin. *E. faecium* E745, Δ*avrB* and Δ*glpO* were grown in the presence and absence of ½ MIC (128 µg/ml) vancomycin for 15 hours and the absorbance (Abs) at 600 nm was measured (Figure 5.19). The lag phase of wildtype *E. faecium* E745 was extended by 5 hours when exposed to ½ MIC vancomycin but the maximal growth increased from Abs 0.5 to 0.9. The Δ*avrB* and Δ*glpO* strains had the same lag phase and initial growth rate as wildtype E745 when grown in the absence of vancomycin but had higher total growth of Abs 0.85 and Abs 0.7 respectively. When the wildtype E745 strain and mutants were exposed to ½ MIC vancomycin, E745 and the Δ*glpO* mutant had an extremely similar lag phase, growth rate and maximal growth however the Δ*avrB* mutant had a lag phase that was 4 hours longer, a decreased growth rate and the maximal growth was around half of the wildtype.

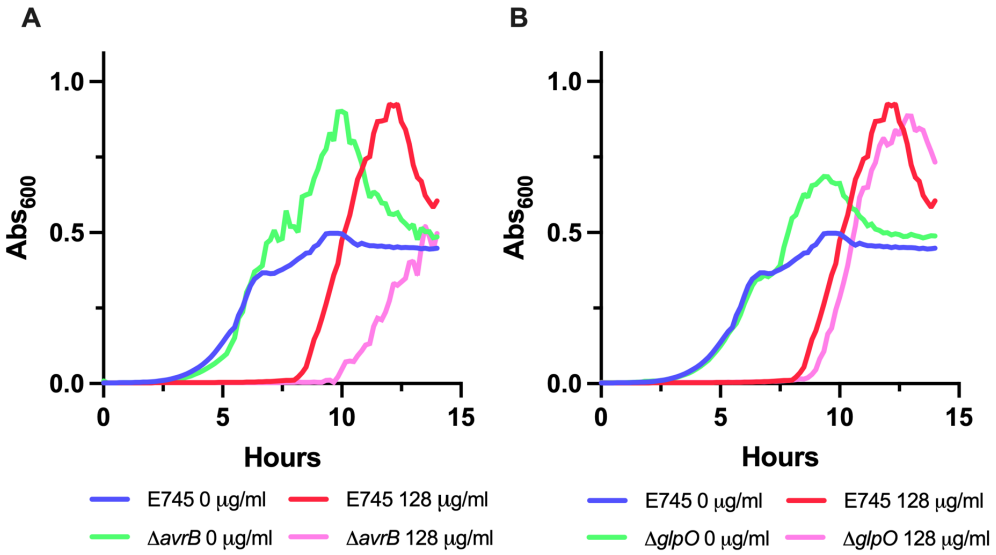

Figure 5.19: Growth kinetics of the *avrB* and *glpO* deletion mutants and wildtype E745 in the presence and absence of ½ MIC (128 µg/ml) vancomycin. Absorbance was measured at 600 nm. $n = 3$.

6.5 Discussion

RNA-seq and Tn-seq were used to investigate the response of vancomycin-resistant *E. faecium* strains exposed to ½ MIC vancomycin. RNA-seq identified wide transcriptional changes in the isolates when exposed to vancomycin and Tn-seq identified a number of plasmid and chromosomal genes that increased fitness in the presence of vancomycin.

Despite E745 and E8202 both being *vanA*-carrying vancomycin-resistant *E. faecium* strains, their transcriptional response to vancomycin was distinct. However, many of the most upregulated genes in E745 and E8202 encoded stress response proteins or stress response-like proteins. The most upregulated gene in E745 was *EfmE745_01745* that encoded a predicted stress response protein. This protein is most closely related to YsnF of *Bacillus subtilis* which is a general stress response protein that is upregulated particularly under high heat, high salt and exposure to ethanol (Petersohn *et al.*, 2001). This protein's mechanism of action is unclear but, in line with other stress response proteins, it may protect the cell by arresting the cells metabolism or altering the expression of the vancomycin resistance genes (Poole, 2012). The second most upregulated chromosomal gene in strain E745 was *EfmE745_01664* that encoded a glycosyltransferase family 8 protein. The function of this particular protein is unknown, but this family of proteins is involved in the biosynthesis of carbohydrate molecules (Campbell *et al.*, 1997). Like in E745, many of the most upregulated chromosomal genes in E8202 were poorly characterised. The chromosomal gene *EQB38_RS12340* was the most upregulated chromosomal gene in E8202, encoding a rhodanese-like domain containing protein. This family of proteins have also been linked to the bacterial stress response (Hofmann, Bucher and Kajava, 1998). The second most upregulated chromosomal genes in E8202 was *EQB38_RS12885,* together with gene *EQB38_RS12880*, these genes encode a glycine betaine/L-proline ABC transporter. Bacteria

use this system to import glycine betaine during periods of stress to protect against osmotic stress (Bouvier *et al.*, 2000). This protein is likely used to buffer any osmotic stress caused by the effects of vancomycin on cell wall integrity. Three LysM domain containing proteins were also highly upregulated in E8202. The LysM domain is involved in the binding of proteins to the peptidoglycan of the bacterial cell wall (Buist *et al.*, 2008). LysM domain containing proteins have not been linked to vancomycin resistance in *E. faecium*. However, the upregulation of genes encoding LysM domain containing proteins is important for the infection of mammalian cells by vancomycin-resistant *E. faecium* strains (Cacaci *et al.*, 2018).

As well as many genes being upregulated, there were also many genes that were downregulated in E745 and E8202 when exposed to vancomycin. The gene encoding the ClpP protease was the most downregulated chromosomal gene in E745. ClpP is linked to biofilm formation, stress tolerance and antibiotic tolerance in bacteria (Zheng *et al.*, 2020). A truncation in ClpP in vancomycin sensitive *S. aureus* has been shown to increase tolerance to vancomycin through the thickening of the cell wall and decreased autolytic activity (Shoji *et al.*, 2011). Galactose metabolism was highly downregulated in E8202. A galactokinase and a galactose mutarotase were downregulated 160-fold when exposed to vancomycin. Downregulation of galactose metabolism has been linked with vancomycin resistance in *S. aureus* and daptomycin resistance in *E. faecium* (Shoji *et al.*, 2011; Sinel *et al.*, 2017). Galactose has also been shown to produce toxic cell wall intermediates in *B. subtilis* especially in the presence of cell wall targeting antibiotics, suggesting that a reduction of galactose import promotes *E. faecium* survival during a vancomycin challenge (Habib *et al.*, 2021). Another metabolism pathway in *E. faecium* that was heavily downregulated was the glycerol metabolism pathway. Proteins in this pathway, particularly GlpO have been linked

to biofilm formation but it is unclear how a reduction in biofilm formation would benefit the cell (Lam *et al.*, 2020). It was not only single genes and small operons that were downregulated, whole systems such as carbohydrate transport and prophage induction were also downregulated.

Exposure of strains E745 and E8202 to vancomycin led to the downregulation of the majority of genes in two chromosomally encoded prophages. Prophages are typically induced by external stressors and enter the lytic cycle to create free bacteriophage particles (Nanda, Thormann and Frunzke, 2015). The fluoroquinolone antibiotics have been shown to induce phage in *Enterococcus faecalis* through a DNA damage mechanism (Matos *et al.*, 2013). As the main action of vancomycin does not lead to DNA damage, it is unlikely to cause phage induction. It is possible that expression of phage genes is downregulated to conserve energy and prevent further damage to the cell during vancomycin exposure.

The expression of genes involved in sugar transport across the cell membrane was strongly downregulated in strain E8202 when exposed to vancomycin. *E. faecium* possesses many phosphotranferase system genes that are used to transport sugars into the cell (Deutscher, Francke and Postma, 2006). 71 of the PTS genes for various sugars were downregulated in response to vancomycin exposure in E8202. The *ptsH* gene was downregulated 2.4-fold, this gene encodes the phosphocarrier protein HPr of PTS and has been linked to the repression of cephalosporin resistance genes (Snyder *et al.*, 2014). It is therefore possible that the downregulation of *ptsH* increases the expression of the vancomycin resistance genes. The *ptsD* gene which encodes a mannose PTS family subunit was downregulated 16.5-fold in the presence of vancomycin. PtsD is a virulence factor that contributes to intestinal colonisation of *E. faecium* (Zhang, Top, *et al.*, 2013). This suggests that certain strains of

vancomycin resistant *E. faecium* may exhibit lower gut colonisation efficiency upon exposure to vancomycin. Another cell-surface targeting antibiotic daptomycin has been shown to downregulate the *ptsD* gene 7.5-fold in the *vanB* carrying strain Aus0004 (Sinel *et al.*, 2017).

Despite the transcriptional response of E745 and E8202 being distinct, there was a set of genes in both strains that were similarly differentially regulated when exposed to vancomycin. The known vancomycin resistance genes *vanRSHAX* were all among the genes with the highest upregulation in both E745 and E8202. This confirmed that RNA-seq identified genes that were involved in the resistance to vancomycin. Neither, E745 nor E8202 possessed the prototypical *vanZ* genes but they did possess *vanZ*-like genes on the vancomycin resistance plasmid in E745 and on the chromosome in E8202. The *vanZ*-like gene in E745 was highly upregulated to similar levels as the *vanHAX* genes. This was surprising as *vanZ* is thought to confer resistance to teicoplanin and not vancomycin (Vimberg *et al.*, 2020). As well as the vancomycin resistance genes, E745 and E8202 both had a cluster of 12 genes that were all upregulated. This gene cluster contained at least 4 stress-related proteins and contained others that are likely involved in counteracting oxidative stress. The most upregulated stress response gene in the gene cluster was *gspA*. This gene encodes the general stress protein A that is also upregulated in the presence of bile salts (Zhang, Bierschenk, *et al.*, 2013). Another gene encoding a Mg^{2+} cation transporter was also highly upregulated in this gene cluster. Import of high concentrations of Mg^{2+} has been reported to decrease susceptibility to vancomycin in *S. aureus* but it is unclear whether a similar mechanism occurs in *E. faecium* (Trachsel *et al.*, 2019). Two other small clusters of upregulated genes were shared between E745 and E8202, with the most notable gene in the first gene cluster encoded a PadR family transcriptional regulator. These genes encode

proteins that sense environmental stressors, the most notable of which is phenolic acid (Park *et al.*, 2017). The other gene cluster contained the *liaX* gene which encodes a daptomycin-sensing surface protein. This protein senses daptomycin and remodels the cell envelope of *E. faecalis* to confer resistance (Khan *et al.*, 2019). Remodelling of the cell envelope may decrease access to the cell wall for vancomycin to bind.

There were fewer downregulated genes that were shared between E745 and E8202 than upregulated genes. However, a small gene cluster including the gene encoding AgrB which is part of a quorum sensing system was strongly downregulated in both strains (Zhang and Ji, 2004). This quorum sensing system has been shown to control the expression of genes related to pathogenicity in *S. aureus* (Wang and Muir, 2016). A LytTR family DNA-binding domain-containing protein and a sensor histidine kinase of unknown function were also strongly downregulated. These genes may also be involved in quorum sensing and virulence. The alpha-galactosidase and beta-galactosidase large subunit genes, particularly in E8202, were strongly downregulated. As discussed earlier, galactose metabolism can lead to toxic cell wall intermediates which may further exacerbate the effects of vancomycin exposure (Habib *et al.*, 2021).

Transposon sequencing identified a number of plasmid and chromosomal genes that contributed to increased fitness in the presence of vancomycin. The vancomycin resistance genes *vanR*, *vanS*, *vanH*, *vanA* and *vanX* were among the genes with the largest effect size. The role of these genes in vancomycin resistance is well-established. However, there were also many chromosomally encoded genes where the link to vancomycin resistance was not immediately obvious. Gene *gshAB* encodes the bi-functional protein GshAB which is involved in the production of the antioxidant glutathione (Janowiak and Griffith, 2005).

Glutathione's primary role in the bacterial cell is to manage the oxidation state of protein thiols however it may also play a role in bacterial virulence (Masip, Veeravalli and Georgiou, 2006). Bactericidal antibiotics of which vancomycin is one, have been shown to kill bacteria by the generation of reactive oxygen species (Belenky *et al.*, 2015). As glutathione is a reductant it may be important for protecting the cell against these reactive oxygen species (Ku and Gan, 2021).

It was notable that a number of the genes that increased fitness in the presence of vancomycin were involved in biofilm formation. The *glpO* gene is part of the glycerol metabolism pathway in *Enterococcus faecium*. Like *E. faecalis, E. faecium* has two glycerol metabolism pathways, but can only use the GlpK pathway (Staerck *et al.*, 2021). GlpO has been linked to biofilm formation in *E. faecalis*, which may provide a mechanism for vancomycin tolerance (Lam *et al.*, 2020). The *mgs* gene is homologous to the *bgsB* gene in *E. faecalis*, which encodes a glycosyltransferase that is involved in glycolipid biosynthesis (Zhang *et al.*, 2017). Deletion of *bgsB* in *E. faecalis* was shown to reduce biofilm production by 50% (Theilacker *et al.*, 2011). The gene *asnB* encodes an asparagine synthetase which catalyses the formation of the amino acid asparagine (Humbert and Simoni, 1980). However, it has also been shown that its amidotransferase activity plays a role in virulence, oxidative stress resistance and drug resistance. AsnA and AsnB, both of which are involved in the synthesis of asparagine, were found to be important for virulence in a *Salmonella enterica* serovar Typhimurium mouse infections model and disruption of *asnB* in *Xanthomonas oryzae* pv. oryzicola (Xoc) reduced the expression of nine virulence-related genes (Qian *et al.*, 2013; Jelsbak *et al.*, 2014). Not only does AsnB regulate the expression of virulence genes in Xoc, it also confers resistance to oxidative stress with deletion mutants being more susceptible to hydrogen peroxide than the wildtype (Qian *et al.*, 2013). Resistance to antibiotic drugs has

been linked with AsnB in both *Mycobacterium smegmatis* and *Clostridioides difficile* (Ren and Liu, 2006; Ammam *et al.*, 2020). Disruption of *asnB* in *M. smegmatis* caused hypersensitivity to rifampin, erythromycin and novobiocin through a currently unknown mechanism (Ren and Liu, 2006). In *C. difficile*, the expression of *asnB* was shown to increase in the presence of vancomycin and that this led to the amidation of peptidoglycan precursors. In *C. difficile*, the vancomycin MIC of the *asnB* deletion strain was not different to the wildtype but overexpression of *asnB* caused the strain to become more susceptible to vancomycin (Ammam *et al.*, 2020). The overlapping pair of small genes *avrA* and *avrB*, are currently uncharacterised. However, the presence of a Sec/SPI signal sequence in AvrA and two possible transmembrane domains in AvrB suggests that these proteins are localised to the cell surface so may interact with vancomycin, the peptidoglycan transglycosylases or other cell wall constituents.

Despite several of the Tn-seq identified genes having orthologs in other species, the alleles identified in *E. faecium* E745 and E8202 were somewhat restricted in the wider *Enterococcus* genus. A search of 1644 clade A *E. faecium* isolate genomes demonstrated that genes *arvA*, *avrB*, *asnB* and *gshAB* were broadly conserved within the species *E. faecium*. However, gene *glpO* was missing from 13 of the isolates tested and the *vanX* gene that is typically found on mobile genetic elements was found in only half of the isolates. Across the *Enterococcus* genus, the α-monoglucosyldiacylglycerol synthase gene *mgs* was found in all *Enterococcus* species and two closely related species suggesting that its function is indispensable for all enterococci. The other genes (*avrB*, *asnB*, *glpO* and *gshAB*) were mainly found in the species most closely related to *E. faecium* suggesting that their function is constrained to this relatively limited number of species among the diversity of *Enterococcus*. Vancomycin resistance is more prevalent in *E. faecium* than in any of the

other enterococci (Public Health England, 2019); it is possible that gene *avrA,* which is found solely in *E. faecium,* may provide a selective advantage for the acquisition or retention of mobile genetic elements carrying vancomycin resistance genes.

To confirm the results of the Tn-seq analysis the genes identified were deleted. A novel CRISPR-Cas9 based system was used to delete the genes in *E. faecium* strain E745 (de Maat *et al.*, 2019). The system was successful in knocking out five of the six genes collected, but the vancomycin resistance plasmid was lost in three of the deletion strains. The vancomycin resistance plasmid in strain E745 contained a pAMβ1 replication initiation gene. The CRISPR sequence containing plasmid pVDM1001 possessed a pSH71/pWV01 replication gene. This replication gene did not belong to the same Inc18 incompatibility group indicating that it did not cause the vancomycin plasmid to be lost. However, plasmid pVPL3004, which is used to deliver Cas9 and the tracrRNA, contained the same plasmid replication gene (pAMβ1) as the vancomycin resistance plasmid of *E. faecium* E745 making it the likely cause of the plasmid loss (Oh and van Pijkeren, 2014). To overcome the incompatibility of the pVLP3004 and vancomycin resistance plasmids, the *cas9* gene should be cloned into the pVDM1001 plasmid, although this could create a large plasmid that may be unstable or difficult to electroporate into *E. faecium*.

Although the vancomycin resistance plasmid was lost from three of the deletion isolates, it was still present in E745 *ΔavrB* and E745 *ΔglpO*. Minimum inhibitory concentration assays and growth kinetics in the presence of vancomycin were used to determine whether the mutants had decreased fitness in the presence of vancomycin. The vancomycin MIC for both the wildtype E745 strain and the mutants was 256 μg/ml. This suggested that the deletion of genes *avrB* and *glpO* did not have a major effect on fitness in the presence of vancomycin.

However, an MIC assay also only provides an endpoint results, it does not show whether a mutation affects the growth dynamics of a strain (Wen *et al.*, 2016). To overcome this limitation, the wildtype E745 strain and mutant isolates were grown for 15 hours in the presence and absence of 128 μg/ml (1/2 MIC) vancomycin to investigate growth dynamics. The *ΔglpO* mutant had a very similar growth dynamic to the parent strain suggesting that the deletion of *glpO* does not affect fitness in the presence of vancomycin. However, the *ΔavrB* mutant had a clear fitness defect in the presence of vancomycin compared to the wildtype isolate. The lag phase in the mutant was 4 hours longer than the wildtype and the maximum growth was approximately 50% lower. These data suggest that *avrB* does contribute to increased fitness in the presence of vancomycin. Complementation of the mutant with *avrB* is still required to confirm the role of this gene in vancomycin resistance.

The genes identified in the Tn-seq analysis as contributing to increased fitness were compared to the genes that were differentially expressed when exposed to vancomycin. In both E745 and E8202, the vancomycin resistance genes *vanRSHAX* were identified by Tn-seq and found to be highly upregulated by RNA-seq in the presence of vancomycin. This is consistent with the inducible nature of the *vanHAX* operon which is only expressed at high levels when the two-component system VanRS senses vancomycin (Stogios and Savchenko, 2020). In strain E745 two genes *yvgN* and *dinB_2* were also identified in both the Tn-seq and RNA-seq analysis. These genes encode a general stress response protein and an error prone DNA polymerase (Wagner *et al.*, 1999; Lei *et al.*, 2009). Upregulation of both genes is likely to provide protection and confer the ability to adapt to vancomycin. In E8202, the polyisoprenyl-teichoic acid--peptidoglycan teichoic acid transferase gene *tagU* increased fitness and was upregulated upon exposure to vancomycin (Kawai *et al.*, 2011). An increase in cell wall teichoic acids may reduce the activity of vancomycin by sequestering the

vancomycin away from its binding site (Peschel *et al.*, 2000). A number of the genes that increased fitness in strain E8202 were downregulated when exposed to vancomycin. This was particularly true of the glycerol metabolism genes *glpO*, *glpK* and *glpF* (Bizzini *et al.*, 2010). Deletion of the *glpO* gene in strain E745 had little effect on growth compared to the wildtype when exposed to vancomycin suggesting that the glycerol metabolism genes do not increase the fitness of *E. faecium* in the presence of vancomycin.

It is notable that exposure to vancomycin led to a strain-specific response with the transcriptional response and vancomycin fitness genes varying greatly between the two *vanA*-type *E. faecium* strains E745 and E8202. Strain-specific responses to external factors have previously been noted for plasmid acquisition and antibiotic exposure (Opijnen, Dedrick and Bento, 2016; Dunn *et al.*, 2021). Exposure of two *Streptococcus pneumoniae* strains to another cell envelope targeting antibiotic daptomycin demonstrated markedly different stress responses with genes that were essential in one strain not being required in the other (Opijnen, Dedrick and Bento, 2016). The strain-dependent response of *E. faecium* to vancomycin demonstrates its adaptability to external stress. Having a strain-dependent response to external stressors likely provides more opportunity for a species to adapt due to the differential gene content of strains. As the stress response is not typically specific to an antibiotic, it is possible that this adaptation could also confer increased fitness to other antibiotics or stressors (Poole, 2012).

Although the genes conferring vancomycin resistance in *E. faecium* have been well characterised, this study demonstrated that there is still much to understand about the mechanisms that underpin vancomycin resistance in *E. faecium*. Identification of genes that increase the fitness of *E. faecium* in the presence of vancomycin may provide novel drug

targets. Drugs that target auxiliary vancomycin resistance proteins, when used in combination with vancomycin may lead to a quicker resolution of disease and better clinical outcomes than vancomycin alone.

5.5.1 Future work

The deletion mutants for genes *avrA*, *asnB*, *gshAB* and *mgs* should be remade without the loss of the vancomycin resistance plasmid. Two approaches could be used to obtain these mutants. Firstly, a larger number of potential deletion mutants should be screened to identify mutants that still carry the vancomycin resistance plasmid. Another approach would be to create a new plasmid that contains the CRISPR sequence, the tracrRNA, the Cas9 gene and the deletion/insertion template and does not have incompatibility issues with the vancomycin resistance plasmid. The genes identified by Tn-seq could be overexpressed to determine whether these genes can change the vancomycin MIC in *E. faecium* E745. Deletion of gene *avrB* appeared to affect the fitness of E745 in the presence of vancomycin but to confirm this observation the deletion will have to be complemented with the wildtype *avrB* gene. As AvrA and AvrB are currently uncharacterised proteins, techniques used to infer the structure such as X-ray crystallography and cryo-EM would provide valuable insights into the possible function of the proteins. Pull-down assays could also be used to identify binding partners with known function, again providing insights into the function of AvrA and AvrB. To confirm that AvrA and AvrB are cell surface proteins, they could be fluorescently tagged, and fluorescence microscopy used to visualise the location of the proteins on the cell. High-throughput knockdown techniques such as CRISPR interference could be used in vancomycin-susceptible strains to target genes identified as being highly downregulated in our RNA-seq data.

