University of Nottingham
UK | CHINA | MALAYSIA

ANTIBIOTIC RESISTANCE IN DAIRY FARM SOIL: A METAGENOMIC SURVEY FROM SLURRY TO FIELD

Alexander D. Williams

B.Sc. (Hons) 2015

Thesis submitted to the University of Nottingham for the degree of

Doctor of Philosophy

Acknowledgements

Firstly, I would like to thank my supervisors; Dr Helen West, Professor Dov Stekel, Professor Christine Dodd, Dr Lisa Avery, Dr Rupert Hough and Dr Steve Hooton for their support and insight throughout my studies. I am also very grateful to my internal assessor Dr Jon Hobman, for his valuable comments during assessment meetings.

I would like to thank Andrew Warry for his advice and willingness to discuss the shifting sands of bioinformatics. I am very grateful to Dr Andrew Millard and the CLIMB-BIG-DATA project for providing access to compute resources.

I am deeply grateful to Dr Matthew Kent for his willingness to help me hone my use of R statistical software and bash command language.

I would also like to thank the Biosciences Technical Staff, including Dr Saúl Vázquez Reina, Laura Holt and John Corrie, without whom laboratories would cease to function.

Special thanks are given to the STARS CDT and NERC (NE-M009106-1), which provided the funding and opportunity to make this project possible. I am very grateful to the STARS CDT for the training, encouragement and community atmosphere they have fostered throughout my studies.

On a personal level, I wish to thank my friends and colleagues who have enriched my experience at The University of Nottingham, my family for their encouragement and Martha Ledger for her companionship and patience.

"Nothing is built on stone; all is built on sand,

but we must build as if the sand were stone."

Jorge Luis Borges

Abstract

Burgeoning antibiotic resistance (AR) threatens to undermine global human health by rendering antibiotic treatments ineffective. The rapid development of AR is widely attributed to excessive antibiotic use, and a greater proportion of global antibiotic sales are associated with livestock relative to human clinical use. In line with the One Health concept, increasing focus has been placed on studying the spread of AR in agricultural environments, including contamination of the human food chain. Livestock waste is a valuable resource for fertilising agricultural land worldwide; however, it also represents a source of unmetabolised antibiotics, assorted antimicrobials and antibiotic resistant bacteria (ARB). In the UK, cattle slurry comprises a substantial proportion of animal waste applied to fields. Consequently, the dispersal of AR on UK dairy farms and their environs warrants investigation.

The body of work presented here therefore aimed to characterise the dynamic spread of AR on a large, high-performance commercial dairy farm in the UK. More specifically, the occurrence and prevalence of AR determinants and bacterial taxa in slurry-amended field soil were evaluated over the course of a year, with sampling occurring in May, July, September and October 2017, as well as January and May 2018. On-site slurry samples were also characterised monthly, between June and October 2017. During the period of study, additional sampling was also carried out in January and May 2018, from a nearby field site with no history of livestock waste application. Environmental DNA was extracted from all samples, processed, and sequenced to produce metagenomic libraries. Bioinformatic analyses were carried out to annotate, quantify, and visualise the ARG profile (resistome) and taxonomic composition of samples.

By bringing together antibiotic usage records, soil physiochemistry, meteorological and metagenomic data, the current work produced a uniquely comprehensive temporal evaluation of AR in field soil following successive amendments of slurry at realistic rates of application.

Through analyses of metagenomic data it was shown that while both slurry and soil contained a diverse array of antibiotic resistance genes (ARGs), the ARG profiles of soil were distinct from those of slurry, irrespective of the origin of the soil in the context of the sites sampled. This is further reflected in the microbial composition of soils which also demonstrated robust correspondence with ARG profiles. Together, these data allowed the identification of slurry biomarkers (ARGs and taxa). ARG slurry biomarkers included gene groups belonging to macrolide, lincosamide and tetracycline resistance gene categories. Furthermore, the wider genetic context of slurry ARGs was characterised by metagenome assembly. Specifically, the beta-lactamase resistance gene *cfxA* was associated with NBU-1-like elements, while the tetracyline resistance gene *tetM* was embedded within Tn*916*-like transposable mobile elements. Similarly, certain ARGs such as *aph(6)* (aminoglycosides)*, aph(3")* (aminoglcosides) and *sul2* (sulfonamides) were co-resident on contigs.

Taxonomic slurry biomarkers included members of phylum Bacteroidetes, Firmicutes, Synergistetes, Tenericutes and Sphaerocheata. In particular, both unassembled read data and metagenomes-assembled genomes (MAGs) indicated *Proteiniphilum* sp. (Bacteroidetes) was a biomarker of slurry application.

Temporal analyses showed that the slurry biomarkers exhibited two distinct lifecycles in soil, one of which appeared to be further modified by season. Firstly, select ARGs were consistently more abundant in soil with a long-term history of slurry application relative to soil from the 'untreated' site (e.g. *tetM*), indicating their prevalence related to long-term enrichment. In contrast, another set of ARGs exhibited transient increases following the first application of slurry in May 2017 (e.g. *lnuB, mefB*). These ARGs declined to pre-treatment levels within eight weeks. In contrast, the same ARGs persisted for >12 weeks after the first application of slurry in February 2018. This suggests the timing of the first slurry application of the season can influence the survival of select slurry associated ARGs and taxa. Given current UK guidelines, the latter scenario may result in the contamination of grass-cut for silage intended for use as cattle-feed and may lead to

a positive feed-back loop of AR in cattle, although this was beyond the scope of the current work. Further field-based and laboratory experiments should be carried out to confirm the broader scalability of these findings.

In summary, the present work explored relationships between taxa, ARGs and mobile genetic elements, as well as the dynamic nature of ARGs within soils on a working dairy farm. Analyses enabled identification of candidate biomarkers of slurry exposure (ARGs and taxa), which can be used to develop more targeted studies in the future.

Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

List of Tables

List of Appendices

List of Supplementary Files

Supplementary file 1 Differentially abundant ARGs by site according to Corncob.

Supplementary file 2 Differentially abundant genera by site according to Corncob (excludes slurry).

Supplementary file 3 Differentially abundant genera according to Corncob for slurry-impacted site soil <24 hours after slurry application and 56 days following slurry application in May 2017.

Supplementary file 4 Differentially abundant genera based on season according to Corncob (untreated site soil).

Supplementary file 5 Differentially abundant genera based on season according to Corncob (slurry tank slurry).

Supplementary file 6 Differentially abundant genera based on season according to Corncob (slurry-impacted site soil).