5.6 Key findings

- RNA-seq identified the known vancomycin resistance genes, *vanRSHAX,* as being among the most upregulated genes in both *E. faecium* E745 and E8202.
- Sugar transport is heavily downregulated in strain E8202, 71 phosphotransferase genes were downregulated when exposed to vancomycin.
- Prophage genes were differentially expressed upon exposure to vancomycin. Genes involved in lysogeny were upregulated in strain E745 and genes involved in head and tail morphogenesis were downregulated, in strain E8202 both the lysogeny related genes and the genes involved in head and tail morphogenesis were downregulated.
- Tn-seq was able to identify the *vanHAX* vancomycin resistance genes among those that gave the greatest fitness advantage in the presence of sub-MIC concentrations of vancomycin.
- Tn-seq also identified a number of chromosomal genes that contribute to fitness in the presence of vancomycin. Two currently uncharacterised genes *avrA* and *avrB,* were highly upregulated in the RNA-seq analysis and were identified as having a large effect size in the Tn-seq analysis. Gene *avrB* was only found in species closely related to *E. faecium* and gene *avrA* was found solely in *E. faecium*. Signal peptide analysis and protein topology analysis suggested that genes *avrA* and *avrB* may encode cell surface-associated proteins.
- Strain E745 with a deletion of the *avrB* gene had a growth defect in the presence of ½ MIC vancomycin. This suggested that gene *avrB* does increase fitness under vancomycin exposure.

Chapter 6
Final Discussion

6.1 Final discussion

Antibiotic resistance is a major threat to public health by endangering the treatment of bacterial infections. To overcome antibiotic resistance, it must be tackled from many different angles utilising a "One Health" approach. The role of the natural and clinical environments in the selection and spread of antibiotic resistance genes and the mechanisms that bacteria use to overcome antibiotics must be studied in order to design novel interventions that not only treat antibiotic resistance infections but also stop them from arising in the first place.

The overuse of antibiotics in healthcare and agri/aquaculture in low-and-middle income countries is a cause for concern. Antibiotics are used widely in aquaculture farms in Bangladesh despite being restricted by law. It was hypothesised that the widespread use of antibiotics in aquaculture would drive antibiotic resistance in the environment. However, in chapter 3 I showed that the use of antibiotics in aquaculture had little effect on the selection of antibiotic resistant bacteria in the environment. It was found instead that urban water bodies including rivers and lakes had much higher levels of antibiotic resistance genes than the rural aquaculture ponds. These urban water bodies also had higher levels of human-associated gut bacteria. Source-sink analysis was able to correlate the levels of antibiotic resistance genes with the levels of human-associated gut bacteria suggesting that the release of untreated human waste into these urban water bodies is driving antibiotic resistance in the environment. As theses bodies of water are often used for personal hygiene and the cleaning of clothes and cooking equipment it is imperative that interventions are put in place that stop the release of human waste into these water bodies. This would likely reduce the incidence of gut colonisation by multidrug-resistant bacteria of those living in Bangladesh and could then reduce the prevalence antibiotic-resistant infections.

As well as the heavy use of antibiotics in aquaculture, antibiotics are often used for the treatment of disease in a clinical setting. In chapter 4, I investigated antibiotic resistance in a hospital through the lens of vancomycin resistant *Enterococcus faecium*. I showed that a local Birmingham hospital had a diverse population of vancomycin resistant *E. faecium* and demonstrated that the sharing of strains between different patients was common during a VRE outbreak in 2016-2017. The isolates contained a diverse population of plasmids, some of which carried antibiotic resistance genes. Long-read sequencing was used in an attempt to contextualise the vancomycin resistance genes. However, this was only possible for a minority of strains. The vancomycin resistance genes were located on plasmids with three different replication initiation gene types which have previously been shown to carry vancomycin resistance genes. Both the vancomycin-resistant and -susceptible *E. faecium* isolates were multidrug resistant with all isolates carrying six or more antibiotic resistance genes. Resistance to vancomycin, erythromycin and aminoglycoside drugs was widespread among the isolates. Three isolates were identified as vancomycin-variable, as they carried vancomycin resistance genes but were phenotypically susceptible to vancomycin. Disruption of the *vanR* gene and promoter region of the vancomycin resistance two-component system was the likely cause of their susceptible phenotype. These isolates were also able to rapidly revert to a vancomycin resistant phenotype when exposed to low concentrations of vancomycin. A novel mechanism of reversion was uncovered that saw the insertion of the vancomycin resistance plasmid into the chromosome and expression of the vancomycin resistance genes driven by a ribosomal RNA operon. Identification of vancomycin-variable isolates that can rapidly revert to a resistant phenotype has worrying implications for the treatment of *E. faecium* infections in the clinic. *E. faecium* isolates phenotyped as susceptible to vancomycin could revert to a resistant phenotype within the

patient leading to a failure of treatment and poor clinical outcomes. Should VVE isolates make up a larger proportion of clinical *E. faecium* isolates it may be necessary to implement targeted PCR approaches or whole genome shotgun sequencing to identify these isolates and prevent the incorrect use of vancomycin.

The mechanism of antibiotic resistance genes is often well characterised in bacterial pathogens. However, the wider transcriptional response and accessory genes that contribute to increased fitness in the presence of the antibiotic have not been well characterised. Vancomycin resistance in *E. faecium* is most often conferred by the *vanA*-type vancomycin resistance genes. These genes replace the terminal D-alanyl D-alanine of the peptidoglycan precursors with D-alanyl D-lactate which vancomycin binds to 1000-fold less strongly (Stogios and Savchenko, 2020). In chapter 5, I used RNA-seq and Tn-seq to identify accessory genes that are involved in the response to vancomycin. RNA-seq identified that genes involved in the response to stress were upregulated and the transport of sugars across the membrane was highly downregulated. In addition, Tn-seq identified genes involved in biofilm formation and detoxification of reactive oxygen as conferring increased fitness in the presence of vancomycin. Although deletion of these genes did not affect the vancomycin MIC, in the case of gene *avrB* it led to a growth defect in the presence of vancomycin. Identification of accessory genes involved in the response to vancomycin could provide novel targets for the development of drugs that work in combination with vancomycin to increase the efficacy of the treatment.

In conclusion, this thesis has demonstrated that although the regulation of antibiotic use is needed in LMIC countries, it is arguably more important that the global population has equal access to sanitation to prevent the development and spread of antibiotic resistance. I have

also demonstrated that more research is needed into the evolution and monitoring of antibiotic resistance in *E. faecium*. Novel mechanisms such as those found in VVE present a real threat to the treatment of infections, and only by studying these mechanisms in the laboratory will we able to identify and treat them accordingly in clinical practice. A final take-home message of this thesis is that in order to overcome antibiotic resistance novel treatments must be developed. These new treatments can only be developed when we fully understand how our current antibiotics work. Understanding this will allow us to identify novel targets for antibiotic drugs but also to design drugs rationally to overcome the evolutionary challenges that our current drugs face.

Chapter 7 References

Abbott, I.J., van Gorp, E., van der Meijden, A., Wijma, R.A., Meletiadis, J., Roberts, J.A., Mouton, J.W. and Peleg, A.Y. (2020) 'Oral Fosfomycin Treatment for Enterococcal Urinary Tract Infections in a Dynamic In Vitro Model', *Antimicrobial Agents and Chemotherapy*, 64(6), pp. e00342-20. doi:10.1128/AAC.00342-20.

Adamecz, Z., Nielsen, K.L., Kirkby, N.S. and Frimodt-Møller, N. (2021) 'Aminoglycoside resistance genes in *Enterococcus faecium*: mismatch with phenotype', *Journal of Antimicrobial Chemotherapy*, 76(8), pp. 2215–2217. doi:10.1093/jac/dkab137.

Adams, M.D., Kelley, J.M., Gocayne, J.D., Dubnick, M., Polymeropoulos, M.H., Xiao, H., Merril, C.R., Wu, A., Olde, B., Moreno, R.F., Kerlavage, A.R., McCombie, W.R. and Venter, J.C. (1991) 'Complementary DNA Sequencing: Expressed Sequence Tags and Human Genome Project', *Science*, 252(5013), pp. 1651-1656. doi:10.1126/science.2047873.

Adams, R.I., Miletto, M., Lindow, S.E., Taylor, J.W. and Bruns, T.D. (2014) 'Airborne bacterial communities in residences: similarities and differences with fungi', *PloS One*, 9(3), p. e91283. doi:10.1371/journal.pone.0091283.

Agramont, J., Gutiérrez-Cortez, S., Joffré, E., Sjöling, Å. and Calderon Toledo, C. (2020) 'Fecal Pollution Drives Antibiotic Resistance and Class 1 Integron Abundance in Aquatic Environments of the Bolivian Andes Impacted by Mining and Wastewater', *Microorganisms*, 8(8), p. 1122. doi:10.3390/microorganisms8081122.

Ahmad, M.K., Islam, S., Rahman, M.S., Haque, M.R. and Islam, M.M. (2010) 'Heavy Metals in Water, Sediment and Some Fishes of Buriganga River, Bangladesh', *International Journal of Environmental Research*, 4(2), pp. 321–332. doi:10.22059/ijer.2010.24.

Ahmed, I., Rabbi, Md.B. and Sultana, S. (2019) 'Antibiotic resistance in Bangladesh: A systematic review', *International Journal of Infectious Diseases*, 80, pp. 54–61. doi:10.1016/j.ijid.2018.12.017.

Ahmed, M.O. and Baptiste, K.E. (2018) 'Vancomycin-Resistant Enterococci: A Review of Antimicrobial Resistance Mechanisms and Perspectives of Human and Animal Health', *Microbial Drug Resistance*, 24(5), pp. 590–606. doi:10.1089/mdr.2017.0147.

Akter, K.S., Kurisu, K. and Hanaki, K. (2017) 'Water Use and Pollution Recognition from the Viewpoint of Local Residents in Dhaka, Bangladesh', *Water*, 9(5), p. 331. doi:10.3390/w9050331.

Alcock, B.P., Raphenya, A.R., Lau, T.T.Y., Tsang, K.K., Bouchard, M., Edalatmand, A., Huynh, W., Nguyen, A.-L.V., Cheng, A.A., Liu, S., Min, S.Y., Miroshnichenko, A., Tran, H.-K., Werfalli, R.E., Nasir, J.A., Oloni, M., Speicher, D.J., Florescu, A., Singh, B., Faltyn, M., Hernandez-Koutoucheva, A., Sharma, A.N., Bordeleau, E., Pawlowski, A.C., Zubyk, H.L., Dooley, D., Griffiths, E., Maguire, F., Winsor, G.L., Beiko, R.G., Brinkman, F.S.L., Hsiao, W.W.L., Domselaar, G.V. and McArthur, A.G. (2020) 'CARD 2020: antibiotic

resistome surveillance with the comprehensive antibiotic resistance database', *Nucleic Acids Research*, 48(D1), pp. D517–D525. doi:10.1093/nar/gkz935.

Ali, H., Rico, A., Murshed-e-Jahan, K. and Belton, B. (2016) 'An assessment of chemical and biological product use in aquaculture in Bangladesh', *Aquaculture*, 454, pp. 199–209. doi:10.1016/j.aquaculture.2015.12.025.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) 'Basic local alignment search tool', *Journal of Molecular Biology*, 215(3), pp. 403–410. doi:10.1016/S0022-2836(05)80360-2.

Amarasinghe, S.L., Su, S., Dong, X., Zappia, L., Ritchie, M.E. and Gouil, Q. (2020) 'Opportunities and challenges in long-read sequencing data analysis', *Genome Biology*, 21(1), p. 30. doi:10.1186/s13059-020-1935-5.

Amin, N., Rahman, M., Raj, S., Ali, S., Green, J., Das, S., Doza, S., Mondol, M.H., Wang, Y., Islam, M.A., Alam, M.-U., Huda, T.M.N., Haque, S., Unicomb, L., Joseph, G. and Moe, C.L. (2019) 'Quantitative assessment of fecal contamination in multiple environmental sample types in urban communities in Dhaka, Bangladesh using SaniPath microbial approach', *PLOS ONE*, 14(12), p. e0221193. doi:10.1371/journal.pone.0221193.

Ammam, F., Patin, D., Coullon, H., Blanot, D., Lambert, T., Mengin-Lecreulx, D. and Candela, T. (2020) 'AsnB is responsible for peptidoglycan precursor amidation in *Clostridium difficile* in the presence of vancomycin', *Microbiology (Reading, England)*, 166(6), pp. 567–578. doi:10.1099/mic.0.000917.

Antonelli, A., D'Andrea, M.M., Brenciani, A., Galeotti, C.L., Morroni, G., Pollini, S., Varaldo, P.E. and Rossolini, G.M. (2018) 'Characterization of *poxtA*, a novel phenicol–oxazolidinone–tetracycline resistance gene from an MRSA of clinical origin', *Journal of Antimicrobial Chemotherapy*, 73(7), pp. 1763–1769. doi:10.1093/jac/dky088.

Arias Granada, Yurani Haque, Sabrina Sharmin, George Joseph, and Monica Yanez Pagans (2018) *Water and sanitation in Dhaka slums : access, quality, and informality in service provision (English)*. Policy Research working paper; no. WPS. World Bank Group.

Arndt, D., Grant, J.R., Marcu, A., Sajed, T., Pon, A., Liang, Y. and Wishart, D.S. (2016) 'PHASTER: a better, faster version of the PHAST phage search tool', *Nucleic Acids Research*, 44(W1), pp. W16-21. doi:10.1093/nar/gkw387.

Arredondo-Alonso, S., Top, J., McNally, A., Puranen, S., Pesonen, M., Pensar, J., Marttinen, P., Braat, J.C., Rogers, M.R.C., Schaik, W. van, Kaski, S., Willems, R.J.L., Corander, J. and Schürch, A.C. (2020) 'Plasmids Shaped the Recent Emergence of the Major Nosocomial Pathogen *Enterococcus faecium*', *mBio*, 11(1), pp. e03284-19. doi:10.1128/mBio.03284-19.

Arredondo-Alonso, S., Willems, R.J., van Schaik, W. and Schürch, A.C. (2017) 'On the (im)possibility of reconstructing plasmids from whole-genome short-read sequencing data', *Microbial Genomics*, 3(10). doi:10.1099/mgen.0.000128.

Arthur, M., Depardieu, F., Cabanié, L., Reynolds, P. and Courvalin, P. (1998) 'Requirement of the VanY and VanX D,D-peptidases for glycopeptide resistance in enterococci', *Molecular Microbiology*, 30(4), pp. 819–830. doi:10.1046/j.1365-2958.1998.01114.x.

Arthur, M., Molinas, C., Depardieu, F. and Courvalin, P. (1993) 'Characterization of Tn1546, a Tn3-related transposon conferring glycopeptide resistance by synthesis of depsipeptide peptidoglycan precursors in *Enterococcus faecium* BM4147', *Journal of Bacteriology*, 175(1), pp. 117–127. doi:10.1128/jb.175.1.117-127.1993.

Auckland, C., Teare, L., Cooke, F., Kaufmann, M.E., Warner, M., Jones, G., Bamford, K., Ayles, H. and Johnson, A.P. (2002) 'Linezolid-resistant enterococci: report of the first isolates in the United Kingdom', *Journal of Antimicrobial Chemotherapy*, 50(5), pp. 743–746. doi:10.1093/jac/dkf246.

Bager, F., Madsen, M., Christensen, J. and Aarestrup, F.M. (1997) 'Avoparcin used as a growth promoter is associated with the occurrence of vancomycin-resistant *Enterococcus faecium* on Danish poultry and pig farms', *Preventive Veterinary Medicine*, 31(1–2), pp. 95–112. doi:10.1016/s0167-5877(96)01119-1.

Baldi, P. and Hatfield, G.W. (2002) *DNA Microarrays and Gene Expression: From Experiments to Data Analysis and Modeling*. Cambridge: Cambridge University Press. pp. 7-28. doi:10.1017/CBO9780511541773.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A. and Pevzner, P.A. (2012) 'SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing', *Journal of Computational Biology*, 19(5), pp. 455–477. doi:10.1089/cmb.2012.0021.

Baquero, F. and Levin, B.R. (2021) 'Proximate and ultimate causes of the bactericidal action of antibiotics', *Nature Reviews Microbiology*, 19(2), pp. 123–132. doi:10.1038/s41579-020-00443-1.

Beattie, R.E., Bandla, A., Swarup, S. and Hristova, K.R. (2020) 'Freshwater Sediment Microbial Communities Are Not Resilient to Disturbance From Agricultural Land Runoff', *Frontiers in Microbiology*, 11. pp. 11:539921. doi: 10.3389/fmicb.2020.539921.

Belenky, P., Ye, J.D., Porter, C.B.M., Cohen, N.R., Lobritz, M.A., Ferrante, T., Jain, S., Korry, B.J., Schwarz, E.G., Walker, G.C. and Collins, J.J. (2015) 'Bactericidal Antibiotics Induce Toxic Metabolic Perturbations that Lead to Cellular Damage', *Cell Reports*, 13(5), pp. 968–980. doi:10.1016/j.celrep.2015.09.059.

Belloso Daza, M.V., Cortimiglia, C., Bassi, D. and Cocconcelli, P.S. 2021 (2021) 'Genome-based studies indicate that the *Enterococcus faecium* Clade B strains belong to *Enterococcus lactis* species and lack of the hospital infection associated markers', *International Journal of Systematic and Evolutionary Microbiology*, 71(8), p. 004948. doi:10.1099/ijsem.0.004948.

Bender, J.K., Fiedler, S., Klare, I. and Werner, G. (2015) 'Complete Genome Sequence of the Gut Commensal and Laboratory Strain *Enterococcus faecium* 64/3', *Genome Announcements*, 3(6), pp. e01275-15. doi:10.1128/genomeA.01275-15.

Berglund, B., Hoang, N.T.B., Tärnberg, M., Le, N.K., Svartström, O., Khu, D.T.K., Nilsson, M., Le, H.T., Welander, J., Olson, L., Larsson, M., Nilsson, L.E. and Hanberger, H. (2018) 'Insertion sequence transpositions and point mutations in *mgrB* causing colistin resistance in a clinical strain of carbapenem-resistant *Klebsiella pneumoniae* from Vietnam', *International Journal of Antimicrobial Agents*, 51(5), pp. 789–793. doi:10.1016/j.ijantimicag.2017.11.012.

Bharti, R. and Grimm, D.G. (2021) 'Current challenges and best-practice protocols for microbiome analysis', *Briefings in Bioinformatics*, 22(1), pp. 178–193. doi:10.1093/bib/bbz155.

Bizzini, A., Zhao, C., Budin-Verneuil, A., Sauvageot, N., Giard, J.-C., Auffray, Y. and Hartke, A. (2010) 'Glycerol Is Metabolized in a Complex and Strain-Dependent Manner in *Enterococcus faecalis*', *Journal of Bacteriology*, 192(3), pp. 779–785. doi:10.1128/JB.00959-09.

Blackwell, G.A., Hunt, M., Malone, K.M., Lima, L., Horesh, G., Alako, B.T.F., Thomson, N.R. and Iqbal, Z. (2021) 'Exploring bacterial diversity via a curated and searchable snapshot of archived DNA sequences', *PLOS Biology*, 19(11), p. e3001421. doi:10.1371/journal.pbio.3001421.

Blair, J.M.A., Webber, M.A., Baylay, A.J., Ogbolu, D.O. and Piddock, L.J.V. (2015) 'Molecular mechanisms of antibiotic resistance', *Nature Reviews Microbiology*, 13(1), pp. 42–51. doi:10.1038/nrmicro3380.

Bodor, A., Bounedjoum, N., Vincze, G.E., Erdeiné Kis, Á., Laczi, K., Bende, G., Szilágyi, Á., Kovács, T., Perei, K. and Rákhely, G. (2020) 'Challenges of unculturable bacteria: environmental perspectives', *Reviews in Environmental Science and Bio/Technology*, 19(1), pp. 1–22. doi:10.1007/s11157-020-09522-4.

Boinett, C.J., Cain, A.K., Hawkey, J., Do Hoang, N.T., Khanh, N.N.T., Thanh, D.P., Dordel, J., Campbell, J.I., Lan, N.P.H., Mayho, M., Langridge, G.C., Hadfield, J., Chau, N.V.V., Thwaites, G.E., Parkhill, J., Thomson, N.R., Holt, K.E. and Baker, S. (2019) 'Clinical and laboratory-induced colistin-resistance mechanisms in *Acinetobacter baumannii*', *Microbial Genomics*, 5(2), p. e000246. doi:10.1099/mgen.0.000246.

Bolger, A.M., Lohse, M. and Usadel, B. (2014) 'Trimmomatic: a flexible trimmer for Illumina sequence data', *Bioinformatics*, 30(15), pp. 2114–2120. doi:10.1093/bioinformatics/btu170.

Bouvier, J., Bordes, P., Romeo, Y., Fourçans, A., Bouvier, I. and Gutierrez, C. (2000) 'Characterization of OpuA, a glycine-betaine uptake system of *Lactococcus lactis*', *Journal of molecular microbiology and biotechnology*, 2(2), pp. 199–205.

Bray, N.L., Pimentel, H., Melsted, P. and Pachter, L. (2016) 'Near-optimal probabilistic RNA-seq quantification', *Nature Biotechnology*, 34(5), pp. 525–527. doi:10.1038/nbt.3519.

Bryan, E.M., Bae, T., Kleerebezem, M. and Dunny, G.M. (2000) 'Improved vectors for nisin-controlled expression in gram-positive bacteria', *Plasmid*, 44(2), pp. 183–190. doi:10.1006/plas.2000.1484.

Buckner, M.M.C., Ciusa, M.L., Meek, R.W., Moorey, A.R., McCallum, G.E., Prentice, E.L., Reid, J.P., Alderwick, L.J., Di Maio, A. and Piddock, L.J.V. (2020) 'HIV Drugs Inhibit Transfer of Plasmids Carrying Extended-Spectrum β-Lactamase and Carbapenemase Genes', *mBio*, 11(1), pp. e03355-19. doi:10.1128/mBio.03355-19.

Buist, G., Steen, A., Kok, J. and Kuipers, O.P. (2008) 'LysM, a widely distributed protein motif for binding to (peptido)glycans', *Molecular Microbiology*, 68(4), pp. 838–847. doi:10.1111/j.1365-2958.2008.06211.x.