All Supplementary files are available at the following link: https://uniofnottm-my.sharepoint.com/:f:/g/personal/alexander_williams_nottingham_ac_uk/Ejuz8mx-zZFKlIeSDGeGuvoBFRhnHvQlk9G_dIYlqcBlTQ?e=0S6nby

Acronyms, Abbreviations and Initialisms

AIDS - acquired immune deficiency syndrome

AMR - antimicrobial resistance (encompasses biocides, antimicrobial metals, detergents etc.)

AR - antibiotic resistance

ARB - antibiotic resistant bacteria

ARGs - antibiotic resistance genes

BLAST - basic local alignment tool

DALYs - disability-adjusted life years

DGGE - denaturing gradient gel electrophoresis

DNA - deoxyribonucleic acid

ESKAPE Pathogens - *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii, Pseudomonas aeruginosa*, *Enterobacter* sp.

FIOs - faecal indicator organisms

GDP - gross domestic product

HGT - horizontal gene transfer

HT-qPCR - high throughput qPCR

ICEs - integrative conjugative elements

LSS - liquid solid separation (of animal manure)

MAGs - metagenome assembled genomes

MGEs - mobile genetic elements

MRGs - metal resistance genes

NVZ - nitrate vulnerable zone

ORFs - open reading frames

qPCR - quantitative polymerase chain reaction

RPPs - ribosomal protection protein

Chapter 1
Introduction

1.1 Antibiotic Resistance: A Global Dilemma

The emergence and spread of antibiotic resistance (AR) have become issues of urgent global concern. Over the last decade, AR has been the subject of many national and international reports, all of which underscore our current reliance on antibiotics to deliver effective healthcare services as a primary motivator for action (O'Neill, 2014, WHO, 2014, FAO, 2016, GOV.UK, 2019, ECDC, 2020). Accordingly, the UK government report into antimicrobial resistance (AMR) estimates that if drug resistance continues to develop unabated, as many as 10 million lives a year could be lost to AMR by 2050, resulting in a 7% reduction in global GDP (O'Neill, 2014). It is important to note that while AMR technically encompasses drugs active against members of any microbial domain, including protozoa and fungi; the work conducted by O'Neill (2014) places emphasis on antibiotic resistant bacteria (ARB), which are the focus of the present work. The significance of ARB can be further demonstrated by the creation of the Global Priority Pathogen List (GPPL) by the WHO (2017b). The list was compiled to galvanise the development of novel antibiotic therapeutics targeting the most internationally concerning antibiotic resistant bacteria (WHO, 2017b). Many bacteria listed under 'critical' and 'high' concern on the GPPL overlap with the ESKAPE group (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* and *Enterobacter* spp.), as previously recognised by the Infectious Diseases Society of America for their increasing capacity to 'escape' existing treatment strategies (Rice, 2008, Boucher et al., 2009, Pendleton et al., 2013). Specifically, the ESKAPE pathogens *A. baumanii* and *P. aeruginosa*, as well as carbapenem and 3rd generation cephalosporin-resistant Enterobacteriaceae (including *K. pneumoniae and Enterobacter* spp.) are assigned to the highest GPPL tier of 'critical'

concern, while *Enterococcus faecium* and *Staphylococcus aureus* fall under the second highest tier of 'high' concern (WHO, 2017b).

Extensively drug resistant *Mycobacterium tuberculosis* was not incorporated into the GPPL on the grounds that this organism was already an established pathogen of concern for which new treatments were being actively sought (WHO, 2017b).

The ascent of ARB as a threat to modern healthcare systems can be explained by the ubiquity of bacteria in our shared environment (including our own bodies) and our dependence on antibiotics to prevent or mitigate bacterial infections which might arise as a result of invasive surgery or immunosuppression (e.g. chemotherapy, AIDS). The vital importance of antibiotics to modern healthcare systems is readily demonstrable. For example, an international meta-analysis by Smaill and Grivell (2014) indicated antibiotic prophylaxis could reduce the incidence of maternal infection following caesarean-section by 60-70%. If the efficacy of antibiotic prophylaxis were reduced by only 30%, Teillant et al. (2015) estimated an additional 120,000 surgical site and post-chemotherapy infections would occur in the US annually, of which 6,300 would likely prove fatal.

Although AMR constitutes a clear public healthcare burden, de Kraker et al. (2016) cautioned that the calculations behind the loss of life projected by O'Neill (2014) were poorly defined and were not submitted to independent scientific review prior to publication. The authors go on to highlight several assumptions common to burden estimates which can inflate or otherwise distort the prevalence of AMR and attributable mortality (de Kraker et al., 2016). Similarly, a meta-analysis by Naylor et al. (2018) underscores marked variability in both economic and health-orientated burden estimates. For example, projected annual increases in AMR-related healthcare costs were shown to extend from no significant increase through to $1 billion 2013 USD (Naylor et al., 2018). Nonetheless, there are many reasons to suspect the true cost of AMR is often underestimated and steadily growing.

Specifically, disability-adjusted life years (DALYs) attempt to quantify important aspects of disease burden which are not encapsulated by mortality figures alone. As defined by WHO (2020), DALYs consider both the years of life lost due to premature death and years of full health lost. DALYs would therefore incorporate the impact of prolonged hospitalisation due to AR infections as well as instances where amputation is necessary to control severe AR infections. In a recent study, Cassini et al. (2019) estimated the contribution of antibiotic resistant infections to DALYs across the EU and European Economic Area (EEA) in 2015 was comparable to the combined DALY rate of HIV, tuberculosis and seasonal influenza. Tacconelli and Pezzani (2019) remarked that while such methods still harbour limitations, the work by Cassini et al. (2019) reiterated the potential societal impact of AR in way which few studies to date have made clear.

Finally, concerns surrounding AR have been further compounded as the rate of novel antibiotic development is unable to match emerging healthcare demands (Freire-Moran et al., 2011, Smith and Coast, 2013, WHO, 2019, Butler and Paterson, 2020). The stewardship of existing antibiotics through careful use and optimisation are therefore archetypal of recommendations made by national and international action plans for managing AR (WHO, 2015, O'Neill, 2016, GOV.UK, 2019).

1.2 The Mechanics of AR: An Overview

1.2.1 Acquisition and Dispersal of ARGs

In bacteria, antibiotic resistance genes (ARGs) emerge and spread by hereditary, spontaneous mutation and horizontal gene transfer (HGT) (Levy and Marshall, 2004, Perichon and Courvalin, 2009). Acquisition by HGT encompasses processes where ARGs are disseminated outside the confines of 'vertical', parent-to-progeny lineages (Levy and Marshall, 2004, Perichon and Courvalin, 2009). Forms of HGT include the uptake of naked DNA from the local environment via transformation, intra- and inter-species

exchange of resistance-encoding mobile genetic elements (MGEs) through conjugation (Levy and Marshall, 2004, Frost et al., 2005, Perichon and Courvalin, 2009), and the insertion of ARGs by bacteriophage, known as transduction (Frost et al., 2005, Balcazar, 2014).

The conjugal transfer of MGEs is thought to dominate ARG dispersal (Von Wintersdorff et al., 2016). Studies reviewing the development of AR in pathogens of critical concern within healthcare settings (e.g. ESKAPE group) further underpin the significance of MGEs (Hegstad et al., 2010, Partridge et al., 2018, De Oliveira et al., 2020). Moreover, recent research in China has empirically demonstrated links between MGEs and ARGs over large geographic areas, including aquatic and terrestrial environments (Yao et al., 2020, Zhang et al., 2020). Considering their importance in the dissemination of ARGs, MGEs will be briefly explored in more detail.

There is a plethora of terms used to categorise MGEs with varying degrees of specificity, and their ability to interact with one another can lead to mosaic structures which defy simple categorisation (Juhas et al., 2009, Bellanger et al., 2014, Johnson and Grossman, 2015, Partridge et al., 2018).

However, one way in which MGEs can be broadly classified is by the extent to which they encode their own inter-cellular mobility. For example, ICEs (integrative conjugative elements), sometimes referred to as conjugative transposons, are self-transmissible by conjugation and are capable of integrating into the host genome or other MGEs such as plasmids or even other ICEs (Bellanger et al., 2014, Johnson and Grossman, 2015, Partridge et al., 2018). Although conjugative plasmids are also self-transmissible, they differ from ICEs in that they are extrachromosomal (Partridge et al., 2018, Botelho and Schulenburg, 2020). Consequently, while both ICEs and conjugal plasmids can be disseminated through vertical transfer to daughter cells and horizontal transfer, plasmids may be exposed to potential loss by segregation during cellular division unless their within-cell copy number is properly maintained (Partridge et al., 2018, Botelho and

Schulenburg, 2020). Conjugative plasmids and ICEs have been described as the major facilitators of conjugal HGT (Koraimann, 2018, Botelho and Schulenburg, 2020).

ICEs have been implicated in the spread of tetracycline and vancomycin resistance in Enterococci (Hegstad et al., 2010, Roberts and Mullany, 2011) as well as carbapenem resistance in *Pseudomonas aeruginosa* (Ding et al., 2018). Meanwhile, conjugative plasmids have been shown to harbour carbapenem and beta-lactam-resistances including *Klebsiella pneumoniae* carbapenamases (KPCs) (Navon-Venezia et al., 2006, Wei et al., 2007, Dang et al., 2020) and extended spectrum beta-lactamases (ESBLs) (Freitag et al., 2017, Poidevin et al., 2018, Liu et al., 2019a).

Other MGEs such as gene cassettes and transposable elements (transposons and insertion sequences) lack the genetic modules necessary for inter-cellular transport, however they can exhibit intra-cellular mobility (Johnson and Grossman, 2015). Integrons (particularly class 1) can accrue a single ARG (gene cassette) or amass several in tandem to form gene cassette arrays (Partridge et al., 2009, Domingues et al., 2012, Partridge et al., 2018). Integrons can therefore function as repositories for multiple ARGs. While integrons themselves are not intrinsically mobile, they can be embedded within transposable elements and plasmids (Holmes et al., 2003, Martínez et al., 2007, Domingues et al., 2012, Gillings, 2014). In this regard, genetic elements with limited mobility can 'piggy-back' within those readily capable of either intra- or inter-cellular transfer (Bellanger et al., 2014, Johnson and Grossman, 2015).

Lastly, it should be noted that MGEs do not exclusively shuttle ARGs; they host any number of genes with various functions (Frost et al., 2005, Juhas et al., 2009, Rankin et al., 2011), some of which, such as virulence factors are also of significant concern in their own right, especially when co-resident with ARGs.

1.2.2 AR Mechanisms and Their Selection

Having summarised the manner in which ARGs can be acquired and propagated, the mechanisms of AR remain to be discussed. Antibiotics are diverse in their specificity of antimicrobial activity and mode of action, however microorganisms have evolved equally diverse mechanisms to negate their effects (Table 1.1). Principally, these mechanisms include antibiotic inactivation by alteration or degradation, antibiotic target modification (typically due to chromosomal mutation) and the ejection of antibiotics by efflux (Levy and Marshall, 2004, Madigan et al., 2006). Examples of ARGs associated with antibiotic inactivation include the widely distributed and diverse group of CTX-M ESBLs, which confer hydrolytic activity against the third-generation cephalosporin cefotaxime (Cantón et al., 2012). Meanwhile resistance to quinolone antibiotics can be mediated by mutations in the target enzymes (Shenagari et al., 2018, Ostrer et al., 2019) or through expression/over-expression of efflux pumps (Pérez-Varela et al., 2018, Azargun et al., 2020). It is therefore evident that the activity of an antibiotic class or even an individual antibiotic may be circumvented by a range of different mechanisms and each mechanism may be represented by an equally diverse array of ARGs (Levy and Marshall, 2004). On the other hand, the glycopeptide antibiotic vancomycin does not enter bacterial cells (negating efflux pumps), nor does it target proteins as many other antibiotics do; accordingly, very specific mechanisms for resistance in the form of *van* genes have worryingly come to the fore in clinical Enterococci and Staphylococci which disrupt the binding of vancomycin to the peptidoglycan membrane of Gram-positive bacteria (Reygaert, 2018, Stogios and Savchenko, 2020). While covering the full gamut of ARGs and the mechanisms by which they confer AR is beyond the practical scope of this introduction, mechanisms associated with specific ARGs of interest will be described in further detail where relevant in subsequent chapters.

Table 1.1 Summary of major antibiotic categories, including subgroups, their mode of action and spectrum of activity. Examples of bacterial resistance mechanisms are also indicated. Based on Levy and Marshall (2004), Madigan et al. (2012) and Roberts et al. (2012).