Bukin, Yu.S., Galachyants, Yu.P., Morozov, I.V., Bukin, S.V., Zakharenko, A.S. and Zemskaya, T.I. (2019) 'The effect of 16S rRNA region choice on bacterial community metabarcoding results', *Scientific Data*, 6, p. 190007. doi:10.1038/sdata.2019.7.

Bussiere, D.E., Pratt, S.D., Katz, L., Severin, J.M., Holzman, T. and Park, C.H. (1998) 'The Structure of VanX Reveals a Novel Amino-Dipeptidase Involved in Mediating Transposon-Based Vancomycin Resistance', *Molecular Cell*, 2(1), pp. 75–84. doi:10.1016/S1097-2765(00)80115-X.

Cacaci, M., Giraud, C., Leger, L., Torelli, R., Martini, C., Posteraro, B., Palmieri, V., Sanguinetti, M., Bugli, F. and Hartke, A. (2018) 'Expression profiling in a mammalian host reveals the strong induction of genes encoding LysM domain-containing proteins in *Enterococcus faecium*', *Scientific Reports*, 8, p. 12412. doi:10.1038/s41598-018-30882-z.

Cain, A.K., Barquist, L., Goodman, A.L., Paulsen, I.T., Parkhill, J. and van Opijnen, T. (2020) 'A decade of advances in transposon-insertion sequencing', *Nature Reviews Genetics*, 21(9), pp. 526–540. doi:10.1038/s41576-020-0244-x.

Campbell, J.A., Davies, G.J., Bulone, V. and Henrissat, B. (1997) 'A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities', *The Biochemical journal*, 326 (Pt 3), pp. 929–939. doi:10.1042/bj3260929u.

Cannatelli, A., Di Pilato, V., Giani, T., Arena, F., Ambretti, S., Gaibani, P., D'Andrea, M.M. and Rossolini, G.M. (2014) 'In vivo evolution to colistin resistance by PmrB sensor kinase mutation in KPC-producing *Klebsiella pneumoniae* is associated with low-dosage colistin treatment', *Antimicrobial Agents and Chemotherapy*, 58(8), pp. 4399–4403. doi:10.1128/AAC.02555-14.

Cantalapiedra, C.P., Hernández-Plaza, A., Letunic, I., Bork, P. and Huerta-Cepas, J. (2021) 'eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale', *Molecular Biology and Evolution*, 38(12), pp. 5825–5829. doi:10.1093/molbev/msab293.

Carattoli, A., Zankari, E., García-Fernández, A., Voldby Larsen, M., Lund, O., Villa, L., Møller Aarestrup, F. and Hasman, H. (2014) 'In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing', *Antimicrobial Agents and Chemotherapy*, 58(7), pp. 3895–3903. doi:10.1128/AAC.02412-14.

Casas, L., Karvonen, A.M., Kirjavainen, P.V., Täubel, M., Hyytiäinen, H., Jayaprakash, B., Lehmann, I., Standl, M., Pekkanen, J. and Heinrich, J. (2019) 'Early life home microbiome and hyperactivity/inattention in school-age children', *Scientific Reports*, 9(1), p. 17355. doi:10.1038/s41598-019-53527-1.

Centers for Disease Control & Prevention (2013) *Antibiotic Resistance Threats in the United States, 2013*.

Cetinkaya, Y., Falk, P. and Mayhall, C.G. (2000) 'Vancomycin-Resistant Enterococci', *Clinical Microbiology Reviews*, 13(4), pp. 686–707. doi:10.1128/CMR.13.4.686

Chain, E., Florey, H.W., Gardner, A.D., Heatley, N.G., Jennings, M.A., Orr-Ewing, J. and Sanders, A.G. (1940) 'Penicillin As a Chemotherapeutic Agent', *The Lancet*, 236(6104), pp. 226–228. doi:10.1016/S0140-6736(01)08728-1.

Chan, J.Z.-M., Halachev, M.R., Loman, N.J., Constantinidou, C. and Pallen, M.J. (2012) 'Defining bacterial species in the genomic era: insights from the genus *Acinetobacter'*, *BMC Microbiology*, 12(1), p. 302. doi:10.1186/1471-2180-12-302.

Chao, M.C., Abel, S., Davis, B.M. and Waldor, M.K. (2016) 'The design and analysis of transposon insertion sequencing experiments', *Nature Reviews Microbiology*, 14(2), pp. 119–128. doi:10.1038/nrmicro.2015.7.

Chen, S., Zhou, Y., Chen, Y. and Gu, J. (2018) 'fastp: an ultra-fast all-in-one FASTQ preprocessor', *Bioinformatics*, 34(17), pp. i884–i890. doi:10.1093/bioinformatics/bty560.

Chen, You-Han, Lin, S.-Y., Lin, Y.-T., Tseng, S.-P., Chang, C.-C., Yu, S.-Y., Hung, W.-W., Jao, Y.-T., Lin, C.-Y., Chen, Yen-Hsu and Hung, W.-C. (2021) 'Emergence of *aac(6')-Ie-aph(2'')-Ia*-positive enterococci with non-high-level gentamicin resistance mediated by IS1216V: adaptation to decreased aminoglycoside usage in Taiwan', *Journal of Antimicrobial Chemotherapy*, 76(7), pp. 1689–1697. doi:10.1093/jac/dkab071.

Chen, Z., Erickson, D.L. and Meng, J. (2020) 'Benchmarking hybrid assembly approaches for genomic analyses of bacterial pathogens using Illumina and Oxford Nanopore sequencing', *BMC Genomics*, 21(1), p. 631. doi:10.1186/s12864-020-07041-8.

Chiu, C.Y. and Miller, S.A. (2019) 'Clinical metagenomics', *Nature Reviews Genetics*, 20(6), pp. 341–355. doi:10.1038/s41576-019-0113-7.

Chokshi, A., Sifri, Z., Cennimo, D. and Horng, H. (2019) 'Global Contributors to Antibiotic Resistance', *Journal of Global Infectious Diseases*, 11(1), pp. 36–42. doi:10.4103/jgid.jgid_110_18.

Chow, J.W. (2000) 'Aminoglycoside Resistance in Enterococci', *Clinical Infectious Diseases*, 31(2), pp. 586–589. doi:10.1086/313949.

Chung, C.T., Niemela, S.L. and Miller, R.H. (1989) 'One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution.', *Proceedings of the National Academy of Sciences*, 86(7), pp. 2172–2175. doi:10.1073/pnas.86.7.2172.

Clewell, D.B., Victoria Francia, M., Flannagan, S.E. and An, F.Y. (2002) 'Enterococcal plasmid transfer: sex pheromones, transfer origins, relaxases, and the *Staphylococcus aureus* issue', *Plasmid*, 48(3), pp. 193–201. doi:10.1016/S0147-619X(02)00113-0.

Collignon, P., Beggs, J.J., Walsh, T.R., Gandra, S. and Laxminarayan, R. (2018) 'Anthropological and socioeconomic factors contributing to global antimicrobial resistance: a univariate and multivariable analysis', *The Lancet Planetary Health*, 2(9), pp. e398–e405. doi:10.1016/S2542-5196(18)30186-4.

Connell, S.R., Tracz, D.M., Nierhaus, K.H. and Taylor, D.E. (2003) 'Ribosomal Protection Proteins and Their Mechanism of Tetracycline Resistance', *Antimicrobial Agents and Chemotherapy*, 47(12), pp. 3675-3681. doi:10.1128/AAC.47.12.3675-3681.2003.

Costa, Y., Galimand, M., Leclercq, R., Duval, J. and Courvalin, P. (1993) 'Characterization of the chromosomal *aac(6')-Ii* gene specific for *Enterococcus faecium*.', *Antimicrobial Agents and Chemotherapy*, 37(9), pp. 1896–1903.

Coyne, S., Rosenfeld, N., Lambert, T., Courvalin, P. and Périchon, B. (2010) 'Overexpression of resistance-nodulation-cell division pump AdeFGH confers multidrug resistance in *Acinetobacter baumannii*', *Antimicrobial Agents and Chemotherapy*, 54(10), pp. 4389–4393. doi:10.1128/AAC.00155-10.

Croucher, N.J., Page, A.J., Connor, T.R., Delaney, A.J., Keane, J.A., Bentley, S.D., Parkhill, J. and Harris, S.R. (2015) 'Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins', *Nucleic Acids Research*, 43(3), p. e15. doi:10.1093/nar/gku1196.

Damier-Piolle, L., Magnet, S., Brémont, S., Lambert, T. and Courvalin, P. (2008) 'AdeIJK, a resistance-nodulation-cell division pump effluxing multiple antibiotics in *Acinetobacter baumannii*', *Antimicrobial Agents and Chemotherapy*, 52(2), pp. 557–562. doi:10.1128/AAC.00732-07.

Darnell, R.L., Knottenbelt, M.K., Rose, F.O.T., Monk, I.R., Stinear, T.P. and Cook, G.M. (2019) 'Genomewide Profiling of the *Enterococcus faecalis* Transcriptional Response to Teixobactin Reveals CroRS as an Essential Regulator of Antimicrobial Tolerance', *mSphere*, 4(3), pp. e00228-19. doi:10.1128/mSphere.00228-19.

D'Arpa, P., Karna, S.L.R., Chen, T. and Leung, K.P. (2021) '*Pseudomonas aeruginosa* transcriptome adaptations from colonization to biofilm infection of skin wounds', *Scientific Reports*, 11(1), p. 20632. doi:10.1038/s41598-021-00073-4.

De Oliveira, D.M.P., Forde, B.M., Kidd, T.J., Harris, P.N.A., Schembri, M.A., Beatson, S.A., Paterson, D.L. and Walker, M.J. (2020) 'Antimicrobial Resistance in ESKAPE

Pathogens', *Clinical Microbiology Reviews*, 33(3), pp. e00181-19. doi:10.1128/CMR.00181-19.

Delcour, A.H. (2009) 'Outer membrane permeability and antibiotic resistance', *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1794(5), pp. 808–816. doi:10.1016/j.bbapap.2008.11.005.

de Maat, V., Stege, P.B., Dedden, M., Hamer, M., van Pijkeren, J.-P., Willems, R.J.L. and van Schaik, W. (2019) 'CRISPR-Cas9-mediated genome editing in vancomycin-resistant *Enterococcus faecium*', *FEMS Microbiology Letters*, 366(22), p. fnz256. doi:10.1093/femsle/fnz256.

Derilus, D., Godoy-Vitorino, F., Rosado, H., Agosto, E., Dominguez-Bello, M.G. and Cavallin, H. (2020) 'An in-depth survey of the microbial landscape of the walls of a neonatal operating room', *PLOS ONE*, 15(4), p. e0230957. doi:10.1371/journal.pone.0230957.

Deutscher, J., Francke, C. and Postma, P.W. (2006) 'How Phosphotransferase System-Related Protein Phosphorylation Regulates Carbohydrate Metabolism in Bacteria', *Microbiology and Molecular Biology Reviews*, 70(4), pp. 939–1031. doi:10.1128/MMBR.00024-06.

Diekema, D.J., Hsueh, P.-R., Mendes, R.E., Pfaller, M.A., Rolston, K.V., Sader, H.S. and Jones, R.N. (2019) 'The Microbiology of Bloodstream Infection: 20-Year Trends from the SENTRY Antimicrobial Surveillance Program', *Antimicrobial Agents and Chemotherapy*, 63(7), pp. e0335519. doi:10.1128/AAC.00355-19.

Ding, L.-J., Zhou, X.-Y. and Zhu, Y.-G. (2020) 'Microbiome and antibiotic resistome in household dust from Beijing, China', *Environment International*, 139, p. 105702. doi:10.1016/j.envint.2020.105702.

Downing, M.A., Xiong, J., Eshaghi, A., McGeer, A., Patel, S.N. and Johnstone, J. (2015) 'Vancomycin-Variable Enterococcal Bacteremia', *Journal of Clinical Microbiology*, 53(12), pp. 3951–3953. doi:10.1128/JCM.02046-15.

Dunn, S., Carrilero, L., Brockhurst, M. and McNally, A. (2021) 'Limited and Strain-Specific Transcriptional and Growth Responses to Acquisition of a Multidrug Resistance Plasmid in Genetically Diverse *Escherichia coli* Lineages', *mSystems*, 6(2), pp. e00083-21. doi:10.1128/mSystems.00083-21.

Dunny, G.M. (2007) 'The peptide pheromone-inducible conjugation system of *Enterococcus faecalis* plasmid pCF10: cell–cell signalling, gene transfer, complexity and evolution', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1483), pp. 1185–1193. doi:10.1098/rstb.2007.2043.

Ebmeyer, S., Kristiansson, E. and Larsson, D.G.J. (2021) 'A framework for identifying the recent origins of mobile antibiotic resistance genes', *Communications Biology*, 4(1), pp. 1–10. doi:10.1038/s42003-020-01545-5.

Edgar, R.C. (2018) 'Accuracy of taxonomy prediction for 16S rRNA and fungal ITS sequences', *PeerJ*, 6, p. e4652. doi:10.7717/peerj.4652.

Egan, S.A., Corcoran, S., McDermott, H., Fitzpatrick, M., Hoyne, A., McCormack, O., Cullen, A., Brennan, G.I., O'Connell, B. and Coleman, D.C. (2020) 'Hospital outbreak of linezolid-resistant and vancomycin-resistant ST80 *Enterococcus faecium* harbouring an *optrA*-encoding conjugative plasmid investigated by whole-genome sequencing', *Journal of Hospital Infection*, 105(4), pp. 726–735. doi:10.1016/j.jhin.2020.05.013.

Egan, S.A., Kavanagh, N.L., Shore, A.C., Mollerup, S., Samaniego Castruita, J.A., O'Connell, B., McManus, B.A., Brennan, G.I., Pinholt, M., Westh, H. and Coleman, D.C. (2022) 'Genomic analysis of 600 vancomycin-resistant *Enterococcus faecium* reveals a high prevalence of ST80 and spread of similar *vanA* regions via IS1216E and plasmid transfer in diverse genetic lineages in Ireland', *Journal of Antimicrobial Chemotherapy*, 77(2), pp. 320–330. doi:10.1093/jac/dkab393.

Elbrecht, V., Vamos, E.E., Meissner, K., Aroviita, J. and Leese, F. (2017) 'Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring', *Methods in Ecology and Evolution*, 8(10), pp. 1265–1275. doi:10.1111/2041-210X.12789.

Ellington, M.J., Ekelund, O., Aarestrup, F.M., Canton, R., Doumith, M., Giske, C., Grundman, H., Hasman, H., Holden, M.T.G., Hopkins, K.L., Iredell, J., Kahlmeter, G., Köser, C.U., MacGowan, A., Mevius, D., Mulvey, M., Naas, T., Peto, T., Rolain, J.-M., Samuelsen, Ø. and Woodford, N. (2017) 'The role of whole genome sequencing in antimicrobial susceptibility testing of bacteria: report from the EUCAST Subcommittee', *Clinical Microbiology and Infection*, 23(1), pp. 2–22. doi:10.1016/j.cmi.2016.11.012.

van Elsas, J.D. and Bailey, M.J. (2002) 'The ecology of transfer of mobile genetic elements', *FEMS Microbiology Ecology*, 42(2), pp. 187–197. doi:10.1111/j.1574-6941.2002.tb01008.x.

Escolà-Vergé, L., Fernández-Hidalgo, N., Rodríguez-Pardo, D., Pigrau, C., González-López, J.J., Bartolomé, R. and Almirante, B. (2019) 'Teicoplanin for treating enterococcal infective endocarditis: A retrospective observational study from a referral centre in Spain', *International Journal of Antimicrobial Agents*, 53(2), pp. 165–170. doi:10.1016/j.ijantimicag.2018.10.003.

Falgenhauer, L., Fritzenwanker, M., Imirzalioglu, C., Steul, K., Scherer, M., Albert-Braun, S., Hunfeld, K.-P., Kempf, V., Kneifel, A., Kukic, A., Jahn-Mühl, B., Madlener, K., Oberdorfer, K., Oeltze, J.-P., Schulze, J., Sotoudeh, N., Tessmann, R., Heudorf, U., Chakraborty, T., and Rhine-Main VREfm study group (2019) 'Near-ubiquitous presence of a vancomycin-resistant *Enterococcus faecium* ST117/CT71/*vanB* –clone in the Rhine-Main metropolitan area of Germany', *Antimicrobial Resistance & Infection Control*, 8(1), p. 128. doi:10.1186/s13756-019-0573-8.

Fang, H., Fröding, I., Ullberg, M. and Giske, C.G. (2021) 'Genomic analysis revealed distinct transmission clusters of vancomycin-resistant *Enterococcus faecium* ST80 in Stockholm, Sweden', *Journal of Hospital Infection*, 107, pp. 12–15. doi:10.1016/j.jhin.2020.10.019.

FAO yearbook. Fishery and Aquaculture Statistics 2016 (2018).

Ferri, M., Ranucci, E., Romagnoli, P. and Giaccone, V. (2017) 'Antimicrobial resistance: A global emerging threat to public health systems', *Critical Reviews in Food Science and Nutrition*, 57(13), pp. 2857–2876. doi:10.1080/10408398.2015.1077192.

Fleming, A. (1929) 'On the Antibacterial Action of Cultures of a Penicillium, with Special Reference to their Use in the Isolation of *B. influenzae*', *British journal of experimental pathology*, 10(3), pp. 226–236.

Flores-Mireles, A.L., Walker, J.N., Caparon, M. and Hultgren, S.J. (2015) 'Urinary tract infections: epidemiology, mechanisms of infection and treatment options', *Nature reviews. Microbiology*, 13(5), pp. 269–284. doi:10.1038/nrmicro3432.

Forbes, J.D., Knox, N.C., Ronholm, J., Pagotto, F. and Reimer, A. (2017) 'Metagenomics: The Next Culture-Independent Game Changer', *Frontiers in Microbiology*, 8, pp. 1069. doi:10.3389/fmicb.2017.01069

Founou, R.C., Founou, L.L. and Essack, S.Y. (2017) 'Clinical and economic impact of antibiotic resistance in developing countries: A systematic review and meta-analysis', *PLOS ONE*, 12(12), p. e0189621. doi:10.1371/journal.pone.0189621.

Freitas, A.R., Tedim, A.P., Novais, C., Coque, T.M. and Peixe, L. (2018) 'Distribution of putative virulence markers in *Enterococcus faecium*: towards a safety profile review', *Journal of Antimicrobial Chemotherapy*, 73(2), pp. 306–319. doi:10.1093/jac/dkx387.

Fujimura, K.E., Johnson, C.C., Ownby, D.R., Cox, M.J., Brodie, E.L., Havstad, S.L., Zoratti, E.M., Woodcroft, K.J., Bobbitt, K.R., Wegienka, G., Boushey, H.A. and Lynch, S.V. (2010) 'Man's best friend? The effect of pet ownership on house dust microbial communities', *Journal of Allergy and Clinical Immunology*, 126(2), pp. 410-412.e3. doi:10.1016/j.jaci.2010.05.042.

Gagetti, P., Bonofiglio, L., García Gabarrot, G., Kaufman, S., Mollerach, M., Vigliarolo, L., von Specht, M., Toresani, I. and Lopardo, H.A. (2019) 'Resistance to β-lactams in enterococci', *Revista Argentina de Microbiología*, 51(2), pp. 179–183. doi:10.1016/j.ram.2018.01.007.

Gandra, S., Tseng, K.K., Arora, A., Bhowmik, B., Robinson, M.L., Panigrahi, B., Laxminarayan, R. and Klein, E.Y. (2019) 'The Mortality Burden of Multidrug-resistant Pathogens in India: A Retrospective, Observational Study', *Clinical Infectious Diseases*, 69(4), pp. 563–570. doi:10.1093/cid/ciy955.

Gao, W., Howden, B.P. and Stinear, T.P. (2018) 'Evolution of virulence in *Enterococcus faecium*, a hospital-adapted opportunistic pathogen', *Current Opinion in Microbiology*, 41, pp. 76–82. doi:10.1016/j.mib.2017.11.030.

Garlock, T., Asche, F., Anderson, J., Bjørndal, T., Kumar, G., Lorenzen, K., Ropicki, A., Smith, M.D. and Tveterås, R. (2020) 'A Global Blue Revolution: Aquaculture Growth Across Regions, Species, and Countries', *Reviews in Fisheries Science & Aquaculture*, 28(1), pp. 107–116. doi:10.1080/23308249.2019.1678111.

Gawronski, J.D., Wong, S.M.S., Giannoukos, G., Ward, D.V. and Akerley, B.J. (2009) 'Tracking insertion mutants within libraries by deep sequencing and a genome-wide screen for *Haemophilus* genes required in the lung', *Proceedings of the National Academy of Sciences*, 106(38), pp. 16422–16427. doi:10.1073/pnas.0906627106.

Gaynes, R. (2017) 'The Discovery of Penicillin—New Insights After More Than 75 Years of Clinical Use', *Emerging Infectious Diseases*, 23(5), pp. 849–853. doi:10.3201/eid2305.161556.

Giannakopoulos, X., Sakkas, H., Ragos, V., Tsiambas, E., Bozidis, P., Evangelou, A.M., Papadopoulou, C. and Petrogian, L. (2019) 'Impact of enterococcal urinary tract infections in immuno- compromised – neoplastic patients', *Journal of the Balkan Union of Oncology*, p. 8.

Gilmore, M.S., Salamzade, R., Selleck, E., Bryan, N., Mello, S.S., Manson, A.L. and Earl, A.M. (2020) 'Genes Contributing to the Unique Biology and Intrinsic Antibiotic Resistance of *Enterococcus faecalis*', *mBio*, 11(6), pp. e02962-20. doi:10.1128/mBio.02962-20.

Goodman, A.L., McNulty, N.P., Zhao, Y., Leip, D., Mitra, R.D., Lozupone, C.A., Knight, R. and Gordon, J.I. (2009) 'Identifying Genetic Determinants Needed to Establish a Human Gut Symbiont in Its Habitat', *Cell Host & Microbe*, 6(3), pp. 279–289. doi:10.1016/j.chom.2009.08.003.

Goodwin, S., McPherson, J.D. and McCombie, W.R. (2016) 'Coming of age: ten years of next-generation sequencing technologies', *Nature Reviews Genetics*, 17(6), pp. 333–351. doi:10.1038/nrg.2016.49.

Gould, K. (2016) 'Antibiotics: from prehistory to the present day', *Journal of Antimicrobial Chemotherapy*, 71(3), pp. 572–575. doi:10.1093/jac/dkv484.