Antibiotic class	Antibiotic/ Antibiotic subgroup	Mode of action (target)	Principal spectrum of activity	Example Mechanisms of Resistance (class level)
Glycopeptides	Vancomycin	Cell wall synthesis	Gram +ves	Modification of drug target (modified peptidoglycan pre-cursors)
β-lactams	Penicillins		Gram +ves, select Gram -ves	Enzymatic inactivation (beta-lactamases and extended-spectrum beta-lactamases)
	Cephalosporins		Gram +ves, select Gram -ves	
	Monobactams		Gram –ves	
	Carbapenems		Gram +ves and Gram-ves	
Quinolones	Nalidixic acid	DNA gyrase (DNA supercoiling)	Gram +ves and Gram -ves	Efflux, alteration of target enzyme
	Ciprofloxacin		Gram +ves and Gram -ves	
Sulfonamide /anti-folate combination	Sulfamethoxazole-Trimethoprim	Folic acid metabolism	Gram +ves and Gram -ves	Alteration of drug target
Macrolides	Erythromycin	Protein synthesis (50S subunit inhibitors)	Gram +ves	Efflux, inactivating enzyme, alteration of drug target (MLS)
Lincosamides	Lincomycin		Gram +ves	
Streptogramins	Streptogramin B		Gram +ves	
Phenicols	Chloramphenicol		Gram +ves and Gram -ves	Efflux, inactivating enzyme, alteration of drug target
Tetracyclines	Oxytetracycline	Protein synthesis (30S subunit)	Gram +ves, select Gram -ves	Efflux, enzymatic inactivation, ribosomal protection
Aminoglycosides	Streptomycin		Gram -ves and *Mycobacteria* spp.	Inactivating enzyme, ribosomal methylation

It is also important to acknowledge some groups of bacteria are naturally resistant to particular antibiotics; for instance, the outer membrane of Gram-negative bacteria renders most impermeable to penicillin (Madigan et al., 2006, Cox and Wright, 2013). Likewise, antibiotic producing bacteria necessarily possess mechanisms to protect them from their own antibiotic arsenal (Peterson and Kaur, 2018). However, Jayaraman (2009) also highlights many ARGs are likely to serve a multitude of functions which are not strictly limited to antibiotic protection, a point illustrated by the export of virulence-factors by select multidrug-resistance (MDR) efflux pumps (Piddock, 2006, Alcalde-Rico et al., 2016). Likewise, a review by Okada and Seyedsayamdost (2017) suggests antibiotics themselves may serve as signalling molecules depending on ecological context.

Selection for antibiotic resistance has occurred for millennia. Studies have identified ARGs redolent of those found in clinical isolates in deep soil cores pre-dating widespread antibiotic use by thousands of years (D'Costa et al., 2011), prairies with no documented history of anthropogenic antibiotic contamination (Durso et al., 2016), and in the gut microbiomes of remote, previously un-contacted indigenous tribes people who have not received modern antibiotic treatment (Clemente et al., 2015). Likewise, Paun et al. (2021) revealed pan-drug-resistant Pseudomonad strains recovered from 13000 year-old cave ice also possessed marked antibacterial activity against clinically relevant pathogens. The possibility that at least some antibiotics and AR mechanisms arose for purposes other than microbial chemical warfare (Jayaraman, 2009), may go some way to explaining their extensive evolutionary prehistory, however it is the strength of more recent anthropogenic antibiotic selection which is of crucial concern to many (Levy and Marshall, 2004, Davies and Davies, 2010, Laxminarayan et al., 2013, Ventola, 2015, Aslam et al., 2018).

Indeed, Levy and Marshall (2004) asserted that it is the collective exposure of humans (and other animals) to antibiotic treatments which drive the selection and eventual dominance of ARB over susceptible bacteria at the population level. However, it is equally important to acknowledge that the use of antibiotics can exert considerable selective pressure within individuals. In particular, consider the empirical use of broad-spectrum antibiotics to treat ICU patients for whom a delay in antibiotic treatment could undermine survival (Karam et al., 2016). The use of broad-spectrum rather than narrow-spectrum antibiotics may increase the risk of selecting for multidrug resistance (Fjalstad et al., 2018) and also enrich AR traits in non-target organisms; a process known as bystander-selection (Tedijanto et al., 2018). More generally, antibiotic treatment has also been shown to stimulate HGT in the human gut (Li et al., 2019). Alternatively, sub-clinical exposure to antibiotics also has the potential to contribute to AR (Gullberg et al., 2011, Sandegren, 2014). It is therefore thought that the replacement of selected AR phenotypes in a population with susceptible equivalents is a slow process, even when

exposure to the original selecting antibiotic is greatly reduced (Levy and Marshall, 2004, Andersson and Hughes, 2011). Lastly, non-antibiotic agents can also indirectly select for AR via co- and cross-selection (Baker-Austin et al., 2006, Wales and Davies, 2015, Davies and Wales, 2019). Briefly, co-selection refers to the acquisition of ARGs because they are located on the same genetic element as functionally unrelated genes which are subject to selection, meanwhile cross-selection describes a scenario where AR phenotypes (such as non-specific efflux pumps) also provide protection from non-antibiotic compounds (Baker-Austin et al., 2006). Nevertheless, while there are many ways in which AR can be selected, Waglechner and Wright (2017) stress all bacteria are not resistant to all antibiotics, highlighting many barriers to AR must also exist. Ultimately, the selection of AR within the wider environment represents complex interplay between natural and anthropogenic factors which research is only beginning to elucidate mechanistically (Singer et al., 2016, Tiedje et al., 2019).

In summary, the majority of national and international reports acknowledge that while AR can arise naturally within the environment, the accelerated development of AR is predominantly attributed to the overuse, misuse and abuse of antibiotics in medicine and to a varying degree, agriculture (FAO, 2016; O'Neill Report, 2016).

1.3 The Role of Agriculture in AR

1.3.1 Antibiotic Use in Agriculture

It is often asserted that extensive use of antibiotics in agriculture can act as a significant selective pressure for the genesis and maintenance of resistance in agricultural environments, with potential impact in clinical settings though contamination of the food chain (Soulsby, 2007, Aarestrup et al., 2008, Marshall and Levy, 2011, Meek et al., 2015, Collignon and McEwen, 2019) and adjoining environmental compartments such as watercourses (Singer et al., 2016, Qiao et al., 2018). These concerns are enshrined with

the concept of 'One Health', which aims to imbue stakeholders with an awareness of the interactions between humans, other animals and the wider environment (McEwen and Collignon, 2018, GOV.UK, 2019, Hernando-Amado et al., 2019, Tiedje et al., 2019).

A substantial quantity of global antibiotic consumption corresponds with antibiotic growth promotion, a practice which arose following evidence in the 1940s that the supplementation of animal feed with subtherapeutic levels of antibiotics increases yield regardless of overt animal health (reviewed by Dibner and Richards, 2005). Until recently, antibiotic growth promotion was widely adopted in both the US (FDA, 2019) and China (Hu and Cowling, 2020), however, establishing the extent of historic antibiotic growth promotion in China remains problematic due to data paucity (Collignon and Voss, 2015, Krishnasamy et al., 2015). Regardless, it is thought the emergence and global dissemination of plasmid-mediated colistin resistance genes (*mcr-1*) in swine and humans (Liu *et al.,* 2016) relates to such agricultural praxis. Critically, colistin is a 'last line' or 'last resort' polymyxin antibiotic (Kaye et al., 2016) for which resistance was previously thought to be restricted to chromosomal mutation (Liu et al., 2016). Moreover, a subsequent correspondence report identified *mcr-1* in *Escherichia coli* isolated from Chinese poultry during the 1980s, around the time when colistin was first utilised in agriculture; however, the recovery of *mcr-1*-positive isolates was shown to be relatively stable until 2009, after which there was a year-on-year increase in the proportion of positive isolates (Shen et al., 2016). This suggests that it may take several years of nascent circulation before an emergent ARG reaches critical mass and spreads rapidly. Lastly, studies from around the world have also established positive correlations between veterinary antibiotic sales and AR on farms (Asai et al., 2005, Chantziaras et al., 2014).

There are others however, who believe the evidence regarding the role of agricultural antibiotic use is less conclusive. For instance, Singer and Williams-Nguyen (2014) contended that while it is clear that agricultural facilities can provide an environment conducive to the development of AR, studies which identify resistant isolates from farm

environments where antibiotics are used demonstrate correlation as opposed to causation. They go on to state that there are many other non-antibiotic compounds which may contribute to AR. For example, biocides and heavy metals, which are routinely encountered in husbandry practice, have been identified as potential co-selective agents for resistance (Baker-Austin et al., 2006, Wales and Davies, 2015, Zhou et al., 2016, Davies and Wales, 2019). Research also raises the possibility that these co-selective agents can select for AR even more strongly than antibiotics. For instance, Song et al. (2017) found that soil microcosms spiked with copper and zinc yielded significant increases in overall bacterial community tolerance to tetracycline and produced a toxic effect on bacterial growth. In contrast, soils spiked with tetracycline (in excess of concentrations realistically encountered in the environment) had limited impact on bacterial community tolerance (Song et al., 2017). However, it is important to appreciate that the behaviour of antibiotic residues in soil may differ markedly depending on soil type (Tasho and Cho, 2016). Another study, this time focusing on the abundance of antibiotic resistance genes (ARGs) in soils amended with cattle manure, found that ARGs increased in soils irrespective of whether the manure was from antibiotic-treated or antibiotic-free animals (Kyselková et al., 2013). This finding was further validated by an additional study where faeces spiked with high and low doses of chlortetracycline yielded similar ARG levels in amended soil; the authors therefore concluded that ARGs can accrue within animal waste-amended soils independent of this antibiotic selective pressure (Kyselková et al., 2015b). Schmitt et al. (2006) also found a diverse array of ARGs in manure regardless of whether intensive or restricted antibiotic treatment regimens were in place. There is therefore ample evidence (as discussed in *AR mechanisms and selection*), that AR is both a natural phenomenon existing in the absence of anthropogenic input and that antibiotic exposure is not the sole means by which anthropogenic activity can select for AR. It is therefore unsurprising that the significance of controlling antibiotic use in agriculture remains an issue of contention for some groups, as outlined by Hoelzer et al. (2017). Indeed, Ghosh and LaPara (2007) found little to distinguish between the AR patterns of soil collected from 10 farms

employing either therapeutic or subtherapeutic antibiotic treatment regimens. However, the aforementioned study relied exclusively on culture techniques to enumerate resistant bacteria and only tetracycline resistance genes were considered. On the other hand, many of the reviews highlighting insufficient causal evidence linking antibiotic use in agriculture with impacts on human health (Phillips et al., 2004, Singer and Williams-Nguyen, 2014, Chang et al., 2015) pre-date the discovery of mobilised resistance to colistin within Chinese swine farms in 2016.

Despite remaining equivocations, there have been concerted efforts to minimise global antibiotic use in agriculture. After several years of phased removal, an EU-wide ban in 2006 prohibited all use of antibiotics for growth promotion in livestock (EC, 2003). More recently, the U.S. Food and Drug Administration (FDA) issued recommendations to remove 'growth promotion' from the listed applications of antibiotics and revoke 'over-the-counter' access to medically important antimicrobials in favour of prescription under licensed veterinarians (FDA, 2013). Indeed, the FDA (2019) indicated a 26% reduction in the overall use of medically important antimicrobials in food-producing animals between 2016-2019; however a 3% increase (largely due to swine production) was reported between 2018-19. Meanwhile China has launched its own action plan to reduce antibiotic use in clinical and agricultural settings (People's Republic of China, 2017, Hu and Cowling, 2020). However, the global consumption of antibiotics by food animals is expected to increase by 11.5% between 2017-2030, signalling that the drivers behind antibiotic use in these sectors may be beyond the immediate control of many nations or their desire/ability to enforce regulations.

1.3.2 Animal Waste as a Reservoir of AR and Associated Selection Pressures

The UK government has expedited a 43% reduction in the total volume of antibiotics sold for animal use between 2015-2019 (UK-VARSS, 2020) through improved animal husbandry and welfare initiatives (e.g. appropriate stocking density, improved nutrition,

and disease monitoring) (Evans and Border, 2018). However, managing the fate of administered antibiotics in the environment should also be considered in plans to address AR. Specifically, animal waste is a potential reservoir for many described resistance factors, including excreted antibiotics (Tasho and Cho, 2016), which can represent as much as 10-90% of the original dose administered to livestock (reviewed in Kumar et al., 2005b). It therefore follows that ARB exhibiting phenotypic resistances to veterinary antibiotics have been recovered from animal waste (Asai et al., 2005, Ibrahim et al., 2016). Likewise, livestock waste is often replete in ARGs and MGEs (Durso et al., 2011, Ma et al., 2016, Lima et al., 2020, Yang et al., 2020). On the other hand, current methods for remediating antibiotic residues and ARGs in animal waste are unlikely to be completely effective (Gros et al., 2019). As a result, land application of livestock waste may facilitate the spread of AR, with the soil-waste interface acting as a hotspot for transfer of ARGs from manure-derived bacteria to indigenous soil microbes (Heuer et al., 2011, He et al., 2020).

Indeed, Andrews Jr et al. (2004) conducted a microcosm experiment which suggested the transfer of *tn916* ICE from slurry-borne bacteria to autochthonous soil bacteria, while Musovic et al. (2014) demonstrated long-term application of manure could promote plasmid transfer. Studies have also demonstrated the application of animal waste can enrich ARG concentrations in soils (Sengeløv et al., 2003, Byrne-Bailey et al., 2009, Heuer et al., 2011, McKinney et al., 2018, Dungan et al., 2019, Zhao et al., 2019). However, waste from animals receiving negligible or no antibiotic treatment has also been shown to elevate ARGs in amended soils, highlighting the importance of non-antibiotic treatment-related factors (Udikovic-Kolic et al., 2014, Kyselková et al., 2015b, Hu et al., 2016).