Gouliouris, T., Coll, F., Ludden, C., Blane, B., Raven, K.E., Naydenova, P., Crawley, C., Török, M.E., Enoch, D.A., Brown, N.M., Harrison, E.M., Parkhill, J. and Peacock, S.J. (2021) 'Quantifying acquisition and transmission of *Enterococcus faecium* using genomic surveillance', *Nature microbiology*, 6(1), pp. 103–111. doi:10.1038/s41564-020-00806-7.

Gouliouris, T., Raven, K.E., Ludden, C., Blane, B., Corander, J., Horner, C.S., Hernandez-Garcia, J., Wood, P., Hadjirin, N.F., Radakovic, M., Holmes, M.A., de Goffau, M., Brown, N.M., Parkhill, J. and Peacock, S.J. (2018) 'Genomic Surveillance of *Enterococcus faecium* Reveals Limited Sharing of Strains and Resistance Genes between Livestock and Humans in the United Kingdom', *mBio*, 9(6), pp. e01780-18. doi:10.1128/mBio.01780-18.

Griffin, M.E., Espinosa, J., Becker, J.L., Luo, J.-D., Carroll, T.S., Jha, J.K., Fanger, G.R. and Hang, H.C. (2021) '*Enterococcus* peptidoglycan remodeling promotes checkpoint inhibitor cancer immunotherapy', *Science (New York, N.Y.)*, 373(6558), pp. 1040–1046. doi:10.1126/science.abc9113.

Guffey, A.A. and Loll, P.J. (2021) 'Regulation of Resistance in Vancomycin-Resistant Enterococci: The VanRS Two-Component System', *Microorganisms*, 9(10), p. 2026. doi:10.3390/microorganisms9102026.

Gupta, C.L., Tiwari, R.K. and Cytryn, E. (2020) 'Platforms for elucidating antibiotic resistance in single genomes and complex metagenomes', *Environment International*, 138, p. 105667. doi:10.1016/j.envint.2020.105667.

Gurevich, A., Saveliev, V., Vyahhi, N. and Tesler, G. (2013) 'QUAST: quality assessment tool for genome assemblies', *Bioinformatics*, 29(8), pp. 1072–1075. doi:10.1093/bioinformatics/btt086.

Gutierrez, M.G., Yoder-Himes, D.R. and Warawa, J.M. (2015) 'Comprehensive identification of virulence factors required for respiratory melioidosis using Tn-seq mutagenesis', *Frontiers in Cellular and Infection Microbiology*, 5(78). doi: 10.3389/fcimb.2015.00078

Guzman Prieto, A.M., van Schaik, W., Rogers, M.R.C., Coque, T.M., Baquero, F., Corander, J. and Willems, R.J.L. (2016) 'Global Emergence and Dissemination of Enterococci as Nosocomial Pathogens: Attack of the Clones?', *Frontiers in Microbiology*, 7, p. 788. doi:10.3389/fmicb.2016.00788.

Guzmán Prieto, A.M., Urbanus, R.T., Zhang, X., Bierschenk, D., Koekman, C.A., van Luit-Asbroek, M., Ouwerkerk, J.P., Pape, M., Paganelli, F.L., Wobser, D., Huebner, J., Hendrickx, A.P.A., Bonten, M.J.M., Willems, R.J.L. and van Schaik, W. (2015) 'The N-terminal domain of the thermo-regulated surface protein PrpA of *Enterococcus faecium* binds to fibrinogen, fibronectin and platelets', *Scientific Reports*, 5(1), p. 18255. doi:10.1038/srep18255.

Ha, D.R., Haste, N.M. and Gluckstein, D.P. (2017) 'The Role of Antibiotic Stewardship in Promoting Appropriate Antibiotic Use', *American Journal of Lifestyle Medicine*, 13(4), pp. 376–383. doi:10.1177/1559827617700824.

Habib, C., Mueller, A., Liu, M., Zhu, J., Schneider, T., Garner, E., Sun, Y. and Chai, Y. (2021) 'Bacterial "galactosemia" is caused by cytoplasmic interference of an essential cell wall biosynthesis glycosyltransferase'. bioRxiv, p. 2021.01.22.426804. doi:10.1101/2021.01.22.426804.

van Hal, S.J., Ip, C.L.C., Ansari, M.A., Wilson, D.J., Espedido, B.A., Jensen, S.O. and Bowden, R. 2016 (2016) 'Evolutionary dynamics of *Enterococcus faecium* reveals complex genomic relationships between isolates with independent emergence of vancomycin resistance', *Microbial Genomics*, 2(1), p. e000048. doi:10.1099/mgen.0.000048.

Hammerum, A.M. (2012) 'Enterococci of animal origin and their significance for public health', *Clinical Microbiology and Infection*, 18(7), pp. 619–625. doi:10.1111/j.1469-0691.2012.03829.x.

Han, M., Zhang, L., Zhang, N., Mao, Y., Peng, Z., Huang, B., Zhang, Y. and Wang, Z. (2022) 'Antibiotic resistome in a large urban-lake drinking water source in middle China: Dissemination mechanisms and risk assessment', *Journal of Hazardous Materials*, 424, p. 127745. doi:10.1016/j.jhazmat.2021.127745.

Hanchi, H., Mottawea, W., Sebei, K. and Hammami, R. (2018) 'The Genus *Enterococcus*: Between Probiotic Potential and Safety Concerns—An Update', *Frontiers in Microbiology*, 9. doi:10.3389/fmicb.2018.01791.

Hancock, L.E. and Gilmore, M.S. (2002) 'The capsular polysaccharide of *Enterococcus faecalis* and its relationship to other polysaccharides in the cell wall', *Proceedings of the National Academy of Sciences*, 99(3), pp. 1574–1579. doi:10.1073/pnas.032448299.

Handwerger, S., Skoble, J., Discotto, L.F. and Pucci, M.J. (1995) 'Heterogeneity of the *vanA* gene cluster in clinical isolates of enterococci from the northeastern United States.', *Antimicrobial Agents and Chemotherapy*, 39(2), pp. 362–368.

Hashemian, S.M.R., Farhadi, T. and Ganjparvar, M. (2018) 'Linezolid: a review of its properties, function, and use in critical care', *Drug Design, Development and Therapy*, 12, pp. 1759–1767. doi:10.2147/DDDT.S164515.

Hasman, H., Clausen, P.T.L.C., Kaya, H., Hansen, F., Knudsen, J.D., Wang, M., Holzknecht, B.J., Samulioniené, J., Røder, B.L., Frimodt-Møller, N., Lund, O. and Hammerum, A.M. (2019) 'LRE-Finder, a Web tool for detection of the 23S rRNA mutations and the *optrA*, *cfr*, *cfr(B)* and *poxtA* genes encoding linezolid resistance in enterococci from whole-genome sequences', *Journal of Antimicrobial Chemotherapy*, 74(6), pp. 1473–1476. doi:10.1093/jac/dkz092.

Hegstad, K., Mikalsen, T., Coque, T.M., Werner, G. and Sundsfjord, A. (2010) 'Mobile genetic elements and their contribution to the emergence of antimicrobial resistant *Enterococcus faecalis* and *Enterococcus faecium*', *Clinical Microbiology and Infection*, 16(6), pp. 541–554. doi:10.1111/j.1469-0691.2010.03226.x.

Hendrickson, R., Urbaniak, C., Minich, J.J., Aronson, H.S., Martino, C., Stepanauskas, R., Knight, R. and Venkateswaran, K. (2021) 'Clean room microbiome complexity impacts planetary protection bioburden', *Microbiome*, 9(1), p. 238. doi:10.1186/s40168-021-01159-x.

Hendrickx, A.P.A., Bonten, M.J.M., van Luit-Asbroek, M., Schapendonk, C.M.E., Kragten, A.H.M. and Willems, R.J.L.Y. 2008 (2008) 'Expression of two distinct types of pili by a hospital-acquired *Enterococcus faecium* isolate', *Microbiology*, 154(10), pp. 3212–3223. doi:10.1099/mic.0.2008/020891-0.

Hendrickx, A.P.A., Willems, R.J.L., Bonten, M.J.M. and van Schaik, W. (2009) 'LPxTG surface proteins of enterococci', *Trends in Microbiology*, 17(9), pp. 423–430. doi:10.1016/j.tim.2009.06.004.

Hiruy, A.M., Mohammed, J., Haileselassie, M.M., Acharya, K., Butte, G., Haile, A.T., Walsh, C. and Werner, D. (2022) 'Spatiotemporal variation in urban wastewater pollution impacts on river microbiomes and associated hazards in the Akaki catchment, Addis Ababa, Ethiopia', *Science of The Total Environment*, 826, p. 153912. doi:10.1016/j.scitotenv.2022.153912.

Hofmann, K., Bucher, P. and Kajava, A.V. (1998) 'A model of Cdc25 phosphatase catalytic domain and Cdk-interaction surface based on the presence of a rhodanese homology domain11', *Journal of Molecular Biology*, 282(1), pp. 195–208. doi:10.1006/jmbi.1998.1998.

Holland, T.L., Baddour, L.M., Bayer, A.S., Hoen, B., Miro, J.M. and Fowler, V.G. (2016) 'Infective endocarditis', *Nature reviews. Disease primers*, 2, p. 16059. doi:10.1038/nrdp.2016.59.

Hollenbeck, B.L. and Rice, L.B. (2012) 'Intrinsic and acquired resistance mechanisms in *enterococcus'*, *Virulence*, 3(5), pp. 421–569. doi:10.4161/viru.21282.

Holman, T.R., Wu, Z., Wanner, B.L. and Walsh, C.T. (1994) 'Identification of the DNA-binding site for the phosphorylated VanR protein required for vancomycin resistance in *Enterococcus faecium*', *Biochemistry*, 33(15), pp. 4625–4631. doi:10.1021/bi00181a024.

Homan, W.L., Tribe, D., Poznanski, S., Li, M., Hogg, G., Spalburg, E., van Embden, J.D.A. and Willems, R.J.L. (2002) 'Multilocus Sequence Typing Scheme for *Enterococcus faecium*', *Journal of Clinical Microbiology*, 40(6), pp. 1963–1971. doi:10.1128/JCM.40.6.1963-1971.2002.

Hooper, D.C. and Jacoby, G.A. (2015) 'Mechanisms of drug resistance: quinolone resistance', *Annals of the New York Academy of Sciences*, 1354(1), pp. 12–31. doi:10.1111/nyas.12830.

Hossain, M.N., Rahaman, A., Hasan, Md.J., Uddin, Md.M., Khatun, N. and Shamsuddin, S.Md. (2021) 'Comparative seasonal assessment of pollution and health risks associated with heavy metals in water, sediment and Fish of Buriganga and Turag River in Dhaka City, Bangladesh', *SN Applied Sciences*, 3(4), p. 509. doi:10.1007/s42452-021-04464-0.

Hsieh, Y.-C., Lee, W.-S., Ou, T.-Y. and Hsueh, P.-R. (2010) 'Clonal spread of CC17 vancomycin-resistant Enterococcus faecium with multilocus sequence type 78 (ST78) and a novel ST444 in Taiwan', *European Journal of Clinical Microbiology & Infectious Diseases: Official Publication of the European Society of Clinical Microbiology*, 29(1), pp. 25–30. doi:10.1007/s10096-009-0810-9.

Hu, T., Chitnis, N., Monos, D. and Dinh, A. (2021) 'Next-generation sequencing technologies: An overview', *Human Immunology*, 82(11), pp. 801–811. doi:10.1016/j.humimm.2021.02.012.

Huang, F., Pan, L., Song, M., Tian, C. and Gao, S. (2018) 'Microbiota assemblages of water, sediment, and intestine and their associations with environmental factors and shrimp physiological health', *Applied Microbiology and Biotechnology*, 102(19), pp. 8585–8598. doi:10.1007/s00253-018-9229-5.

Huang, W., Chen, X., Jiang, X. and Zheng, B. (2017) 'Characterization of sediment bacterial communities in plain lakes with different trophic statuses', *MicrobiologyOpen*, 6(5), p. e00503. doi:10.1002/mbo3.503.

Huggins, L.G., Koehler, A.V., Ng-Nguyen, D., Wilcox, S., Schunack, B., Inpankaew, T. and Traub, R.J. (2019) 'Assessment of a metabarcoding approach for the characterisation of vector-borne bacteria in canines from Bangkok, Thailand', *Parasites & Vectors*, 12(1), p. 394. doi:10.1186/s13071-019-3651-0.

Hughes, A., Ballard, S., Sullivan, S. and Marshall, C. (2019) 'An outbreak of *vanA* vancomycin-resistant *Enterococcus faecium* in a hospital with endemic *vanB* VRE', *Infection, Disease & Health*, 24(2), pp. 82–91. doi:10.1016/j.idh.2018.12.002.

Humbert, R. and Simoni, R.D. (1980) 'Genetic and biomedical studies demonstrating a second gene coding for asparagine synthetase in *Escherichia coli*', *Journal of Bacteriology*, 142(1), pp. 212–220. doi:10.1128/jb.142.1.212-220.1980.

Hussain, F., Clasen, T., Akter, S., Bawel, V., Luby, S.P., Leontsini, E., Unicomb, L., Barua, M.K., Thomas, B. and Winch, P.J. (2017) 'Advantages and limitations for users of double pit pour-flush latrines: a qualitative study in rural Bangladesh', *BMC Public Health*, 17(1), p. 515. doi:10.1186/s12889-017-4412-7.

Inkster, T., Coia, J., Meunier, D., Doumith, M., Martin, K., Pike, R., Imrie, L., Kane, H., Hay, M., Wiuff, C., Wilson, J., Deighan, C., Hopkins, K.L., Woodford, N. and Hill, R. (2017) 'First outbreak of colonization by linezolid- and glycopeptide-resistant *Enterococcus faecium* harbouring the *cfr* gene in a UK nephrology unit', *Journal of Hospital Infection*, 97(4), pp. 397–402. doi:10.1016/j.jhin.2017.07.003.

Jain, C., Rodriguez-R, L.M., Phillippy, A.M., Konstantinidis, K.T. and Aluru, S. (2018) 'High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries', *Nature Communications*, 9(1), p. 5114. doi:10.1038/s41467-018-07641-9.

Jaluria, P., Konstantopoulos, K., Betenbaugh, M. and Shiloach, J. (2007) 'A perspective on microarrays: current applications, pitfalls, and potential uses', *Microbial Cell Factories*, 6, p. 4. doi:10.1186/1475-2859-6-4.

Janowiak, B.E. and Griffith, O.W. (2005) 'Glutathione Synthesis in Streptococcus agalactiae: 'One protein accounts for γ-glutamylcysteine synthetase and glutathione synthetase activities', *Journal of Biological Chemistry*, 280(12), pp. 11829–11839. doi:10.1074/jbc.M414326200.

Jelsbak, L., Hartman, H., Schroll, C., Rosenkrantz, J.T., Lemire, S., Wallrodt, I., Thomsen, L.E., Poolman, M., Kilstrup, M., Jensen, P.R. and Olsen, J.E. (2014) 'Identification of Metabolic Pathways Essential for Fitness of *Salmonella Typhimurium* In Vivo', *PLOS ONE*, 9(7), p. e101869. doi:10.1371/journal.pone.0101869.

Jeon, Y.-S., Chun, J. and Kim, B.-S. (2013) 'Identification of household bacterial community and analysis of species shared with human microbiome', *Current Microbiology*, 67(5), pp. 557–563. doi:10.1007/s00284-013-0401-y.

Jolley, K.A., Bray, J.E. and Maiden, M.C.J. (2018) 'Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications', *Wellcome Open Research*, 3, p. 124. doi:10.12688/wellcomeopenres.14826.1.

Jolley, K.A. and Maiden, M.C. (2010) 'BIGSdb: Scalable analysis of bacterial genome variation at the population level', *BMC Bioinformatics*, 11(1), p. 595. doi:10.1186/1471-2105-11-595.

Jonas, O.G., Irwin, Al., Berthe, F.C.J., Le Gall, F.G. and Marquez, P.P. (2017) *Drug-resistant infections : a threat to our economic future (Vol. 2) : final report (English).*

Jozefíková, A., Valček, A., Šoltys, K., Nováková, E. and Bujdáková, H. (2022) 'Persistence and multi-ward dissemination of vancomycin-resistant *Enterococcus faecium* ST17 clone in hospital settings in Slovakia 2017–2020', *International Journal of Antimicrobial Agents*, 59(4), p. 106561. doi:10.1016/j.ijantimicag.2022.106561.

Käll, L., Krogh, A. and Sonnhammer, E.L.L. (2007) 'Advantages of combined transmembrane topology and signal peptide prediction—the Phobius web server', *Nucleic Acids Research*, 35(Web Server issue), pp. W429–W432. doi:10.1093/nar/gkm256.

Kaminski, J., Gibson, M.K., Franzosa, E.A., Segata, N., Dantas, G. and Huttenhower, C. (2015) 'High-Specificity Targeted Functional Profiling in Microbial Communities with ShortBRED', *PLOS Computational Biology*, 11(12), p. e1004557. doi:10.1371/journal.pcbi.1004557.

Karkman, A., Pärnänen, K. and Larsson, D.G.J. (2019) 'Fecal pollution can explain antibiotic resistance gene abundances in anthropogenically impacted environments', *Nature Communications*, 10(1), p. 80. doi:10.1038/s41467-018-07992-3.

Katz, L.S., Griswold, T., Morrison, S.S., Caravas, J.A., Zhang, S., Bakker, H.C. den, Deng, X. and Carleton, H.A. (2019) 'Mashtree: a rapid comparison of whole genome sequence files', *Journal of Open Source Software*, 4(44), p. 1762. doi:10.21105/joss.01762.

Kawai, Y., Marles-Wright, J., Cleverley, R.M., Emmins, R., Ishikawa, S., Kuwano, M., Heinz, N., Bui, N.K., Hoyland, C.N., Ogasawara, N., Lewis, R.J., Vollmer, W., Daniel, R.A. and Errington, J. (2011) 'A widespread family of bacterial cell wall assembly proteins', *The EMBO journal*, 30(24), pp. 4931–4941. doi:10.1038/emboj.2011.358.

Kawsar, A., Alam, T., Ahamed, S. and Mou, H. (2018) 'Aqua drugs and antibiotics used in freshwater aquaculture of North Chittagong, Bangladesh', *International Journal of Fisheries and Aquatic Studies*, 7(1), p. 7.

Khan, A., Davlieva, M., Panesso, D., Rincon, S., Miller, W.R., Diaz, L., Reyes, J., Cruz, M.R., Pemberton, O., Nguyen, A.H., Siegel, S.D., Planet, P.J., Narechania, A., Latorre, M., Rios, R., Singh, K.V., Ton-That, H., Garsin, D.A., Tran, T.T., Shamoo, Y. and Arias, C.A. (2019) 'Antimicrobial sensing coupled with cell membrane remodeling mediates antibiotic resistance and virulence in *Enterococcus faecalis*', *Proceedings of the National Academy of Sciences*, 116(52), pp. 26925–26932. doi:10.1073/pnas.1916037116.

Khan, A.U., Maryam, L. and Zarrilli, R. (2017) 'Structure, Genetics and Worldwide Spread of New Delhi Metallo-β-lactamase (NDM): a threat to public health', *BMC Microbiology*, 17(1), p. 101. doi:10.1186/s12866-017-1012-8.

Khan, M.A., Northwood, J.B., Loor, R.G.J., Tholen, A.T.R., Riera, E., Falcón, M., Network, P.A., van Belkum, A., van Westreenen, M. and Hays, J.P. (2010) 'High prevalence of ST-78 infection-associated vancomycin-resistant *Enterococcus faecium* from hospitals in Asunción, Paraguay', *Clinical Microbiology and Infection*, 16(6), pp. 624–627. doi:10.1111/j.1469-0691.2009.02898.x.

Kingsford, C.L., Ayanbule, K. and Salzberg, S.L. (2007) 'Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake', *Genome Biology*, 8(2), p. R22. doi:10.1186/gb-2007-8-2-r22.

Kohler, P., Eshaghi, A., Kim, H.C., Plevneshi, A., Green, K., Willey, B.M., McGeer, A., Patel, S.N., and Toronto Invasive Bacterial Diseases Network (TIBDN) (2018) 'Prevalence of vancomycin-variable *Enterococcus faecium* (VVE) among *vanA*-positive sterile site isolates and patient factors associated with VVE bacteremia', *PloS One*, 13(3), p. e0193926. doi:10.1371/journal.pone.0193926.

Köhler, T., Epp, S.F., Curty, L.K. and Pechère, J.C. (1999) 'Characterization of MexT, the regulator of the MexE-MexF-OprN multidrug efflux system of *Pseudomonas aeruginosa*', *Journal of Bacteriology*, 181(20), pp. 6300–6305.

Kohler, V., Vaishampayan, A. and Grohmann, E. (2018) 'Broad-host-range Inc18 plasmids: Occurrence, spread and transfer mechanisms', *Plasmid*, 99, pp. 11–21. doi:10.1016/j.plasmid.2018.06.001.

Koonin, E.V., Makarova, K.S. and Wolf, Y.I. (2021) 'Evolution of Microbial Genomics: Conceptual Shifts over a Quarter Century', *Trends in Microbiology*, 29(7), pp. 582–592. doi:10.1016/j.tim.2021.01.005.

de Kraker, M.E.A., Stewardson, A.J. and Harbarth, S. (2016) 'Will 10 Million People Die a Year due to Antimicrobial Resistance by 2050?', *PLoS Medicine*, 13(11), p. e1002184. doi:10.1371/journal.pmed.1002184.

Kramer, T.S., Remschmidt, C., Werner, S., Behnke, M., Schwab, F., Werner, G., Gastmeier, P. and Leistner, R. (2018) 'The importance of adjusting for *enterococcus* species when assessing the burden of vancomycin resistance: a cohort study including over 1000 cases of enterococcal bloodstream infections', *Antimicrobial Resistance & Infection Control*, 7(1), p. 133. doi:10.1186/s13756-018-0419-9.

Krawczyk, P.S., Lipinski, L. and Dziembowski, A. (2018) 'PlasFlow: predicting plasmid sequences in metagenomic data using genome signatures', *Nucleic Acids Research*, 46(6), pp. e35–e35. doi:10.1093/nar/gkx1321.

Ku, J.W.K. and Gan, Y.-H. (2021) 'New roles for glutathione: Modulators of bacterial virulence and pathogenesis', *Redox Biology*, 44, p. 102012. doi:10.1016/j.redox.2021.102012.

Kukurba, K.R. and Montgomery, S.B. (2015) 'RNA Sequencing and Analysis', *Cold Spring Harbor protocols*, 2015(11), pp. 951–969. doi:10.1101/pdb.top084970.

Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K. (2018) 'MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms', *Molecular Biology and Evolution*, 35(6), pp. 1547–1549. doi:10.1093/molbev/msy096.