Another layer of complexity is introduced when considering the processing of animal waste on farms, whereby waste can pass through several distinct environmental compartments before being applied to fields for the purpose of fertilisation (Figure 1.1). These compartments possess a unique set of conditions, which may promote or hinder

AR development, and can exhibit varying degrees of spatiotemporal dynamism. Many variables have been identified as factors influencing the survival of manure-derived bacteria, the abundance of ARGs and/or putative (co) selective pressures. These factors can include: the duration of waste storage (Joy et al., 2014, Baker et al., 2016, Muurinen et al., 2017), the texture (Srinivasan and Sarmah, 2014, Blau et al., 2018) and moisture (Wang et al,. 2004) of fields receiving slurry, and the method by which slurry is applied (Hutchison et al., 2004, Hodgson et al., 2016). The slurry application method can also impact the transport of ARGs in runoff, as reported by Joy et al. (2013). Meanwhile, liquid-solid separation (LSS) of farmyard animal waste, a common waste-management practice, has been shown to promote the development of distinct bacterial communities in the different fractions (Pandey et al., 2018).

In order to devise effective mitigation strategies, it is therefore essential that agricultural sources of AR, with special reference to animal waste, are studied within the context of the entire farm environment and not merely a single factor, such as antibiotic use.

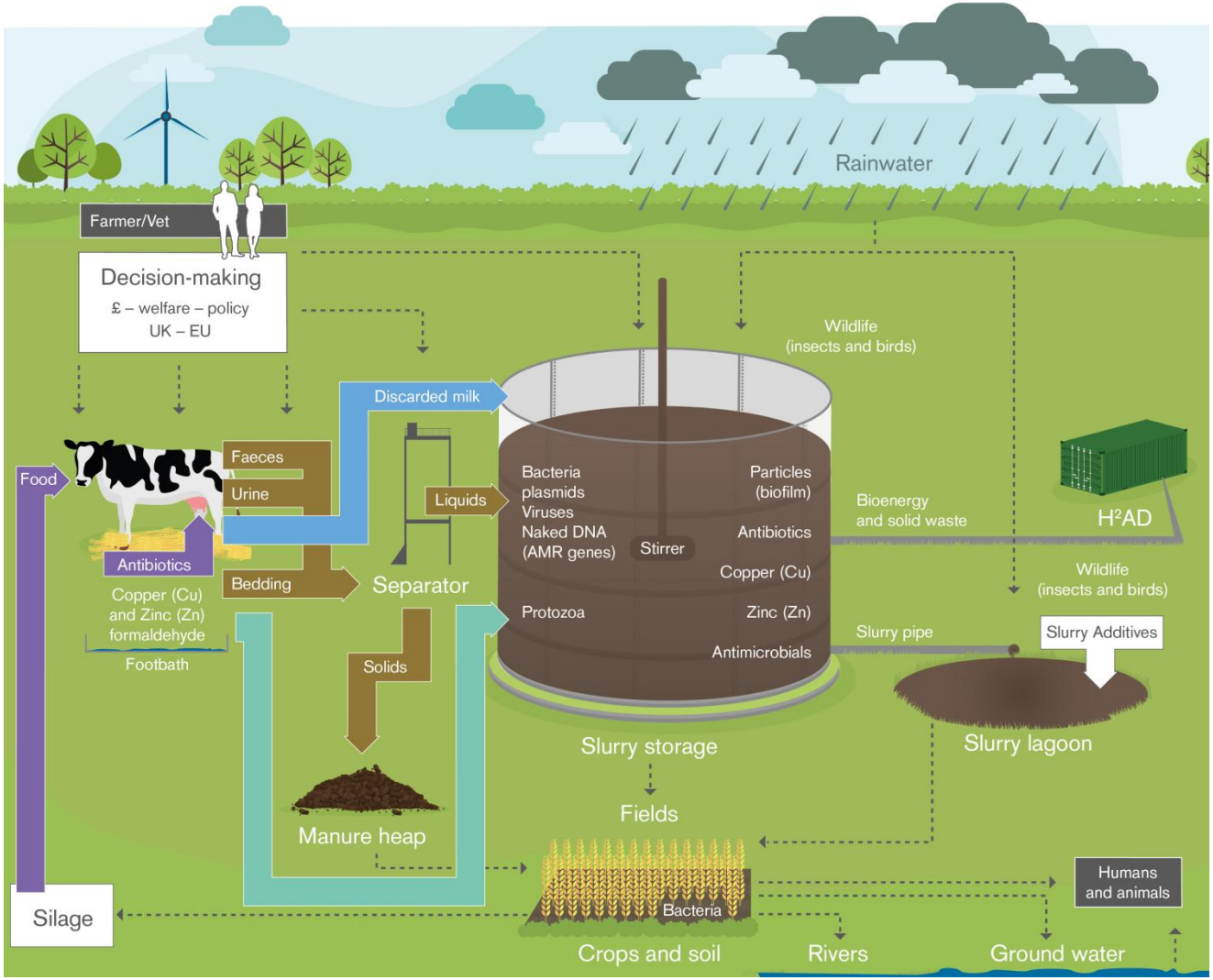


Figure 1.1 Infographic showing the various environmental compartments (slurry tank, slurry lagoon, manure heap and fields) animal waste enters while on a commercial dairy farm. Abiotic and biotic factors which may influence antibiotic resistance in the slurry tank are also highlighted (infographic courtesy of Dov Stekel).

1.4 Dairy Farms and AR

1.4.1 Contextual Significance of UK Dairy Farms

Studies into AR in agriculture frequently centre on livestock operations, which have become routinely associated with intensive farming and high antibiotic use; namely swine and poultry. Although dairy farms typically use fewer antibiotics, they are of interest for several reasons.

Firstly, a number of studies have highlighted that the minimum selective concentration (MSC) for certain antibiotics is substantially less than that of the minimum inhibitory concentration (MIC) (Gullberg et al., 2011, Andersson and Hughes, 2012, Sandegren, 2014); suggesting that it is possible for AR to arise in environments due to limited, but persistent antibiotic contamination, as might be the case on dairy farms.

The expected productive lifespan of cattle on dairy farms far exceeds that of both swine and poultry (approximately 4 years, 5-6 months and 5-7 weeks respectively) (FTP, 2020). As such, while the dairy industry as a whole consumes smaller quantities of antibiotics than swine or poultry operations, dairy cattle are likely to receive antibiotics over a longer timeframe. Their longer farm lifespan may also provide extended opportunity for AR to develop within their gut microbiomes in response to antibiotic dosage or increase the probability of acquiring naturally resistant pathogens from the immediate environment. Furthermore, once an animal presents ARB it may continue to shed these in faeces until successful treatment or removal from the population. As animal waste is an established reservoir of AR, it is noteworthy that dairy cattle are estimated to comprise 80% (~67Mt) of annual UK animal waste production (Smith and Williams, 2016), with more cattle waste applied to fields in England and Wales than swine and poultry combined (DEFRA, 2016). This is astonishing as swine numbers greatly outnumber dairy cattle (DEFRA, 2016).