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. and Raoult, D. (2012) 'Microbial culturomics: paradigm shift in the human gut microbiome study', *Clinical Microbiology and Infection*, 18(12), pp. 1185–1193. doi:10.1111/1469-0691.12023.

Lagier, J.-C., Dubourg, G., Million, M., Cadoret, F., Bilen, M., Fenollar, F., Levasseur, A., Rolain, J.-M., Fournier, P.-E. and Raoult, D. (2018) 'Culturing the human microbiota and culturomics', *Nature Reviews Microbiology*, 16(9), pp. 540–550. doi:10.1038/s41579-018-0041-0.

Lam, L.N., Wong, J.J., Chong, K.K.L. and Kline, K.A. (2020) '*Enterococcus faecalis* Manganese Exporter MntE Alleviates Manganese Toxicity and Is Required for Mouse Gastrointestinal Colonization', *Infection and Immunity*, 88(6), pp. e00058-20. doi:10.1128/IAI.00058-20.

Lam, M.M.C., Seemann, T., Bulach, D.M., Gladman, S.L., Chen, H., Haring, V., Moore, R.J., Ballard, S., Grayson, M.L., Johnson, P.D.R., Howden, B.P. and Stinear, T.P. (2012) 'Comparative Analysis of the First Complete *Enterococcus faecium* Genome', *Journal of Bacteriology*, 194(9), pp. 2334–2341. doi:10.1128/JB.00259-12.

Lamba, M., Gupta, S., Shukla, R., Graham, D.W., Sreekrishnan, T.R. and Ahammad, S.Z. (2018) 'Carbapenem resistance exposures via wastewaters across New Delhi', *Environment International*, 119, pp. 302–308. doi:10.1016/j.envint.2018.07.004.

Landete, J.M., Peirotén, Á., Medina, M., Arqués, J.L. and Rodríguez-Mínguez, E. (2018) 'Virulence and Antibiotic Resistance of Enterococci Isolated from Healthy Breastfed Infants', *Microbial Drug Resistance*, 24(1), pp. 63–69. doi:10.1089/mdr.2016.0320.

Langridge, G.C., Phan, M.-D., Turner, D.J., Perkins, T.T., Parts, L., Haase, J., Charles, I., Maskell, D.J., Peters, S.E., Dougan, G., Wain, J., Parkhill, J. and Turner, A.K. (2009) 'Simultaneous assay of every *Salmonella Typhi* gene using one million transposon mutants', *Genome Research*, 19(12), pp. 2308–2316. doi:10.1101/gr.097097.109.

Larivière, D., Wickham, L., Keiler, K., Nekrutenko, A., and The Galaxy Team (2021) 'Reproducible and accessible analysis of transposon insertion sequencing in Galaxy for qualitative essentiality analyses', *BMC Microbiology*, 21(1), p. 168. doi:10.1186/s12866-021-02184-4.

Law, C.W., Chen, Y., Shi, W. and Smyth, G.K. (2014) 'voom: precision weights unlock linear model analysis tools for RNA-seq read counts', *Genome Biology*, 15(2), p. R29. doi:10.1186/gb-2014-15-2-r29.

Lax, S., Smith, D.P., Hampton-Marcell, J., Owens, S.M., Handley, K.M., Scott, N.M., Gibbons, S.M., Larsen, P., Shogan, B.D., Weiss, S., Metcalf, J.L., Ursell, L.K., Vázquez-

Baeza, Y., Van Treuren, W., Hasan, N.A., Gibson, M.K., Colwell, R., Dantas, G., Knight, R. and Gilbert, J.A. (2014) 'Longitudinal analysis of microbial interaction between humans and the indoor environment', *Science (New York, N.Y.)*, 345(6200), pp. 1048–1052. doi:10.1126/science.1254529.

Laxminarayan, R., Duse, A., Wattal, C., Zaidi, A.K.M., Wertheim, H.F.L., Sumpradit, N., Vlieghe, E., Hara, G.L., Gould, I.M., Goossens, H., Greko, C., So, A.D., Bigdeli, M., Tomson, G., Woodhouse, W., Ombaka, E., Peralta, A.Q., Qamar, F.N., Mir, F., Kariuki, S., Bhutta, Z.A., Coates, A., Bergstrom, R., Wright, G.D., Brown, E.D. and Cars, O. (2013) 'Antibiotic resistance—the need for global solutions', *The Lancet Infectious Diseases*, 13(12), pp. 1057–1098. doi:10.1016/S1473-3099(13)70318-9.

Lebreton, F., Manson, A.L., Saavedra, J.T., Straub, T.J., Earl, A.M. and Gilmore, M.S. (2017) 'Tracing the Enterococci from Paleozoic Origins to the Hospital', *Cell*, 169(5), pp. 849-861.e13. doi:10.1016/j.cell.2017.04.027.

Lebreton, F., Schaik, W. van, McGuire, A.M., Godfrey, P., Griggs, A., Mazumdar, V., Corander, J., Cheng, L., Saif, S., Young, S., Zeng, Q., Wortman, J., Birren, B., Willems, R.J.L., Earl, A.M. and Gilmore, M.S. (2013) 'Emergence of Epidemic Multidrug-Resistant *Enterococcus faecium* from Animal and Commensal Strains', *mBio*, 4(4), pp. e00534-13. doi:10.1128/mBio.00534-13.

Lee, K., Kim, D.-W., Lee, D.-H., Kim, Y.-S., Bu, J.-H., Cha, J.-H., Thawng, C.N., Hwang, E.-M., Seong, H.J., Sul, W.J., Wellington, E.M.H., Quince, C. and Cha, C.-J. (2020) 'Mobile resistome of human gut and pathogen drives anthropogenic bloom of antibiotic resistance', *Microbiome*, 8(1), p. 2. doi:10.1186/s40168-019-0774-7.

Leenhouts, K., Buist, G., Bolhuis, A., ten Berge, A., Kiel, J., Mierau, I., Dabrowska, M., Venema, G. and Kok, J. (1996) 'A general system for generating unlabelled gene replacements in bacterial chromosomes', *Molecular and General Genetics MGG*, 253(1), pp. 217–224. doi:10.1007/s004380050315.

Lei, J., Zhou, Y.-F., Li, L.-F. and Su, X.-D. (2009) 'Structural and biochemical analyses of YvgN and YtbE from *Bacillus subtilis*', *Protein Science : A Publication of the Protein Society*, 18(8), pp. 1792–1800. doi:10.1002/pro.178.

Lemonidis, K., Salih, T.S., Dancer, S.J., Hunter, I.S. and Tucker, N.P. (2019) 'Emergence of an Australian-like *pstS*-null vancomycin resistant *Enterococcus faecium* clone in Scotland', *PLOS ONE*, 14(6), p. e0218185. doi:10.1371/journal.pone.0218185.

Li, D., Liu, C.-M., Luo, R., Sadakane, K. and Lam, T.-W. (2015) 'MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph', *Bioinformatics*, 31(10), pp. 1674–1676. doi:10.1093/bioinformatics/btv033.

Li, H. and Durbin, R. (2009) 'Fast and accurate short read alignment with Burrows–Wheeler transform', *Bioinformatics*, 25(14), pp. 1754–1760. doi:10.1093/bioinformatics/btp324.

Li, Q., Chang, W., Zhang, H., Hu, D. and Wang, X. (2019) 'The Role of Plasmids in the Multiple Antibiotic Resistance Transfer in ESBLs-Producing Escherichia coli Isolated From

Wastewater Treatment Plants', *Frontiers in Microbiology*, 10. doi:10.3389/fmicb.2019.00633.

Li, Y., Mima, T., Komori, Y., Morita, Y., Kuroda, T., Mizushima, T. and Tsuchiya, T. (2003) 'A new member of the tripartite multidrug efflux pumps, MexVW-OprM, in *Pseudomonas aeruginosa*', *The Journal of Antimicrobial Chemotherapy*, 52(4), pp. 572–575. doi:10.1093/jac/dkg390.

Lim, C., Takahashi, E., Hongsuwan, M., Wuthiekanun, V., Thamlikitkul, V., Hinjoy, S., Day, N.P., Peacock, S.J. and Limmathurotsakul, D. (2016) 'Epidemiology and burden of multidrug-resistant bacterial infection in a developing country', *eLife*. 5, p. e18082. doi:10.7554/eLife.18082.

Liu, L., Wang, Y., Che, Y., Chen, Y., Xia, Y., Luo, R., Cheng, S.H., Zheng, C. and Zhang, T. (2020) 'High-quality bacterial genomes of a partial-nitritation/anammox system by an iterative hybrid assembly method', *Microbiome*, 8(1), p. 155. doi:10.1186/s40168-020-00937-3.

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. and Shen, J. (2016) 'Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study', *The Lancet Infectious Diseases*, 16(2), pp. 161–168. doi:10.1016/S1473-3099(15)00424-7.

Logsdon, G.A., Vollger, M.R. and Eichler, E.E. (2020) 'Long-read human genome sequencing and its applications', *Nature Reviews Genetics*, 21(10), pp. 597–614. doi:10.1038/s41576-020-0236-x.

Long, K.S., Poehlsgaard, J., Kehrenberg, C., Schwarz, S. and Vester, B. (2006) 'The Cfr rRNA Methyltransferase Confers Resistance to Phenicols, Lincosamides, Oxazolidinones, Pleuromutilins, and Streptogramin A Antibiotics', *Antimicrobial Agents and Chemotherapy*, 50(7), pp. e00131-06. doi:10.1128/AAC.00131-06.

Lorenz, R., Bernhart, S.H., Höner zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P.F. and Hofacker, I.L. (2011) 'ViennaRNA Package 2.0', *Algorithms for Molecular Biology*, 6(1), p. 26. doi:10.1186/1748-7188-6-26.

Love, M.I., Huber, W. and Anders, S. (2014) 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2', *Genome Biology*, 15(12), p. 550. doi:10.1186/s13059-014-0550-8.

Lowder, S.K., Skoet, J. and Raney, T. (2016) 'The Number, Size, and Distribution of Farms, Smallholder Farms, and Family Farms Worldwide', *World Development*, 87, pp. 16–29. doi:10.1016/j.worlddev.2015.10.041.

Lowe, R., Shirley, N., Bleackley, M., Dolan, S. and Shafee, T. (2017) 'Transcriptomics technologies', *PLOS Computational Biology*, 13(5), p. e1005457. doi:10.1371/journal.pcbi.1005457.

Lu, J., Breitwieser, F.P., Thielen, P. and Salzberg, S.L. (2017) 'Bracken: estimating species abundance in metagenomics data', *PeerJ Computer Science*, 3, p. e104. doi:10.7717/peerj-cs.104.

Lulijwa, R., Rupia, E.J. and Alfaro, A.C. (2020) 'Antibiotic use in aquaculture, policies and regulation, health and environmental risks: a review of the top 15 major producers', *Reviews in Aquaculture*, 12(2), pp. 640–663. doi:10.1111/raq.12344.

de Maat, V. (2022) *Adaptations of hospital-acquired Enterococcus faecium*. UMC Utrecht.

de Maat, V., Arredondo-Alonso, S., Willems, R.J.L. and van Schaik, W. (2020) 'Conditionally essential genes for survival during starvation in *Enterococcus faecium* E745', *BMC Genomics*, 21(1), p. 568. doi:10.1186/s12864-020-06984-2.

Magi, G., Capretti, R., Paoletti, C., Pietrella, M., Ferrante, L., Biavasco, F., Varaldo, P.E. and Facinelli, B. (2003) 'Presence of a *vanA*-Carrying Pheromone Response Plasmid (pBRG1) in a Clinical Isolate of *Enterococcus faecium*', *Antimicrobial Agents and Chemotherapy*, 47(5), pp. 1571-1576. doi:10.1128/AAC.47.5.1571-1576.2003.

Mahnert, A., Moissl-Eichinger, C., Zojer, M., Bogumil, D., Mizrahi, I., Rattei, T., Martinez, J.L. and Berg, G. (2019) 'Man-made microbial resistances in built environments', *Nature Communications*, 10(1), p. 968. doi:10.1038/s41467-019-08864-0.

Makałowski, W., Gotea, V., Pande, A. and Makałowska, I. (2019) 'Transposable Elements: Classification, Identification, and Their Use As a Tool For Comparative Genomics', in Anisimova, M. (ed.) *Evolutionary Genomics: Statistical and Computational Methods*. New York, NY: Springer (Methods in Molecular Biology), pp. 177–207. doi:10.1007/978-1-4939-9074-0_6.

Mamun, K.Z., Shears, P. and Hart, C.A. (1993) 'The prevalence and genetics of resistance to commonly used antimicrobial agents in faecal Enterobacteriaceae from children in Bangladesh.', *Epidemiology and Infection*, 110(3), pp. 447–458.

Mantere, T., Kersten, S. and Hoischen, A. (2019) 'Long-Read Sequencing Emerging in Medical Genetics', *Frontiers in Genetics*, 10:426. doi:10.3389/fgene.2019.00426.

Marmen, S., Fadeev, E., Al Ashhab, A., Benet-Perelberg, A., Naor, A., Patil, H.J., Cytryn, E., Viner-Mozzini, Y., Sukenik, A., Lalzar, M. and Sher, D. (2021) 'Seasonal Dynamics Are the Major Driver of Microbial Diversity and Composition in Intensive Freshwater Aquaculture', *Frontiers in Microbiology*, 12: 679743. doi:10.3389/fmicb.2021.679743.

Martin, M. (2011) 'Cutadapt removes adapter sequences from high-throughput sequencing reads', *EMBnet.journal*, 17(1), pp. 10–12. doi:10.14806/ej.17.1.200.

Martínez, J.L. (2008) 'Antibiotics and Antibiotic Resistance Genes in Natural Environments', *Science*, 321(5887), pp. 365–367. doi:10.1126/science.1159483.

Masip, L., Veeravalli, K. and Georgiou, G. (2006) 'The Many Faces of Glutathione in Bacteria', Mary Ann Liebert, 8(5-6). doi:10.1089/ars.2006.8.753.

Matelska, D., Shabalin, I.G., Jabłońska, J., Domagalski, M.J., Kutner, J., Ginalski, K. and Minor, W. (2018) 'Classification, substrate specificity and structural features of D-2-hydroxyacid dehydrogenases: 2HADH knowledgebase', *BMC Evolutionary Biology*, 18(1), p. 199. doi:10.1186/s12862-018-1309-8.

Matos, R.C., Lapaque, N., Rigottier-Gois, L., Debarbieux, L., Meylheuc, T., Gonzalez-Zorn, B., Repoila, F., Lopes, M. de F. and Serror, P. (2013) '*Enterococcus faecalis* Prophage Dynamics and Contributions to Pathogenic Traits', *PLOS Genetics*, 9(6), p. e1003539. doi:10.1371/journal.pgen.1003539.

McCombie, W.R., McPherson, J.D. and Mardis, E.R. (2019) 'Next-Generation Sequencing Technologies', *Cold Spring Harbor Perspectives in Medicine*, 9(11), p. a036798. doi:10.1101/cshperspect.a036798.

McEwen, S.A. and Collignon, P.J. (2018) 'Antimicrobial Resistance: a One Health Perspective', *Microbiology Spectrum*, 6(2). doi:10.1128/microbiolspec.ARBA-0009-2017.

McGhee, J.J., Rawson, N., Bailey, B.A., Fernandez-Guerra, A., Sisk-Hackworth, L. and Kelley, S.T. (2020) 'Meta-SourceTracker: application of Bayesian source tracking to shotgun metagenomics', *PeerJ*, 8. doi:10.7717/peerj.8783.

McInnes, R.S., McCallum, G.E., Lamberte, L.E. and van Schaik, W. (2020) 'Horizontal transfer of antibiotic resistance genes in the human gut microbiome', *Current Opinion in Microbiology*, 53, pp. 35–43. doi:10.1016/j.mib.2020.02.002.

McMurry, L.M., Park, B.H., Burdett, V. and Levy, S.B. (1987) 'Energy-dependent efflux mediated by class L (*tetL*) tetracycline resistance determinant from streptococci', *Antimicrobial Agents and Chemotherapy*, 31(10), pp. 1648–1650. doi:10.1128/AAC.31.10.1648.

Meziti, A., Rodriguez-R, L.M., Hatt, J.K., Peña-Gonzalez, A., Levy, K. and Konstantinidis, K.T. (2021) 'The Reliability of Metagenome-Assembled Genomes (MAGs) in Representing Natural Populations: Insights from Comparing MAGs against Isolate Genomes Derived from the Same Fecal Sample', *Applied and Environmental Microbiology,* 87(6), pp.e02593-20. doi:10.1128/AEM.02593-20.

Michael, C.A., Dominey-Howes, D. and Labbate, M. (2014) 'The Antimicrobial Resistance Crisis: Causes, Consequences, and Management', *Frontiers in Public Health*, 2(145). doi:10.3389/fpubh.2014.00145.

Millett, J. (1937) 'The Intrathecal Use of Prontosil Soluble: Report of a Case of Type III *Pneumococcus* Meningitis and Septicemia Treated with Prontosil Soluble, With Complete Autopsy Report', *Journal of the American Medical Association*, 109(26), pp. 2138–2139. doi:10.1001/jama.1937.02780520028007.

Modi, A., Vai, S., Caramelli, D. and Lari, M. (2021) 'The Illumina Sequencing Protocol and the NovaSeq 6000 System', in Mengoni, A., Bacci, G., and Fondi, M. (eds) *Bacterial Pangenomics: Methods and Protocols*. New York, NY: Springer US (Methods in Molecular Biology), pp. 15–42. doi:10.1007/978-1-0716-1099-2_2.

Moet, G.J., Jones, R.N., Biedenbach, D.J., Stilwell, M.G. and Fritsche, T.R. (2007) 'Contemporary causes of skin and soft tissue infections in North America, Latin America, and Europe: Report from the SENTRY Antimicrobial Surveillance Program (1998–2004)', *Diagnostic Microbiology and Infectious Disease*, 57(1), pp. 7–13. doi:10.1016/j.diagmicrobio.2006.05.009.

Monira, S., Shabnam, S.A., Ali, Sk.I., Sadique, A., Johura, F.-T., Rahman, K.Z., Alam, N.H., Watanabe, H. and Alam, M. (2017) 'Multi-drug resistant pathogenic bacteria in the gut of young children in Bangladesh', *Gut Pathogens*, 9(1), p. 19. doi:10.1186/s13099-017-0170-4.

Montravers, P., Mohler, J., Julien, L.S. and Carbon, C. (1997) 'Evidence of the proinflammatory role of *Enterococcus faecalis* in polymicrobial peritonitis in rats', *Infection and Immunity*, 65(1), pp. 144-149. doi:10.1128/iai.65.1.144-149.1997.

Munita, J.M. and Arias, C.A. (2016) 'Mechanisms of Antibiotic Resistance', *Microbiology Spectrum*, 4(2). doi:10.1128/microbiolspec.VMBF-0016-2015.

Murray, C.J., Ikuta, K.S., Sharara, F., Swetschinski, L., Aguilar, G.R., Gray, A., Han, C., Bisignano, C., Rao, P., Wool, E., Johnson, S.C., Browne, A.J., Chipeta, M.G., Fell, F., Hackett, S., Haines-Woodhouse, G., Hamadani, B.H.K., Kumaran, E.A.P., McManigal, B., Agarwal, R., Akech, S., Albertson, S., Amuasi, J., Andrews, J., Aravkin, A., Ashley, E., Bailey, F., Baker, S., Basnyat, B., Bekker, A., Bender, R., Bethou, A., Bielicki, J., Boonkasidecha, S., Bukosia, J., Carvalheiro, C., Castañeda-Orjuela, C., Chansamouth, V., Chaurasia, S., Chiurchiù, S., Chowdhury, F., Cook, A.J., Cooper, B., Cressey, T.R., Criollo-Mora, E., Cunningham, M., Darboe, S., Day, N.P.J., Luca, M.D., Dokova, K., Dramowski, A., Dunachie, S.J., Eckmanns, T., Eibach, D., Emami, A., Feasey, N., Fisher-Pearson, N., Forrest, K., Garrett, D., Gastmeier, P., Giref, A.Z., Greer, R.C., Gupta, V., Haller, S., Haselbeck, A., Hay, S.I., Holm, M., Hopkins, S., Iregbu, K.C., Jacobs, J., Jarovsky, D., Javanmardi, F., Khorana, M., Kissoon, N., Kobeissi, E., Kostyanev, T., Krapp, F., Krumkamp, R., Kumar, A., Kyu, H.H., Lim, C., Limmathurotsakul, D., Loftus, M.J., Lunn, M., Ma, J., Mturi, N., Munera-Huertas, T., Musicha, P., Mussi-Pinhata, M.M., Nakamura, T., Nanavati, R., Nangia, S., Newton, P., Ngoun, C., Novotney, A., Nwakanma, D., Obiero, C.W., Olivas-Martinez, A., Olliaro, P., Ooko, E., Ortiz-Brizuela, E., Peleg, A.Y., Perrone, C., Plakkal, N., Ponce-de-Leon, A., Raad, M., Ramdin, T., Riddell, A., Roberts, T., Robotham, J.V., Roca, A., Rudd, K.E., Russell, N., Schnall, J., Scott, J.A.G., Shivamallappa, M., Sifuentes-Osornio, J., Steenkeste, N., Stewardson, A.J., Stoeva, T., Tasak, N., Thaiprakong, A., Thwaites, G., Turner, C., Turner, P., Doorn, H.R. van, Velaphi, S., Vongpradith, A., Vu, H., Walsh, T., Waner, S., Wangrangsimakul, T., Wozniak, T., Zheng, P., Sartorius, B., Lopez, A.D., Stergachis, A., Moore, C., Dolecek, C. and Naghavi, M. (2022) 'Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis', *The Lancet*, 399(10325), pp. 629-655. doi:10.1016/S0140-6736(21)02724-0.

Muziasari, W.I., Pitkänen, L.K., Sørum, H., Stedtfeld, R.D., Tiedje, J.M. and Virta, M. (2017) 'The Resistome of Farmed Fish Feces Contributes to the Enrichment of Antibiotic Resistance Genes in Sediments below Baltic Sea Fish Farms', *Frontiers in Microbiology*, 7. doi:10.3389/fmicb.2016.02137.

Nanda, A.M., Thormann, K. and Frunzke, J. (2015) 'Impact of Spontaneous Prophage Induction on the Fitness of Bacterial Populations and Host-Microbe Interactions', *Journal of Bacteriology*, 197(3), pp. 410–419. doi:10.1128/JB.02230-14.