According to survey estimates, the number of UK dairy farms has more than halved (-55.14%) between 2000 and 2018, while the average size of dairy cattle herds nearly doubled during the same period (+45.59%) (AHDB, 2020a, AHDB, 2020b).

The largest UK herds can exceed 900 (Hanks and Kossaibati, 2019). This appears to reflect the general trend that many smaller dairy farms are being replaced by fewer, larger dairy farming operations, likely in response to intensifying competition and overhead costs. The continued development of large scale, intensive farming in the dairy sector could have implications for AR. For instance, larger herds producing greater volumes of waste may be more likely to contaminate the local environment with waste-borne resistance determinants. Furthermore, it stands to reason that larger farms may hire more employees, potentially increasing contact events between AR-carrying cattle and different members of staff, facilitating resistance gene dissemination. Alternatively, larger dairy operations may increase automation thereby reducing human-animal interaction. Nonetheless, a number of studies have also shown that disease prevalence and/or persistence can increase with herd size (Bartlett et al., 1992, Brooks-Pollock and Keeling, 2009), which may exacerbate AR indirectly through greater antibiotic and biocide use. However, in US dairy farms, Hill et al. (2009) reported that within-herd disease prevalence tended to decrease with increasing herd size. This may correspond to larger farms being able to make greater investment in prevention and mitigation strategies (Hoe and Ruegg, 2006, Haskell et al., 2009). However, Hill et al. (2009) also found that larger farms used broad-spectrum antibiotics more frequently than smaller farms, which could potentially aid selection for resistance in a wider array of bacteria.

Mastitis is one of the most common afflictions affecting dairy cattle (Bradley, 2002, Hillerton and Berry, 2005), and the second most costly disease to the UK dairy industry (CHAWG, 2020). Accordingly, cases of mastitis make a substantial contribution to antibiotic use in the EU (De Briyne et al., 2014). Treatment predominantly involves beta-lactam antibiotics such as penicillin, 1st and 2nd generation cephalosporins and to a lesser extent, antibiotics considered of critical importance to human medicine (De Briyne et al., 2014). These include 3rd and 4th generation cephalosporins and macrolide antibiotics (Brunton et al., 2012, De Briyne et al., 2014). The aetiology of mastitis is often associated with both Gram-negative and Gram-positive bacteria, namely *Escherichia coli* and *Streptococcus uberis* (Bradley et al., 2007). However, over 100 possible causal agents have been documented (Watts, 1988). Evidence of increasing AR in mastitis pathogens as a result of antibiotic use on dairy farms remains elusive (Oliver and Murinda, 2012). Nonetheless, a US study demonstrated the recovery of mastitis pathogens resistant to beta-lactam antibiotics declined following the transition of one dairy farm from conventional to organic antibiotic usage practices (Park et al., 2012). Although growth promotion is presently prohibited in the UK, antibiotic dry cow therapy remains permissible (Biggs et al., 2016, Higgins et al., 2017). Antibiotic dry cow therapy relates to the practice of administering groups of ostensibly healthy, non-lactating cows with preventative antibiotic treatment for mastitis during periods of heightened disease risk. Dry cow therapy can involve a variety of beta-lactam antibiotics ranging from simple penicillin to 4th generation cephalosporins (see Cephaguard DC, Orbenin Dry Cow, Orbenin Extra Dry Cow and Ultrapen LA in Table 2.1). However, following calls to adopt more selective use of antibiotics during the drying-off period (Breen et al., 2014, Biggs et al., 2016, Higgins et al., 2017) the dairy industry appears to be discouraging this practice (CHAWG, 2020).

Meanwhile, antimicrobial metals (e.g. copper, zinc) and biocides, which have the potential to co-select for antibiotic resistance (Pal et al., 2015, Davies and Wales, 2019) are used to prevent and treat lameness (Griffiths et al., 2018) which is the most costly disease in UK dairy farms (CHAWG, 2020). These agents can contaminate livestock

stalls, and may even be directly disposed of in slurry tanks. Consequently, UK dairy farm environments may be consistently exposed to various antibiotic (co) selective pressures.

1.4.2 AR in Cattle Waste-amended Soils

Many studies concerning antibiotic resistance in soils and crops following animal waste amendment have focussed on swine and poultry; the key aspects and findings of investigations involving dairy cattle waste will now be discussed.

A number of studies have demonstrated a transient increase in the concentration of ARGs and/or resistance-associated MGEs in soil amended with cattle waste (Fahrenfeld et al., 2014, Hu et al., 2016, Nõlvak et al., 2016, Muurinen et al., 2017, Macedo et al., 2020). The duration of elevated ARG concentrations in soils post-amendment is variable. For instance, although some studies suggest concentrations approach or a return to background levels within 2 months (Fahrenfeld et al., 2014, Muurinen et al., 2017), others demonstrate elevation in ARGs can persist for >5 months (Hu et al., 2016, Nõlvak et al., 2016). In contrast, another study found the long-term application of cattle manure had minimal impact on the ARG profile of soil when compared to other livestock waste (Peng et al., 2017). On the other hand, a study by McKinney et al. (2018) demonstrated a significant increase in tetracycline and sulphonamide ARGs within soil following application of dairy manure. There is also evidence to suggest that different ARGs do not exhibit the same decay rates (Fahrenfeld et al., 2014, Kyselková et al., 2015b, Sandberg and LaPara, 2016, Lin et al., 2019); however, their behaviour is not always consistent across studies (Kyselková et al., 2013, Fahrenfeld et al., 2014, Macedo et al., 2020). The large range of environmental conditions represented in the literature and variation in experimental design (e.g., field or microcosm scale) as well as the lack of a standardised definition of what constitutes the baseline abundance of resistance determinants in the environment (Rothrock Jr et al., 2016) may explain the lack of congruence to some degree.

Site-specific management practice can also govern environmental exposure to antibiotic resistance determinants. For example, the duration of manure storage has been linked to ARG concentrations in models and site studies (Baker et al., 2016, Ruuskanen et al., 2016, Muurinen et al., 2017). The latter study indicates some ARGs can accumulate in waste over extended storage periods (Muurinen *et al.*, 2017). However, few studies have assessed changes in resistome structure over time *in situ*. One such study by Hurst et al. (2019) documented seasonal shifts in cattle manure-associated ARGs, thereby alluding to possible interaction effects between the length of storage and when storage occurs.