Narayanan, N., Rai, R., Vaidya, P., Desai, A., Bhowmick, T. and Weinstein, M.P. (2019) 'Comparison of linezolid and daptomycin for the treatment of vancomycin-resistant enterococcal bacteremia', *Therapeutic Advances in Infectious Disease*, 6, p. 2049936119828964. doi:10.1177/2049936119828964.

Nho, S.W., Abdelhamed, H., Paul, D., Park, S., Mauel, M.J., Karsi, A. and Lawrence, M.L. (2018) 'Taxonomic and Functional Metagenomic Profile of Sediment From a Commercial Catfish Pond in Mississippi', *Frontiers in Microbiology*, 9. doi:10.3389/fmicb.2018.02855.

Nicolaou, K.C. and Rigol, S. (2018) 'A brief history of antibiotics and select advances in their synthesis', *The Journal of Antibiotics*, 71(2), pp. 153–184. doi:10.1038/ja.2017.62.

Niebel, M., Perera, M.T.P.R., Shah, T., Marudanayagam, R., Martin, K., Oppenheim, B.A. and David, M.D. (2016) 'Emergence of linezolid resistance in hepatobiliary infections caused by *Enterococcus faecium*', *Liver Transplantation*, 22(2), pp. 201–208. doi:10.1002/lt.24328.

Nji, E., Kazibwe, J., Hambridge, T., Joko, C.A., Larbi, A.A., Damptey, L.A.O., Nkansa-Gyamfi, N.A., Stålsby Lundborg, C. and Lien, L.T.Q. (2021) 'High prevalence of antibiotic resistance in commensal *Escherichia coli* from healthy human sources in community settings', *Scientific Reports*, 11(1), p. 3372. doi:10.1038/s41598-021-82693-4.

O'Driscoll, T. and Crank, C.W. (2015) 'Vancomycin-resistant enterococcal infections: epidemiology, clinical manifestations, and optimal management', *Infection and Drug Resistance*, 8, pp. 217–230. doi:10.2147/IDR.S54125.

Oh, J.-H. and van Pijkeren, J.-P. (2014) 'CRISPR–Cas9-assisted recombineering in *Lactobacillus reuteri*', *Nucleic Acids Research*, 42(17), p. e131. doi:10.1093/nar/gku623.

Oksanen, J., Guillaume, B., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Henry, M., Stevens, H., Szoecs, E. and Wagner, Helene (2019) 'Vegan: Community Ecology Package. R'. Available at: https://CRAN.R-project.org/package=vegan.

O'Neill, J. (2014) *Antimicrobial resistance: tackling a crisis for the health and wealth of nations*. Available at: http://amr-review.org/Publications.

van Opijnen, T., Bodi, K.L. and Camilli, A. (2009) 'Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms', *Nature Methods*, 6(10), pp. 767–772. doi:10.1038/nmeth.1377.

van Opijnen, T. and Camilli, A. (2013) 'Transposon insertion sequencing: a new tool for systems-level analysis of microorganisms', *Nature Reviews Microbiology*, 11(7), pp. 435–442. doi:10.1038/nrmicro3033.

Opijnen, T. van, Dedrick, S. and Bento, J. (2016) 'Strain Dependent Genetic Networks for Antibiotic-Sensitivity in a Bacterial Pathogen with a Large Pan-Genome', *PLOS Pathogens*, 12(9), p. e1005869. doi:10.1371/journal.ppat.1005869.

Ouattara, N.K., de Brauwere, A., Billen, G. and Servais, P. (2013) 'Modelling faecal contamination in the Scheldt drainage network', *Journal of Marine Systems*, 128, pp. 77–88. doi:10.1016/j.jmarsys.2012.05.004.

Paerl, H.W. and Tucker, C.S. (1995) 'Ecology of Blue-Green Algae in Aquaculture Ponds', *Journal of the World Aquaculture Society*, 26(2), pp. 109–131. doi:10.1111/j.1749-7345.1995.tb00235.x.

Paganelli, F.L., de Been, M., Braat, J.C., Hoogenboezem, T., Vink, C., Bayjanov, J., Rogers, M.R.C., Huebner, J., Bonten, M.J.M., Willems, R.J.L. and Leavis, H.L. (2015) 'Distinct SagA from Hospital-Associated Clade A1 *Enterococcus faecium* Strains Contributes to Biofilm Formation', *Applied and Environmental Microbiology*, 81(19), pp. 6873–6882. doi:10.1128/AEM.01716-15.

Paganelli, F.L., Huebner, J., Singh, K.V., Zhang, X., van Schaik, W., Wobser, D., Braat, J.C., Murray, B.E., Bonten, M.J.M., Willems, R.J.L. and Leavis, H.L. (2016) 'Genome-wide Screening Identifies Phosphotransferase System Permease BepA to Be Involved in *Enterococcus faecium* Endocarditis and Biofilm Formation', *The Journal of Infectious Diseases*, 214(2), pp. 189–195. doi:10.1093/infdis/jiw108.

Palencia-Gándara, C., Getino, M., Moyano, G., Redondo, S., Fernández-López, R., González-Zorn, B. and de la Cruz, F. (2021) 'Conjugation Inhibitors Effectively Prevent Plasmid Transmission in Natural Environments', *mBio*, 12(4), p. e0127721. doi:10.1128/mBio.01277-21.

Palmer, K.L., Kos, V.N. and Gilmore, M.S. (2010) 'Horizontal Gene Transfer and the Genomics of Enterococcal Antibiotic Resistance', *Current opinion in microbiology*, 13(5), pp. 632–639. doi:10.1016/j.mib.2010.08.004.

Pandya, S., Ravi, K., Srinivas, V., Jadhav, S., Khan, A., Arun, A., Riley, L.W. and Madhivanan, P. (2017) 'Comparison of culture-dependent and culture-independent molecular methods for characterization of vaginal microflora', *Journal of Medical Microbiology*, 66(2), pp. 149–153. doi:10.1099/jmm.0.000407.

Park, S.C., Kwak, Y.M., Song, W.S., Hong, M. and Yoon, S. (2017) 'Structural basis of effector and operator recognition by the phenolic acid-responsive transcriptional regulator PadR', *Nucleic Acids Research*, 45(22), pp. 13080–13093. doi:10.1093/nar/gkx1055.

Parte, A.C., Sardà Carbasse, J., Meier-Kolthoff, J.P., Reimer, L.C. and Göker, M. 2020 (2020) 'List of Prokaryotic names with Standing in Nomenclature (LPSN) moves to the DSMZ', *International Journal of Systematic and Evolutionary Microbiology*, 70(11), pp. 5607–5612. doi:10.1099/ijsem.0.004332.

Partridge, S.R., Kwong, S.M., Firth, N. and Jensen, S.O. (2018) 'Mobile Genetic Elements Associated with Antimicrobial Resistance', *Clinical Microbiology Reviews*, 31(4), pp. e00088-17. doi:10.1128/CMR.00088-17.

Payne, A., Holmes, N., Rakyan, V. and Loose, M. (2019) 'BulkVis: a graphical viewer for Oxford nanopore bulk FAST5 files', *Bioinformatics*, 35(13), pp. 2193–2198. doi:10.1093/bioinformatics/bty841.

Peal, A., Evans, B., Blackett, I., Hawkins, P. and Heymans, C. (2014) 'Fecal sludge management: a comparative analysis of 12 cities', *Journal of Water, Sanitation and Hygiene for Development*, 4(4), pp. 563–575. doi:10.2166/washdev.2014.026.

Pedicord, V.A., Lockhart, A.A.K., Rangan, K.J., Craig, J.W., Loschko, J., Rogoz, A., Hang, H.C. and Mucida, D. (2016) 'Exploiting a host-commensal interaction to promote intestinal barrier function and enteric pathogen tolerance', *Science Immunology*, 1(3), p. eaai7732. doi:10.1126/sciimmunol.aai7732.

Pericàs, J.M., Llopis, J., Muñoz, P., Gálvez-Acebal, J., Kestler, M., Valerio, M., Hernández-Meneses, M., Goenaga, M.Á., Cobo-Belaustegui, M., Montejo, M., Ojeda-Burgos, G., Sousa-Regueiro, M.D., de Alarcón, A., Ramos-Martínez, A., Miró, J.M., and GAMES Investigators (2020) 'A Contemporary Picture of Enterococcal Endocarditis', *Journal of the American College of Cardiology*, 75(5), pp. 482–494. doi:10.1016/j.jacc.2019.11.047.

Peschel, A., Vuong, C., Otto, M. and Götz, F. (2000) 'The d-Alanine Residues of *Staphylococcus aureus* Teichoic Acids Alter the Susceptibility to Vancomycin and the Activity of Autolytic Enzymes', *Antimicrobial Agents and Chemotherapy*, 44(10), pp. 2845–2847.

Petersohn, A., Brigulla, M., Haas, S., Hoheisel, J.D., Völker, U. and Hecker, M. (2001) 'Global Analysis of the General Stress Response of *Bacillus subtilis*', *Journal of Bacteriology*, 183(19), pp. 5617–5631. doi:10.1128/JB.183.19.5617-5631.2001.

Peterson, E. and Kaur, P. (2018) 'Antibiotic Resistance Mechanisms in Bacteria: Relationships Between Resistance Determinants of Antibiotic Producers, Environmental Bacteria, and Clinical Pathogens', *Frontiers in Microbiology*, 9:2928, doi:10.3389/fmicb.2018.02928.

Poole, K. (2012) 'Bacterial stress responses as determinants of antimicrobial resistance', *Journal of Antimicrobial Chemotherapy*, 67(9), pp. 2069–2089. doi:10.1093/jac/dks196.

Potter, S.C., Luciani, A., Eddy, S.R., Park, Y., Lopez, R. and Finn, R.D. (2018) 'HMMER web server: 2018 update', *Nucleic Acids Research*, 46(W1), pp. W200–W204. doi:10.1093/nar/gky448.

Pratama, R., Beukers, A.G., McIver, C.J., Keighley, C.L., Taylor, P.C. and van Hal, S.J. (2021) 'A *vanA* vancomycin-resistant *Enterococcus faecium* ST80 outbreak resulting from a single importation event', *The Journal of Antimicrobial Chemotherapy*, 77(1), pp. 31–37. doi:10.1093/jac/dkab379.

Prestinaci, F., Pezzotti, P. and Pantosti, A. (2015) 'Antimicrobial resistance: a global multifaceted phenomenon', *Pathogens and Global Health*, 109(7), pp. 309–318. doi:10.1179/2047773215Y.0000000030.

Public Health England (2019) *Laboratory surveillance of Enterococcus spp. bacteraemia in England, Wales and Northern Ireland: 2018*. Health Protection Report.

Qian, G., Liu, C., Wu, G., Yin, F., Zhao, Y., Zhou, Y., Zhang, Y., Song, Z., Fan, J., Hu, B. and Liu, F. (2013) 'AsnB, regulated by diffusible signal factor and global regulator Clp, is involved in aspartate metabolism, resistance to oxidative stress and virulence in *Xanthomonas oryzae* pv. *oryzicola*', *Molecular Plant Pathology*, 14(2), pp. 145–157. doi:10.1111/j.1364-3703.2012.00838.x.

Qin, H., Lo, N.W.-S., Loo, J.F.-C., Lin, X., Yim, A.K.-Y., Tsui, S.K.-W., Lau, T.C.-K., Ip, M. and Chan, T.-F. (2018) 'Comparative transcriptomics of multidrug-resistant *Acinetobacter baumannii* in response to antibiotic treatments', *Scientific Reports*, 8(1), p. 3515. doi:10.1038/s41598-018-21841-9.

Quince, C., Walker, A.W., Simpson, J.T., Loman, N.J. and Segata, N. (2017) 'Shotgun metagenomics, from sampling to analysis', *Nature Biotechnology*, 35(9), pp. 833–844. doi:10.1038/nbt.3935.

Quinn, R. (2013) 'Rethinking Antibiotic Research and Development: World War II and the Penicillin Collaborative', *American Journal of Public Health*, 103(3), pp. 426–434. doi:10.2105/AJPH.2012.300693.

R Core Team (2017) 'R: A language and environment for statistical computing.', *R Foundation for Statistical Computing, Vienna, Austria*.

Ramos, Y., Sansone, S. and Morales, D.K. (2021) 'Sugarcoating it: Enterococcal polysaccharides as key modulators of host–pathogen interactions', *PLOS Pathogens*, 17(9), p. e1009822. doi:10.1371/journal.ppat.1009822.

Rangan, K.J., Pedicord, V.A., Wang, Y.-C., Kim, B., Lu, Y., Shaham, S., Mucida, D. and Hang, H.C. (2016) 'A secreted bacterial peptidoglycan hydrolase enhances tolerance to enteric pathogens', *Science (New York, N.Y.)*, 353(6306), pp. 1434–1437. doi:10.1126/science.aaf3552.

Raskin, D.M., Seshadri, R., Pukatzki, S.U. and Mekalanos, J.J. (2006) 'Bacterial Genomics and Pathogen Evolution', *Cell*, 124(4), pp. 703–714. doi:10.1016/j.cell.2006.02.002.

Rathnayake, I.U., Hargreaves, M. and Huygens, F. (2012) 'Antibiotic resistance and virulence traits in clinical and environmental *Enterococcus faecalis* and *Enterococcus faecium* isolates', *Systematic and Applied Microbiology*, 35(5), pp. 326–333. doi:10.1016/j.syapm.2012.05.004.

Raven, K.E., Gouliouris, T., Parkhill, J. and Peacock, S.J. (2018) 'Genome-Based Analysis of *Enterococcus faecium* Bacteremia Associated with Recurrent and Mixed-Strain Infection', *Journal of Clinical Microbiology*, 56(3), pp. e01520-17. doi:10.1128/JCM.01520-17.

Raven, K.E., Reuter, S., Reynolds, R., Brodrick, H.J., Russell, J.E., Török, M.E., Parkhill, J. and Peacock, S.J. (2016) 'A decade of genomic history for healthcare-associated *Enterococcus faecium* in the United Kingdom and Ireland', *Genome Research*, 26(10), pp. 1388–1396. doi:10.1101/gr.204024.116.

Ren, H. and Liu, J. (2006) 'AsnB is involved in natural resistance of *Mycobacterium smegmatis* to multiple drugs', *Antimicrobial Agents and Chemotherapy*, 50(1), pp. 250–255. doi:10.1128/AAC.50.1.250-255.2006.

Reyes, K., Bardossy, A.C. and Zervos, M. (2016) 'Vancomycin-Resistant Enterococci: Epidemiology, Infection Prevention, and Control', *Infectious Disease Clinics of North America*, 30(4), pp. 953–965. doi:10.1016/j.idc.2016.07.009.

Reynolds, E., Ross, J.I. and Cove, J.H. (2003) 'Msr(A) and related macrolide/streptogramin resistance determinants: incomplete transporters?', *International Journal of Antimicrobial Agents*, 22(3), pp. 228–236. doi:10.1016/S0924-8579(03)00218-8.

Reynolds, L.J., Sala-Comorera, L., Martin, N.A., Nolan, T.M., Stephens, J.H., Gitto, A., O'Hare, G.M.P., O'Sullivan, J.J. and Meijer, W.G. (2020) 'Correlation between antimicrobial resistance and faecal contamination in small urban streams and bathing waters', *Science of The Total Environment*, 739, p. 140242. doi:10.1016/j.scitotenv.2020.140242.

Rios, R., Reyes, J., Carvajal, L.P., Rincon, S., Panesso, D., Echeverri, A.M., Dinh, A., Kolokotronis, S.-O., Narechania, A., Tran, T.T., Munita, J.M., Murray, B.E., Planet, P.J., Arias, C.A. and Diaz, L. (2020) 'Genomic Epidemiology of Vancomycin-Resistant *Enterococcus faecium* (VREfm) in Latin America: Revisiting The Global VRE Population Structure', *Scientific Reports*, 10(1), p. 5636. doi:10.1038/s41598-020-62371-7.

Robinson, T.P., Bu, D.P., Carrique-Mas, J., Fèvre, 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. and Woolhouse, M.E.J. (2016) 'Antibiotic resistance is the quintessential One Health issue', *Transactions of The Royal Society of Tropical Medicine and Hygiene*, 110(7), pp. 377–380. doi:10.1093/trstmh/trw048.

Roper, D.I., Huyton, T., Vagin, A. and Dodson, G. (2000) 'The molecular basis of vancomycin resistance in clinically relevant Enterococci: Crystal structure of d-alanyl-d-lactate ligase (VanA)', *Proceedings of the National Academy of Sciences*, 97(16), pp. 8921–8925. doi:10.1073/pnas.150116497.

Rossi, F., Diaz, L., Wollam, A., Panesso, D., Zhou, Y., Rincon, S., Narechania, A., Xing, G., Di Gioia, T.S.R., Doi, A., Tran, T.T., Reyes, J., Munita, J.M., Carvajal, L.P., Hernandez-Roldan, A., Brandão, D., van der Heijden, I.M., Murray, B.E., Planet, P.J., Weinstock, G.M. and Arias, C.A. (2014) 'Transferable Vancomycin Resistance in a Community-Associated MRSA Lineage', *New England Journal of Medicine*, 370(16), pp. 1524–1531. doi:10.1056/NEJMoa1303359.

Rousham, E.K., Islam, M.A., Nahar, P., Lucas, P.J., Naher, N., Ahmed, S.M., Nizame, F.A. and Unicomb, L. (2019) 'Pathways of antibiotic use in Bangladesh: qualitative protocol for the PAUSE study', *BMJ Open*, 9(1), p. e028215. doi:10.1136/bmjopen-2018-028215.

Rubin, I.M.C., Pedersen, M.S., Mollerup, S., Kaya, H., Petersen, A.M., Westh, H. and Pinholt, M. (2020) 'Association between vancomycin-resistant *Enterococcus faecium* colonization and subsequent infection: a retrospective WGS study', *Journal of Antimicrobial Chemotherapy*, 75(7), pp. 1712–1715. doi:10.1093/jac/dkaa074.

Rubinstein, E. and Keynan, Y. (2014) 'Vancomycin Revisited – 60 Years Later', *Frontiers in Public Health*, 2(217). doi:10.3389/fpubh.2014.00217.

Ruiz-Masó, J.Á., Luengo, L.M., Moreno-Córdoba, I., Díaz-Orejas, R. and del Solar, G. (2017) 'Successful Establishment of Plasmids R1 and pMV158 in a New Host Requires the Relief of the Transcriptional Repression of Their Essential rep Genes', *Frontiers in Microbiology*, 8(2367). doi:10.3389/fmicb.2017.02367.

Santajit, S. and Indrawattana, N. (2016) 'Mechanisms of Antimicrobial Resistance in ESKAPE Pathogens', *BioMed Research International*, 2016, p. 2475067. doi:10.1155/2016/2475067.

Santiago, M., Lee, W., Fayad, A.A., Coe, K.A., Rajagopal, M., Do, T., Hennessen, F., Srisuknimit, V., Müller, R., Meredith, T.C. and Walker, S. (2018) 'Genome-wide mutant profiling predicts the mechanism of a Lipid II binding antibiotic', *Nature Chemical Biology*, 14(6), pp. 601–608. doi:10.1038/s41589-018-0041-4.

Sarker Md, J. (2016) 'A Study on the Determination of Heavy Metals in Sediment of Fish Farms in Bangladesh', *Fisheries and Aquaculture Journal*, 07(01). doi:10.4172/2150-3508.1000159.

Sattari-Maraji, A., Jabalameli, F., Node Farahani, N., Beigverdi, R. and Emaneini, M. (2019) 'Antimicrobial resistance pattern, virulence determinants and molecular analysis of *Enterococcus faecium* isolated from children infections in Iran', *BMC Microbiology*, 19(1), p. 156. doi:10.1186/s12866-019-1539-y.

van Schaik, W. (2015) 'The human gut resistome', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1670), p. 20140087. doi:10.1098/rstb.2014.0087.

Schildkraut, J.A., Coolen, J.P.M., Burbaud, S., Sangen, J.J.N., Kwint, M.P., Floto, R.A., Camp, H.J.M. op den, Brake, L.H.M. te, Wertheim, H.F.L., Neveling, K., Hoefsloot, W. and Ingen, J. van (2021) 'RNA Sequencing Elucidates Drug-Specific Mechanisms of Antibiotic Tolerance and Resistance in *Mycobacterium abscessus*', *Antimicrobial Agents and Chemotherapy*, 66(1), pp. e01509-21. doi:10.1128/AAC.01509-21.

Scholz, P., Haring, V., Wittmann-Liebold, B., Ashman, K., Bagdasarian, M. and Scherzinger, E. (1989) 'Complete nucleotide sequence and gene organization of the broad-host-range plasmid RSF1010', *Gene*, 75(2), pp. 271–288. doi:10.1016/0378-1119(89)90273-4.

Seemann, T. (2014) 'Prokka: rapid prokaryotic genome annotation', *Bioinformatics*, 30(14), pp. 2068–2069. doi:10.1093/bioinformatics/btu153.

Seiler, C. and Berendonk, T.U. (2012) 'Heavy metal driven co-selection of antibiotic resistance in soil and water bodies impacted by agriculture and aquaculture', *Frontiers in Microbiology*, 3. doi:10.3389/fmicb.2012.00399.

Semenov, M.V. (2021) 'Metabarcoding and Metagenomics in Soil Ecology Research: Achievements, Challenges, and Prospects', *Biology Bulletin Reviews*, 11(1), pp. 40–53. doi:10.1134/S2079086421010084.

Sepkowitz, K.A. (2011) 'One Hundred Years of Salvarsan', *New England Journal of Medicine*, 365(4), pp. 291–293. doi:10.1056/NEJMp1105345.

Shahi, S.K., Zarei, K., Guseva, N.V. and Mangalam, A.K. (2019) 'Microbiota Analysis Using Two-step PCR and Next-generation 16S rRNA Gene Sequencing', *Journal of visualized experiments : JoVE*, (152), p. 10.3791/59980. doi:10.3791/59980.

Shenhav, L., Thompson, M., Joseph, T.A., Briscoe, L., Furman, O., Bogumil, D., Mizrahi, I., Pe'er, I. and Halperin, E. (2019) 'FEAST: fast expectation-maximization for microbial source tracking', *Nature Methods*, 16(7), pp. 627–632. doi:10.1038/s41592-019-0431-x.

Shoji, M., Cui, L., Iizuka, R., Komoto, A., Neoh, H., Watanabe, Y., Hishinuma, T. and Hiramatsu, K. (2011) '*walK* and *clpP* mutations confer reduced vancomycin susceptibility in *Staphylococcus aureus'*, *Antimicrobial Agents and Chemotherapy*, 55(8), pp. 3870–3881. doi:10.1128/AAC.01563-10.

Sinel, C., Augagneur, Y., Sassi, M., Bronsard, J., Cacaci, M., Guérin, F., Sanguinetti, M., Meignen, P., Cattoir, V. and Felden, B. (2017) 'Small RNAs in vancomycin-resistant *Enterococcus faecium* involved in daptomycin response and resistance', *Scientific Reports*, 7, p. 11067. doi:10.1038/s41598-017-11265-2.

Sivertsen, A., Pedersen, T., Larssen, K.W., Bergh, K., Rønning, T.G., Radtke, A. and Hegstad, K. (2016) 'A Silenced *vanA* Gene Cluster on a Transferable Plasmid Caused an Outbreak of Vancomycin-Variable Enterococci', *Antimicrobial Agents and Chemotherapy*, 60(7), pp. 4119-4127. doi:10.1128/AAC.00286-16.

Snyder, H., Kellogg, S.L., Skarda, L.M., Little, J.L. and Kristich, C.J. (2014) 'Nutritional Control of Antibiotic Resistance via an Interface between the Phosphotransferase System and a Two-Component Signaling System', *Antimicrobial Agents and Chemotherapy*, 58(2), pp. 957–965. doi:10.1128/AAC.01919-13.

Sparo, M., Delpech, G. and García Allende, N. (2018) 'Impact on Public Health of the Spread of High-Level Resistance to Gentamicin and Vancomycin in Enterococci', *Frontiers in Microbiology*, 9:3073. doi:10.3389/fmicb.2018.03073

Srikumar, S., Kröger, C., Hébrard, M., Colgan, A., Owen, S.V., Sivasankaran, S.K., Cameron, A.D.S., Hokamp, K. and Hinton, J.C.D. (2015) 'RNA-seq Brings New Insights to the Intra-Macrophage Transcriptome of *Salmonella Typhimurium*', *PLOS Pathogens*, 11(11), p. e1005262. doi:10.1371/journal.ppat.1005262.

Staerck, C., Wasselin, V., Budin-Verneuil, A., Rincé, I., Cacaci, M., Weigel, M., Giraud, C., Hain, T., Hartke, A. and Riboulet-Bisson, E. (2021) 'Analysis of glycerol and dihydroxyacetone metabolism in *Enterococcus faecium*', *FEMS Microbiology Letters*, 368(8), p. fnab043. doi:10.1093/femsle/fnab043.

Stamatakis, A. (2014) 'RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies', *Bioinformatics*, 30(9), pp. 1312–1313. doi:10.1093/bioinformatics/btu033.

Stanton, I.C., Murray, A.K., Zhang, L., Snape, J. and Gaze, W.H. (2020) 'Evolution of antibiotic resistance at low antibiotic concentrations including selection below the minimal selective concentration', *Communications Biology*, 3(1), pp. 1–11. doi:10.1038/s42003-020-01176-w.

Sterling, A.J., Snelling, W.J., Naughton, P.J., Ternan, N.G. and Dooley, J.S.G. (2020) 'Competent but complex communication: The phenomena of pheromone-responsive plasmids', *PLoS Pathogens*, 16(4), p. e1008310. doi:10.1371/journal.ppat.1008310.

Stogios, P.J. and Savchenko, A. (2020) 'Molecular mechanisms of vancomycin resistance', *Protein Science*, 29(3), pp. 654–669. doi:10.1002/pro.3819.

Sze, M.A. and Schloss, P.D. (2019) 'The Impact of DNA Polymerase and Number of Rounds of Amplification in PCR on 16S rRNA Gene Sequence Data', *mSphere*, 4(3), pp. e00163-19. doi:10.1128/mSphere.00163-19.

Tan, G., Opitz, L., Schlapbach, R. and Rehrauer, H. (2019) 'Long fragments achieve lower base quality in Illumina paired-end sequencing', *Scientific Reports*, 9(1), p. 2856. doi:10.1038/s41598-019-39076-7.

Tanimoto, K. and Ike, Y. (2008) 'Complete nucleotide sequencing and analysis of the 65-kb highly conjugative Enterococcus faecium plasmid pMG1: identification of the transfer-related region and the minimum region required for replication', *FEMS Microbiology Letters*, 288(2), pp. 186–195. doi:10.1111/j.1574-6968.2008.01342.x.

Tattevin, P., Levy Hara, G., Toumi, A., Enani, M., Coombs, G., Voss, A., Wertheim, H., Poda, A., Daoud, Z., Laxminarayan, R., Nathwani, D. and Gould, I. (2020) 'Advocacy for Increased International Efforts for Antimicrobial Stewardship Actions in Low-and Middle-Income Countries on Behalf of Alliance for the Prudent Use of Antimicrobials (APUA), Under the Auspices of the International Society of Antimicrobial Chemotherapy (ISAC)', *Frontiers in Medicine*, 7:503. doi:10.3389/fmed.2020.00503.

Taylor, R.G., Walker, D.C. and McInnes, R.R. (1993) '*E. coli* host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing.', *Nucleic Acids Research*, 21(7), pp. 1677–1678.

Teng, F., Kawalec, M., Weinstock, G.M., Hryniewicz, W. and Murray, B.E. (2003) 'An *Enterococcus faecium* Secreted Antigen, SagA, Exhibits Broad-Spectrum Binding to Extracellular Matrix Proteins and Appears Essential for *E. faecium* Growth', *Infection and Immunity*, 71(9), pp. 5033–5041. doi:10.1128/IAI.71.9.5033-5041.2003.

TESSy & ECDC (2021) *Data from The European Surveillance System – TESSy, and released by ECDC*.

Teufel, F., Almagro Armenteros, J.J., Johansen, A.R., Gíslason, M.H., Pihl, S.I., Tsirigos, K.D., Winther, O., Brunak, S., von Heijne, G. and Nielsen, H. (2022) 'SignalP 6.0 predicts all five types of signal peptides using protein language models', *Nature Biotechnology*, pp. 1–3. doi:10.1038/s41587-021-01156-3.

Thaker, M.N., Kalan, L., Waglechner, N., Eshaghi, A., Patel, S.N., Poutanen, S., Willey, B., Coburn, B., McGeer, A., Low, D.E. and Wright, G.D. (2015) 'Vancomycin-Variable Enterococci Can Give Rise to Constitutive Resistance during Antibiotic Therapy', *Antimicrobial Agents and Chemotherapy*, 59(3), pp. 1405–1410. doi:10.1128/AAC.04490-14.

Theilacker, C., Sava, I., Sanchez-Carballo, P., Bao, Y., Kropec, A., Grohmann, E., Holst, O. and Huebner, J. (2011) 'Deletion of the glycosyltransferase *bgsB* of *Enterococcus faecalis* leads to a complete loss of glycolipids from the cell membrane and to impaired biofilm formation', *BMC Microbiology*, 11(1), p. 67. doi:10.1186/1471-2180-11-67.

Theuretzbacher, U., Outterson, K., Engel, A. and Karlén, A. (2020) 'The global preclinical antibacterial pipeline', *Nature Reviews Microbiology*, 18(5), pp. 275–285. doi:10.1038/s41579-019-0288-0.

Thomas, T., Gilbert, J. and Meyer, F. (2012) 'Metagenomics - a guide from sampling to data analysis', *Microbial Informatics and Experimentation*, 2, p. 3. doi:10.1186/2042-5783-2-3.

Tomita, H. and Ike, Y. (2008) 'Genetic Analysis of the *Enterococcus* Vancomycin Resistance Conjugative Plasmid pHTβ: Identification of the Region Involved in Cell Aggregation and *traB*, a Key Regulator Gene for Plasmid Transfer and Cell Aggregation', *Journal of Bacteriology*, 190(23), pp. 7739-7753. doi:10.1128/JB.00361-08.

Tomita, H., Tanimoto, K., Hayakawa, S., Morinaga, K., Ezaki, K., Oshima, H. and Ike, Y. (2003) 'Highly Conjugative pMG1-Like Plasmids Carrying Tn1546-Like Transposons That Encode Vancomycin Resistance in *Enterococcus faecium*', *Journal of Bacteriology*, 185(23), pp. 7024–7028. doi:10.1128/JB.185.23.7024-7028.2003.

Tomlinson, B.R., Malof, M.E. and Shaw, L.N.Y. 2021 (2021) 'A global transcriptomic analysis of *Staphylococcus aureu*s biofilm formation across diverse clonal lineages', *Microbial Genomics*, 7(7), p. 000598. doi:10.1099/mgen.0.000598.

Top, J., Willems, R. and Bonten, M. (2008) 'Emergence of CC17 *Enterococcus faecium*: from commensal to hospital-adapted pathogen', *FEMS Immunology & Medical Microbiology*, 52(3), pp. 297–308. doi:10.1111/j.1574-695X.2008.00383.x.

Trachsel, E., Redder, P., Linder, P. and Armitano, J. (2019) 'Genetic screens reveal novel major and minor players in magnesium homeostasis of *Staphylococcus aureus*', *PLoS Genetics*, 15(8), p. e1008336. doi:10.1371/journal.pgen.1008336.

Tram, G., Klare, W.P., Cain, J.A., Mourad, B., Cordwell, S.J., Day, C.J. and Korolik, V. (2020) 'Assigning a role for chemosensory signal transduction in *Campylobacter jejuni*

biofilms using a combined omics approach', *Scientific Reports*, 10(1), p. 6829. doi:10.1038/s41598-020-63569-5.

Truong, D.T., Franzosa, E.A., Tickle, T.L., Scholz, M., Weingart, G., Pasolli, E., Tett, A., Huttenhower, C. and Segata, N. (2015) 'MetaPhlAn2 for enhanced metagenomic taxonomic profiling', *Nature Methods*, 12(10), pp. 902–903. doi:10.1038/nmeth.3589.

Turco, E.R.D., Bartoletti, M., Dahl, A., Cervera, C. and Pericàs, J.M. (2021) 'How do I manage a patient with enterococcal bacteraemia?', *Clinical Microbiology and Infection*, 27(3), pp. 364–371. doi:10.1016/j.cmi.2020.10.029.

Valdezate, S., Labayru, C., Navarro, A., Mantecón, M.A., Ortega, M., Coque, T.M., García, M. and Saéz-Nieto, J.A. (2009) 'Large clonal outbreak of multidrug-resistant CC17 ST17 *Enterococcus faecium* containing Tn5382 in a Spanish hospital', *The Journal of Antimicrobial Chemotherapy*, 63(1), pp. 17–20. doi:10.1093/jac/dkn449.

Valdezate, S., Miranda, C., Navarro, A., Freitas, A.R., Cabrera, J.J., Carrasco, G., Coque, T.M., Jiménez-Romano, E. and Saéz-Nieto, J.A. (2012) 'Clonal outbreak of ST17 multidrug-resistant *Enterococcus faecium* harbouring an Inc18-like::Tn1546 plasmid in a haemo-oncology ward of a Spanish hospital', *Journal of Antimicrobial Chemotherapy*, 67(4), pp. 832–836. doi:10.1093/jac/dkr545.

Van Goethem, N., Descamps, T., Devleesschauwer, B., Roosens, N.H.C., Boon, N.A.M., Van Oyen, H. and Robert, A. (2019) 'Status and potential of bacterial genomics for public health practice: a scoping review', *Implementation Science*, 14(1), p. 79. doi:10.1186/s13012-019-0930-2.

Van Tyne, D., Martin, M.J. and Gilmore, M.S. (2013) 'Structure, Function, and Biology of the *Enterococcus faecalis* Cytolysin', *Toxins*, 5(5), pp. 895–911. doi:10.3390/toxins5050895.

Vimberg, V., Zieglerová, L., Buriánková, K., Branny, P. and Balíková Novotná, G. (2020) 'VanZ Reduces the Binding of Lipoglycopeptide Antibiotics to *Staphylococcus aureus* and *Streptococcus pneumoniae* Cells', *Frontiers in Microbiology*, 11, p. 566. doi:10.3389/fmicb.2020.00566.

Viswanath, L.S., Sugumar, M., Chandra Murthy Peela, S., Walia, K. and Sistla, S. (2022) 'Detection of vancomycin variable enterococci (VVE) among clinical isolates of *Enterococcus faecium* collected across India-first report from the subcontinent', *Indian Journal of Medical Microbiology*, 40(2), pp. 285-288. doi:10.1016/j.ijmmb.2021.12.011.

Wagner, J., Gruz, P., Kim, S.R., Yamada, M., Matsui, K., Fuchs, R.P. and Nohmi, T. (1999) 'The *dinB* gene encodes a novel *E. coli* DNA polymerase, DNA pol IV, involved in mutagenesis', *Molecular Cell*, 4(2), pp. 281–286. doi:10.1016/s1097-2765(00)80376-7.

Wagner, T.M., Janice, J., Sivertsen, A., Sjögren, I., Sundsfjord, A. and Hegstad, K. (2021) 'Alternative *vanHAX* promoters and increased *vanA*-plasmid copy number resurrect silenced glycopeptide resistance in *Enterococcus faecium*', *Journal of Antimicrobial Chemotherapy*, 76(4), pp. 876–882. doi:10.1093/jac/dkaa541.

Wang, B. and Muir, T.W. (2016) 'Regulation of virulence in *Staphylococcus aureus*: molecular mechanisms and remaining puzzles', *Cell chemical biology*, 23(2), pp. 214–224. doi:10.1016/j.chembiol.2016.01.004.

Wang, R., Dorp, L. van, Shaw, L.P., Bradley, P., Wang, Q., Wang, X., Jin, L., Zhang, Q., Liu, Y., Rieux, A., Dorai-Schneiders, T., Weinert, L.A., Iqbal, Z., Didelot, X., Wang, H. and Balloux, F. (2018) 'The global distribution and spread of the mobilized colistin resistance gene *mcr-1*', *Nature Communications*, 9(1), pp. 1–9. doi:10.1038/s41467-018-03205-z.

Wang, Y., Lv, Y., Cai, J., Schwarz, S., Cui, L., Hu, Z., Zhang, R., Li, J., Zhao, Q., He, T., Wang, D., Wang, Z., Shen, Y., Li, Y., Feßler, A.T., Wu, C., Yu, H., Deng, X., Xia, X. and Shen, J. (2015) 'A novel gene, *optrA*, that confers transferable resistance to oxazolidinones and phenicols and its presence in *Enterococcus faecalis* and *Enterococcus faecium* of human and animal origin', *Journal of Antimicrobial Chemotherapy*, 70(8), pp. 2182–2190. doi:10.1093/jac/dkv116.

Wang, Z., Gerstein, M. and Snyder, M. (2009) 'RNA-Seq: a revolutionary tool for transcriptomics', *Nature reviews. Genetics*, 10(1), pp. 57–63. doi:10.1038/nrg2484.

Wardal, E., Kuch, A., Gawryszewska, I., Żabicka, D., Hryniewicz, W. and Sadowy, E. (2017) 'Diversity of plasmids and Tn1546-type transposons among VanA *Enterococcus faecium* in Poland', *European Journal of Clinical Microbiology & Infectious Diseases*, 36(2), pp. 313–328. doi:10.1007/s10096-016-2804-8.

Warr, A.R., Hubbard, T.P., Munera, D., Blondel, C.J., Wiesch, P.A. zur, Abel, S., Wang, X., Davis, B.M. and Waldor, M.K. (2019) 'Transposon-insertion sequencing screens unveil requirements for EHEC growth and intestinal colonization', *PLOS Pathogens*, 15(8), p. e1007652. doi:10.1371/journal.ppat.1007652.

Weaver, K.E., Kwong, S.M., Firth, N. and Francia, M.V. (2009) 'The RepA_N replicons of Gram-positive bacteria: a family of broadly distributed but narrow host range plasmids', *Plasmid*, 61(2), pp. 94–109. doi:10.1016/j.plasmid.2008.11.004.

Wen, X., Gehring, R., Stallbaumer, A., Riviere, J.E. and Volkova, V.V. (2016) 'Limitations of MIC as sole metric of pharmacodynamic response across the range of antimicrobial susceptibilities within a single bacterial species', *Scientific Reports*, 6(1), p. 37907. doi:10.1038/srep37907.

Wick, R.R., Judd, L.M., Gorrie, C.L. and Holt, K.E. (2017) 'Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads', *PLOS Computational Biology*, 13(6), p. e1005595. doi:10.1371/journal.pcbi.1005595.

Wilhelm, S.W., LeCleir, G.R., Bullerjahn, G.S., McKay, R.M., Saxton, M.A., Twiss, M.R. and Bourbonniere, R.A. (2014) 'Seasonal changes in microbial community structure and activity imply winter production is linked to summer hypoxia in a large lake', *FEMS Microbiology Ecology*, 87(2), pp. 475–485. doi:10.1111/1574-6941.12238.

Willems, R.J.L., Top, J., van Santen, M., Robinson, D.A., Coque, T.M., Baquero, F., Grundmann, H. and Bonten, M.J.M. (2005) 'Global Spread of Vancomycin-resistant

Enterococcus faecium from Distinct Nosocomial Genetic Complex', *Emerging Infectious Diseases*, 11(6), pp. 821–828. doi:10.3201/eid1106.041204.

Wood, D.E., Lu, J. and Langmead, B. (2019) 'Improved metagenomic analysis with Kraken 2', *Genome Biology*, 20(1), p. 257. doi:10.1186/s13059-019-1891-0.

Woolhouse, M., Ward, M., van Bunnik, B. and Farrar, J. (2015) 'Antimicrobial resistance in humans, livestock and the wider environment', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1670), p. 20140083. doi:10.1098/rstb.2014.0083.

Workman, R.E., Tang, A.D., Tang, P.S., Jain, M., Tyson, J.R., Razaghi, R., Zuzarte, P.C., Gilpatrick, T., Payne, A., Quick, J., Sadowski, N., Holmes, N., de Jesus, J.G., Jones, K.L., Soulette, C.M., Snutch, T.P., Loman, N., Paten, B., Loose, M., Simpson, J.T., Olsen, H.E., Brooks, A.N., Akeson, M. and Timp, W. (2019) 'Nanopore native RNA sequencing of a human poly(A) transcriptome', *Nature Methods*, 16(12), pp. 1297–1305. doi:10.1038/s41592-019-0617-2.

Wright, G.D., Molinas, C., Arthur, M., Courvalin, P. and Walsh, C.T. (1992) 'Characterization of *vanY*, a DD-carboxypeptidase from vancomycin-resistant *Enterococcus faecium* BM4147.', *Antimicrobial Agents and Chemotherapy*, 36(7), pp. 1514–1518.

Yang, J., Jiang, Y., Guo, L., Ye, Li., Ma, Y. and Luo, Y. (2016) 'Prevalence of Diverse Clones of Vancomycin-Resistant *Enterococcus faecium* ST78 in a Chinese Hospital', *Microbial Drug Resistance (Larchmont, N.Y.)*, 22(4), pp. 294–300. doi:10.1089/mdr.2015.0069.

Yong, D., Toleman, M.A., Giske, C.G., Cho, H.S., Sundman, K., Lee, K. and Walsh, T.R. (2009) 'Characterization of a New Metallo-β-Lactamase Gene, blaNDM-1, and a Novel Erythromycin Esterase Gene Carried on a Unique Genetic Structure in *Klebsiella pneumoniae* Sequence Type 14 from India', *Antimicrobial Agents and Chemotherapy*, 53(12), pp. 5046-5054. doi:10.1128/AAC.00774-09.

Yu, Y., Wu, G., Wang, C., Lu, N., Yuan, X. and Zhu, X. (2019) 'Pollution characteristics of antibiotics and antibiotic resistance of coliform bacteria in the Yitong River, China', *Environmental Monitoring and Assessment*, 191(8), p. 516. doi:10.1007/s10661-019-7660-4.

Zaffiri, L., Gardner, J. and Toledo-Pereyra, L.H. (2012) 'History of Antibiotics. From Salvarsan to Cephalosporins', *Journal of Investigative Surgery*, 25(2), pp. 67–77. doi:10.3109/08941939.2012.664099.

Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., Aarestrup, F.M. and Larsen, M.V. (2012) 'Identification of acquired antimicrobial resistance genes', *Journal of Antimicrobial Chemotherapy*, 67(11), pp. 2640–2644. doi:10.1093/jac/dks261.

Zavascki, A.P., Goldani, L.Z., Li, J. and Nation, R.L. (2007) 'Polymyxin B for the treatment of multidrug-resistant pathogens: a critical review', *Journal of Antimicrobial Chemotherapy*, 60(6), pp. 1206–1215. doi:10.1093/jac/dkm357.

Zhang, L. and Ji, G. (2004) 'Identification of a Staphylococcal AgrB Segment(s) Responsible for Group-Specific Processing of AgrD by Gene Swapping', *Journal of Bacteriology*, 186(20), pp. 6706–6713. doi:10.1128/JB.186.20.6706-6713.2004.

Zhang, X., Bierschenk, D., Top, J., Anastasiou, I., Bonten, M.J., Willems, R.J. and van Schaik, W. (2013) 'Functional genomic analysis of bile salt resistance in *Enterococcus faecium*', *BMC Genomics*, 14(1), p. 299. doi:10.1186/1471-2164-14-299.

Zhang, X., Jiang, C. and Zhou, C. (2021) 'Diagnosis of *Enterococcus faecalis* meningitis associated with long-term cerebrospinal fluid rhinorrhoea using metagenomics next-generation sequencing: a case report', *BMC Infectious Diseases*, 21(1), p. 1105. doi:10.1186/s12879-021-06797-y.

Zhang, X., de Maat, V., Guzmán Prieto, A.M., Prajsnar, T.K., Bayjanov, J.R., de Been, M., Rogers, M.R.C., Bonten, M.J.M., Mesnage, S., Willems, R.J.L. and van Schaik, W. (2017) 'RNA-seq and Tn-seq reveal fitness determinants of vancomycin-resistant *Enterococcus faecium* during growth in human serum', *BMC Genomics*, 18(1), p. 893. doi:10.1186/s12864-017-4299-9.

Zhang, X, Paganelli, F.L., Bierschenk, D., Kuipers, A. and Bonten, M.J.M. (2012) 'Genome-Wide Identification of Ampicillin Resistance Determinants in *Enterococcus faecium*', *PLoS Genet*, 8(6), p. 1002804. doi:10.1371/journal.pgen.1002804.

Zhang, Xinglin, Paganelli, F.L., Bierschenk, D., Kuipers, A., Bonten, M.J.M., Willems, R.J.L. and van Schaik, W. (2012) 'Genome-wide identification of ampicillin resistance determinants in *Enterococcus faecium*', *PLoS genetics*, 8(6), p. e1002804. doi:10.1371/journal.pgen.1002804.

Zhang, X., Top, J., de Been, M., Bierschenk, D., Rogers, M., Leendertse, M., Bonten, M.J.M., van der Poll, T., Willems, R.J.L. and van Schaik, W. (2013) 'Identification of a Genetic Determinant in Clinical *Enterococcus faecium* Strains That Contributes to Intestinal Colonization During Antibiotic Treatment', *The Journal of Infectious Diseases*, 207(11), pp. 1780–1786. doi:10.1093/infdis/jit076.

Zhao, S., Fung-Leung, W.-P., Bittner, A., Ngo, K. and Liu, X. (2014) 'Comparison of RNA-Seq and Microarray in Transcriptome Profiling of Activated T Cells', *PLOS ONE*, 9(1), p. e78644. doi:10.1371/journal.pone.0078644.

Zheng, J., Wu, Y., Lin, Z., Wang, G., Jiang, S., Sun, X., Tu, H., Yu, Z. and Qu, D. (2020) 'ClpP participates in stress tolerance, biofilm formation, antimicrobial tolerance, and virulence of *Enterococcus faecalis*', *BMC Microbiology*, 20(1), p. 30. doi:10.1186/s12866-020-1719-9.

Zhong, F., Gao, Y., Yu, T., Zhang, Y., Xu, D., Xiao, E., He, F., Zhou, Q. and Wu, Z. (2011) 'The management of undesirable cyanobacteria blooms in channel catfish ponds using a constructed wetland: Contribution to the control of off-flavor occurrences', *Water Research*, 45(19), pp. 6479–6488. doi:10.1016/j.watres.2011.09.044.

Zhou, X., Willems, R.J.L., Friedrich, A.W., Rossen, J.W.A. and Bathoorn, E. (2020) '*Enterococcus faecium*: from microbiological insights to practical recommendations for infection control and diagnostics', *Antimicrobial Resistance & Infection Control*, 9(1), p. 130. doi:10.1186/s13756-020-00770-1.

Zhu, W., Murray, P.R., Huskins, W.C., Jernigan, J.A., McDonald, L.C., Clark, N.C., Anderson, K.F., McDougal, L.K., Hageman, J.C., Olsen-Rasmussen, M., Frace, M., Alangaden, G.J., Chenoweth, C., Zervos, M.J., Robinson-Dunn, B., Schreckenberger, P.C., Reller, L.B., Rudrik, J.T. and Patel, J.B. (2010) 'Dissemination of an *Enterococcus* Inc18-Like *vanA* Plasmid Associated with Vancomycin-Resistant *Staphylococcus aureus*', *Antimicrobial Agents and Chemotherapy*, 54(10), pp. 4314-4320. doi:10.1128/AAC.00185-10.

Chapter 8 Appendices

Table 8.1: DNA sequences of the gBlocks used to delete genes that were identified in the Tn-seq analysis from *E. faecium* E745.

Deletion construct	Sequence (5' -> 3')
asnB	TGCAAAAGAAAAGCATGGGATGCGTTGGACCAAGTATCGAGGCTTGGAAAAAGTCGCGACGCATACGATGCTAGTGTTTGCTGCAATGAAT TTAAAGAAATTAGCCACCTGGCTATGGAAAGGCAAGGAGCCTTTATTTTTTTGTTCAAAAATTCGAAACGAAGTAGATAAAAAGCTGTTCC AAGCTCGAGTTACGAGTTTGGAACAGCTTTTGTCTACAGTCTGAGACACGATCTTTAGGTCGTGTCTTTTTTTGCTGGTTTTTTTGGAAAT AGCATCCACTAATATTTTTTTAACACGAAGAGAATAAGGTAGCCTCAAAAAAGAAAACATGGTAAATTAGATATTGGGACTTTATTTTCTG TAAAAACAGAAACTTTCATCTTTTGTTTTCACGTGACGAAATTGTAAGAGAGAAAATATTCGTCCGCTATTTATTCTATGATAAGATGCAT ATGAATGAAATAAACAGTAATTTTATATTAGGATGTGAAAGTTAAAGATGAAAATCAAATATCAATAAAAAAGAGGTTTTCTCCTACCGTT TAACACTGAGTGATAAGAAACAATTTCAAAAAACAGCTCTTCCCGTTTTATAGATATTATGACTTATGGTTCGGAGTTGGCAGAGGAACAC CATTGTTTCGAACATAAGGGGGTTGTGACACGATTTTCGTCACAACCCCTTTACATATGTTCAATAATCGATTTCTTTTTCTTTTGTCCCA TCTTTTTTATAAATGAACAAAGTTGATTCTTTGTTTTGTGCTACTTGCTTTCCTTTTTCGATTGCTTCTTTTTTTGTATCAAAATGGTTGC TTGCCTTTTTTGCCCCTTTAGATTTCACGATCCATTGTTCTTCCTCATATTCTACGATCACATCTGAATCAAGCAGTTTTCCAGCACGAGG ATTCGTATCGTGTTCATCCGTCTTGCTTGGATTTTTTTCTTTTTCGAAAGTTCGTTTTTCTGATTCACTAGCATCTTCATACCATTTTTTT
avrA	GAAAAAGAAGGTGGTGCTCGCTGCAGTGCATGCTTCAATATGCGATTAGATCTTGTTGCGCAAAAAGCGCAAGAACTAGGATACGATTACT TTGGAAGTGCCCTGACCATCTCGCCAAAGAAAAATGCAGCTTTGATCAATCAGATCGGAATGGATATCCAAAAGATTTACGGGACACAATA CTTGCCAAGTGATTTCAAGAAAAACAAAGGCTATGAACGCTCTTTGCAAATGTGCAAAGAGTATGATATTTATCGCCAATGTTATTGTGGT TGCGTGTTCGCTGCCAAGCAGCAAGGATGCGATCTGAAAGAAATCAATCGGGAAGCAAAAGCTTACTTAAAAACATCAACCGTAGTTTTTG AGGAAAATTAAAGAAATCTTTCGGAAAAACGTAAGAAAAAAATAAAGAATTTATGGTAGAAATAAGGTAGACTGTATAATTTGTCCGTAGT AAAGTAAGGAAACAAGAATAAATTTCAATGAAGGAAGAGATAAGAGACATTTTTACTAGACCTTATCAACGGGGTGCATAAATTCCTTGGC TATTTAGACATCAGTCCTAAATATCTTAACCGTGGATATACGATTTTAAGTGTGATTCCCACGCTTTATTTATTGAGAATTGTTTATGGAC TTTGGCAAAACCAAAACTATCTTCAATTTTTCTTATATGGACTTGCATTTCTAGTGCTTGTTTATTTTACTGTTTTAAATGTATTTTATTA CTTTTTAGATAAAAACACGAAAGCCGATGTCACTCAGCTTTTTGTGAAGTATTTACCTGATGAAGCGTTCAACATCCAGACAGAAGAATCG ATAGATAATGGAAATTCGATCGATAAAGTGAATACACAAGAGGTCATGGTTTCTTATGACGAAGATTATCAATTGAAATTAGCCGAGAATA TGCGGTATTTGATTGGAAATGGTGAGATCAAAACAAATGAACTCGGTCGCATGGATGGATTTCTGGTCGATAGAAATACATTGTATCCTT
avrB	ATTTCAATGAAGGAAGAGATAAGATGCGAAAAGTAGCACCTATTTTGAAGCCAATCGCTGTATTAACAATCTCGACGATTGTCGCGACTAC TACAGGTATTGTGAATAGTTCCAGCTACCGTGCACCAAGTTCAAGTTCTAGTACTGTAGAAGAAACAACACCAACTACTTCAGTCTCGACA TCTGAACCAACAACAACGGAAAGTTTGAATGAAACCACTACGACAAGTACATGGGACGATACAACAACTAGTTCGTCGTATGATGAGACAT ATACAACTGACAGCTCAAGCGAAACAAGTAATACGGTACCAGATGGCACTCCATCGGATTCAACCAATCCAAGTACAAGTCCGGATACAAG TGCTAGCTCCGAAACGATACCAAGCAGTACGACTGTAGATGATAACACGCAAAATGATGATGCAGCACGTACAGCACAGCAAAACCAAGAT TCAGGAAATACCCAAACGTATAACCAGCAAGCAGGTGATTCTCAAAATGTGATGAAACGCCGTGTATCAGTGGTTGCGATAAAAACGTTTA TTCGGAATAAGGGGGATGTGACACGTCTTTTGTCACGTTCCTTTTTTATATAACGAGAAAGATAAGACGAACACATCACTTATTCTTTCCT TTCATTTATAAACATTGTAGAATAAAGAAAGTACAACGAAGTATTGTGAAAAAACGTAATAAATAACGAGGAAAATACAAGTAGGAGGTAA GCTGTGAAAGAATTTTTGAAAAATTGGGGGATTTTGATCCTTGTATTAGCGGCTATTTTATTAGCTAGAGTGTACGTGTTTACACCAGTTA

	CTGTAAACGGGCATTCGATGGACCCTACATTAAGTGATGGACAACGATTGATTTCATCTAAGATCTCAAACTATGAACGAATGGATATTAT TACTACAAAGGAACCGGGGGATGAAGAACGGATGATTGTTAAACGGATCATCGGTATGCCTGGAGATACAGTCAAGATGGAAAATGACCA
glpO	ACAAGCGAAGATGAGGTCTATGTTGTTCCTGCATTTGTTGGACTGGGTGCTCCTTATTGGGATCAGGCAGCACGAGGGGCTATGTTTGGTT TGACACGTGGAACGACAAAAGAAGATATCATTAAAGCTACTTTGCAGTCGATTGCCTATCAAGTGAGAGATATCATCGATACGATGCAAGA TGACACTGGAATCAAGATTCCTGTATTGAAAGTAGATGGCGGTGCAGCAAACAATGAATATCTGATGCAATTCCAAACAGATATCTTGAAT GTACCGATCCAACGAGCAGAAAACTTGGAAACAACGGCGCTAGGGGCAGCATTTTTAGCAGGATTAGCTGTCGGTTATTGGAAAGATACAG ACGAGATCCGCGAATTTTATGAAGCTGGCAAGCTATTTGAAGTTCAAATGGAGGAAGAGCGCCGAGAAAAACTTTACAACGGTTGGAAAAA AGCAGTAAAAGCTACCCAAGCATTTGAATAAGGAGGAAATCCATAAAATGAACAGTGATATGACTCAAATCATGGGGGAGTTTATTGGGAC ACTGATCCTAGTTTTGCTTGGAGATGGTGTGTGTGCGGCTGTGAATTTAAACAAAAGTAAAGCACAGGCTTCTGGCTGGATTGTCATCGCA TTTGGTTGGGGACTCGCTGTAACAATGGCTGTCTATATTTCTGGATTCATGGGCCCAGCTCATTTGAATCCTGCAGTCTCTCTAGCTATGG CGATGACAGGGGCAATTAGTTGGAATCTTGTCGTTCCTTTCATCATTGCGCAAGTTTTAGGAGCTTTTGCAGGGGCCATTCTTGTTTGGCT TTCTTATTTGCCACATTGGAATGCAACGAAAGATGAAAGTGCGATTTTAGGCACGTTTGCAACAGGTCCTGCTATACGGAATTATCCAGCA AATGTGATCACAGAATTAATTGGTACATTTGTATTAGTTCTTGGGTTGCTAGCTTTTGGACAAAACGAATTTGCACCTGGAACGAACGTA
gshAB	TAGAAAAAATAAAACGAGTAATCCAGCGGGAATTACCAGAAGCACCCCAAGAAACACTTCTTGTACTAGATGCAACAACTGGTCAAAACGC GATGGTTCAAGCAAAACAATTCAAAGAAACAACAGATGTGACCGGTCTTGTTTTGACAAAATTAGATGGTACAGCAAAGGGCGGTATCGTA CTTGCTATCCGTAATGAATTGCATTTACCAGTGAAATTGGTCGGTCTTGGTGAAGGAATCGATGACTTGGAGCCGTTTGATCCAAATGATT TTGTCGTTGGATTGTTCAAAGGATTACTAAAAGAAGAATAACCCATGTCTAAAAAGCAGGAGCTGTGACAAAAGTTGCAGCTCCTTGTTTG AATGATTGTTTTAACATAGCTGTGTTTTTGATGCCTTCATTTTTATACTTATTTGTAAGCGGTTCTCAAACAATTTCGTTCTTTTTCTTAG ACTTTTTGATTATAATTAGATTCATATGAATTGGAGGTCACTTAAAAAAGAAAAAGCTTTTGACAGATTGTCCTATTTGATGTCAATCTGA TCAAAAGCTTTTTTGCATGTAAAAATTTATGTGTTTTTTCGTTTTAAAATAAACGATCCCAGAAGCTTTTTTTCTTTTTTGGTTTTTCTTT TTCAGCTGGCGAAGAGTTTTCTTCTGGGTATTTTTTTTCGATATCAATCACAGGTTCATTTACTGCTTCTTTTGCTGTTAAGAGCAATCCG AATTCCTCAGGATTTGCAGTGACGTAATCGTTGACTACAGTAAATCGGCTATTGTTTTTCGTTGCTAGCTGGATATATTGGTTTTGGACAG AAAGGGGTAGAGCCCCATTCAAGAGTATGGAAGCATCTGGATGTTTTTTGAGTTCTTTTAAAAAGTTTTCTTTGTTTTTTGGTTGGATCAT TTCAGCAACAGTCATACTAAGATAGCAGCGTTCCCTGAATGTCCCTAAATATTTGTGCTGTTCTTCAGGGTTCACTAAAGGTGTGCCATA

The underlined sequence highlights the upstream homology arm and the sequence that is not underlined highlights the downstream homology arm.

Table 8.2: Tn-seq significant genes – E745

Locus tag	Gene name	Fold change	Adjusted *P*-value
EfmE745_03026	*vanX*	-171.3	4.78E-202
EfmE745_03027	*vanA*	-53.5	3.62E-09
EfmE745_03030	*vanR*	-45.0	8.96E-19
EfmE745_03029	*vanS*	-20.8	1.19E-22
EfmE745_02202	*mgs*	-12.7	2.25E-51
EfmE745_02840	*EfmE745_02840*	-8.4	1.05E-19
EfmE745_02810	*topB_3*	-5.5	4.94E-17
EfmE745_00312	*npr_1*	-4.5	0.000263566
EfmE745_00852	*avrA*	-4.4	4.57E-07
EfmE745_02436	*rarA*	-3.8	1.37E-15
EfmE745_00838	*gmuF*	-3.7	2.32E-05
EfmE745_03028	*vanH*	-3.4	5.07E-17
EfmE745_01540	*tig*	-3.3	0.008147024
EfmE745_02606	*ispE*	-3.2	2.65E-05
EfmE745_02815	*ssb_3*	-2.9	0.00173128
EfmE745_02071	*yvgN*	-2.7	0.000416657
EfmE745_00116	*gshAB*	-2.7	5.59E-05
EfmE745_00315	*glpO*	-2.7	1.29E-10
EfmE745_01958	*EfmE745_01958*	-2.5	9.08E-05
EfmE745_00313	*EfmE745_00313*	-2.5	3.69E-10
EfmE745_00827	*pgl_1*	-2.4	0.004315705
EfmE745_00314	*glpK*	-2.3	3.58E-06
EfmE745_02354	*EfmE745_02354*	-2.3	0.00017301
EfmE745_00337	*artM_2*	-2.2	0.017712923
EfmE745_00150	*treR*	-2.2	3.39E-10
EfmE745_01049	*zapA*	-2.1	0.007956696
EfmE745_00662	*dinB_2*	-2.1	1.49E-05
EfmE745_02634	*yfmO_1*	-2.1	3.52E-06
EfmE745_01129	*clsA_1*	-2	1.36E-06
EfmE745_00687	*algB*	-2	3.25E-06

Table 8.3: Tn-seq significant genes – E8202

Locus Tag	Gene name	Fold Change	Adjusted *P*-value
EQB38_RS16800	*vanX*	-95.2	< 2.225074E-308
EQB38_RS09075	*asnB*	-32.0	1.34E-05
EQB38_RS00600	*gshAB*	-27.6	1.76E-22
EQB38_RS16795	*vanA*	-24.0	4.64E-31
EQB38_RS01810	*glpO*	-17.3	1.78E-51
EQB38_RS00075	*ybiT_1*	-13.6	3.97E-53
EQB38_RS02775	*yfdH*	-12.9	3.20E-12
EQB38_RS16790	*vanH*	-12.2	3.39E-35
EQB38_RS05815	*EQB38_RS05815*	-10.9	7.29E-52
EQB38_RS12230	*avrB*	-8.9	6.48E-58
EQB38_RS12235	*avrA*	-8.6	1.08E-48
EQB38_RS08910	*pyrH*	-7.4	0.01530318
EQB38_RS00260	*hin*	-5.9	0.00032117
EQB38_RS05735	*ypfD*	-5.5	0.01508155
EQB38_RS12490	*groL*	-5.5	0.00108307
EQB38_RS06185	*lytG*	-5.2	2.91E-29
EQB38_RS10990	*est*	-4.7	4.59E-19
EQB38_RS00775	*treR*	-4.5	1.16E-64
EQB38_RS12360	*pgl*	-4.4	0.0048834
EQB38_RS15510	*gor*	-4.4	4.12E-18
EQB38_RS01805	*glpK*	-4.4	2.10E-06
EQB38_RS02010	*EQB38_RS02010*	-4.2	4.01E-63
EQB38_RS06695	*EQB38_RS06695*	-4.1	3.56E-16
EQB38_RS07070	*yfbR*	-4.0	7.59E-09
EQB38_RS09555	*srrA_1*	-3.6	7.62E-25
EQB38_RS11455	*EQB38_RS11455*	-3.4	3.17E-36
EQB38_RS05555	*EQB38_RS05555*	-3.3	2.24E-10
EQB38_RS00105	*EQB38_RS00105*	-3.3	1.07E-07
EQB38_RS11645	*thiI*	-3.2	8.51E-08
EQB38_RS06510	*EQB38_RS06510*	-3.1	2.60E-13
EQB38_RS07120	*tagU_2*	-3.0	7.97E-11
EQB38_RS02210	*treA*	-3.0	0.01556366
EQB38_RS13140	*mecA*	-2.9	2.23E-18
EQB38_RS06810	*cysK*	-2.8	4.28E-07
EQB38_RS11060	*rpoN1*	-2.8	2.27E-53
EQB38_RS06525	*EQB38_RS06525*	-2.7	2.07E-16
EQB38_RS12040	*EQB38_RS12040*	-2.7	2.87E-07
EQB38_RS14580	*yacL*	-2.7	4.75E-27
EQB38_RS15680	*misCA*	-2.7	1.18E-07
EQB38_RS08940	*argS*	-2.7	0.04615
EQB38_RS12370	*EQB38_RS12370*	-2.6	6.46E-52
EQB38_RS17515	*hpf*	-2.5	0.00400829
EQB38_RS11650	*iscS_2*	-2.5	0.00185625

EQB38_RS16070	*xerS*	-2.5	2.01E-23
EQB38_RS01815	*glpF*	-2.5	0.00019337
EQB38_RS13175	*yfiC*	-2.5	1.28E-09
EQB38_RS05865	*rnhB*	-2.4	0.00031419
EQB38_RS09505	*ltaS1_2*	-2.3	0.01159265
EQB38_RS09550	*rcsC_1*	-2.3	7.20E-12
EQB38_RS06180	*EQB38_RS06180*	-2.3	7.50E-07
EQB38_RS10715	*ponA_2*	-2.3	0.00910194
EQB38_RS00630	*EQB38_RS00630*	-2.2	0.0008398
EQB38_RS06405	*EQB38_RS06405*	-2.2	0.01840874
EQB38_RS13625	*EQB38_RS13625*	-2.2	6.26E-10
EQB38_RS10580	*pgpH*	-2.2	1.71E-07
EQB38_RS00780	*msrB*	-2.1	3.51E-08
EQB38_RS06860	*ponA_1*	-2.1	3.82E-08
EQB38_RS08780	*EQB38_RS08780*	-2.1	0.00058457
EQB38_RS04045	*xseA*	-2.1	0.00146188
EQB38_RS07520	*napA*	-2.0	2.42E-11
EQB38_RS12110	*spoIIIE*	-2.0	0.01306184

Table 8.4: Tn-seq significant genes – Aus0004

Locus Tag	Gene name	Fold Change	Adjusted *P*-value
EFAU004_00372	*EFAU004_00372*	-647.1	5.45E-08
EFAU004_00889	*mdxG*	-496.1	3.28E-06
EFAU004_02055	*EFAU004_02055*	-444.4	1.81E-06
EFAU004_00352	*EFAU004_00352*	-333.9	0.00011968
EFAU004_00549	*EFAU004_00549*	-240.4	1.51E-08
EFAU004_01064	*EFAU004_01064*	-208.0	0.00254203
EFAU004_01733	*argS*	-147.7	0.00327615
EFAU004_00354	*EFAU004_00354*	-137.4	0.00283777
EFAU004_00593	*EFAU004_00593*	-101.9	0.00283015
EFAU004_00604	*EFAU004_00604*	-86.1	0.00904907
EFAU004_02417	*groL*	-85.5	0.00555526
EFAU004_02309	*EFAU004_02309*	-81.7	1.68E-05
EFAU004_00650	*EFAU004_00650*	-79.8	0.00555526
EFAU004_00104	*rpsQ*	-48.8	0.00013869
EFAU004_00293	*EFAU004_00293*	-42.2	0.01392982
EFAU004_01029	*EFAU004_01029*	-28.4	1.37E-37
EFAU004_00754	*EFAU004_00754*	-20.3	0.00904907
EFAU004_01792	*EFAU004_01792*	-13.1	0.00852446
EFAU004_02789	*EFAU004_02789*	-11.9	0.0327985
EFAU004_02645	*EFAU004_02645*	-11.8	0.02153222
EFAU004_01087	*recQ*	-11.8	1.90E-14
EFAU004_01233	*tuf_1*	-11.7	0.02201044
EFAU004_00024	*EFAU004_00024*	-10.8	0.03231993
EFAU004_01883	*EFAU004_01883*	-8.4	0.04358289

EFAU004_00533	*EFAU004_00533*	-7.2	0.02153222
EFAU004_02526	*EFAU004_02526*	-7.2	0.01066094
EFAU004_02523	*ecsA*	-4.7	3.93E-14
EFAU004_00411	*EFAU004_00411*	-2.3	0.0209231
EFAU004_p1049	*EFAU004_p1049*	-2.0	0.01379867