Similarly, application method may play a role in ARG and ARB survival. Joy et al. (2013) identified that broadcast application of swine slurry leads to a greater prevalence of ARGs in run-off than incorporation methods at a single field site. Meanwhile Hodgson et al. (2016) showed the survivability of faecal bacteria in soil following shallow injection of diary slurry was increased compared to broadcast application, although ARGs abundances were not evaluated. Given that HGT of ARGs from exogenous bacteria in dairy waste to indigenous soil bacteria is thought to contribute to the maintenance of ARGs in soil (Heuer et al., 2011), the prolonged survival of introduced organisms is of relevance as it would afford greater opportunity for transfer events to occur. Indeed, although Peng et al. (2017) found that long-term animal waste failed to generate significant changes in pathogen community structure, manures were broadcast rather than injected.

Further research is needed to clarify the role of soil physiochemical properties in determining the fate of antibiotic resistance determinants in applied waste. For example, cattle waste can affect soil pH, while the abundance of sulfonamide resistance genes and proxies for horizontal gene transfer (HGT) such as *intl1* and *intl2* have been correlated with specific pH levels in slurry-amended soils (Nõlvak et al., 2016). The association of soil properties with proxies for HGT is particularly relevant as the authors of a study modelling AMR in slurry tanks cite the rate of HGT as a significant driver of AMR (Baker et al., 2016). Similarly, soil texture has been shown to impact both faecal coliform survival (Cools et al., 2001) and post-amendment ARG concentrations (Blau et al., 2018,

Zhang et al., 2018a). Likewise, Guron et al. (2019) reported soil texture affected the composition and abundance of ARGs detected on vegetables grown in field plots fertilised with cattle manure.

Another vital facet to consider is the means by which the occurrence of AR is estimated in soil receiving cattle-derived waste. This has been borne out by a recent study which sought to compare and contrast the most popular methods for characterising antibiotic resistance measures in dairy waste-amended soil. Specifically, these included shotgun metagenomics, qPCR and selective culture (Wind et al., 2020). Molecular techniques tend to dominate this area of research, primarily because <1% of environmental prokaryotic species are thought to be cultivatable (Rastogi and Sani, 2011), and clinically significant AR phenotypes may be present within a small proportion of the total bacterial population. The latter can be circumvented by antibiotic selection, but this raises its own quandaries and opportunity for bias (e.g., which media and selecting antibiotic to employ). On the other hand, metagenomic and PCR-based techniques offer the opportunity to access the genetic material of a much greater section of bacterial communities (total environmental DNA, in theory). On the other hand, the genotypic resistance traits identified by qPCR and metagenomics may not confer phenotypic resistance. Likewise, not all phenotypes may possess fully described reference genotypes (Davis et al., 2011). Additionally, molecular techniques rarely distinguish between the genetic material of living and dead organisms (Carini et al., 2016), however when considering the possibility of transformation (uptake of extracellular DNA), this material could remain relevant to the study of environmental ARGs.

While several studies have assessed cattle-manure impacted soil using qPCR (Nõlvak et al., 2016, Muurinen et al., 2017, Zhang et al., 2017b, McKinney et al., 2018, Chen et al., 2019b), the adoption of a metagenomic approach can enable the genetic contextualisation of ARGs and provide a more robust association between ARGs and host organisms. This added context can be particularly useful when assessing the mobility of specific ARGs (Slizovskiy et al., 2020) and developing mitigation strategies to pinpoint emergent bacteria of concern. Finally, although qPCR can be more sensitive (reviewed in

Waseem et al., 2019), metagenomic techniques do not require the assembly of target-limiting gene arrays. It is therefore unsurprising that Wind et al. (2020) recommended a combined approach to determining AR in the soil environment. It is noteworthy that such multi-technique studies remain rare.

1.5 Research Aim and Objectives

Aim: The overall aim of this project was to explore the composition and temporal succession of antibiotic resistance determinants and bacterial taxa in soil and slurry associated with a commercial UK dairy farm over the course of a year.

By combining a field sampling campaign with a metagenomics approach, the main aim was addressed by the following ***specific objectives***:

1. Quantify the ARGs present in the farm slurry tank, in field soil amended with slurry, and in 'clean' field soil with no history of slurry amendment over a year. Chapter 2 addresses this objective.
2. To determine if slurry application and season resulted in taxonomic shifts in the field soils and slurry tank. This objective is the focus of Chapter 3.
3. To evaluate associations between the ARGs, bacterial taxa and the mobilome. Chapter 4 addresses this objective.
4. To attempt to quantify how the studied facets of the resistome impact the risk posed by AR in the dairy farm environment. This objective is addressed in Chapter 4.
5. To consolidate the data and examine the findings of the work in the context of government policy. This is discussed in Chapter 5.

Chapter 2
Evaluating ARGs in Soil and Slurry Metagenomes

2.1 Introduction

The extent to which the application of animal waste to agricultural soil can enhance the soil resistome has been subject to intensive research and numerous reviews (Heuer et al., 2011, Williams-Nguyen et al., 2016, Xie et al., 2018b, Tyrrell et al., 2019). Typically, the addition of animal waste coincides with the enrichment of antibiotic resistance genes (ARGs) in soil (Heuer et al., 2011, Sandberg and LaPara, 2016, Zhang et al., 2017b, McKinney et al., 2018). However, only transient increases in entrained ARGs are generally observed, with most studies indicating that introduced ARGs fail to establish in the environment under normal agricultural application rates (Sengeløv et al., 2003, Ghosh and LaPara, 2007, Hu et al., 2016, Liu et al., 2017, Muurinen et al., 2017, Chen et al., 2019a, Cheng et al., 2019). On the other hand, studies have also reported that long-term exposure to animal waste can result in the accumulation of associated ARGs in soil (Peng et al., 2017, Xie et al., 2018a, Dungan et al., 2019, Lu et al., 2020).

Comparative studies have demonstrated that the impact of swine and poultry waste on the soil resistome is greater than that of cattle waste and dissipates more slowly (Sandberg and LaPara, 2016, Peng et al., 2017, Zhang et al., 2017b). In addition , swine and poultry waste has been shown to harbour a more diverse and populous resistome than dairy cattle waste (Wang et al., 2016a, Peng et al., 2017, Zhang et al., 2017b, Qian et al., 2018, He et al., 2020). Whether these trends correspond to antibiotic practice or other factors such as species-specific physiology is still unclear (Sandberg and LaPara, 2016, He et al., 2020).

Although cattle waste may represent a modest pool of ARGs relative to swine and poultry waste, the impact of its use as fertiliser should not be dismissed. For example, Wichmann et al. (2014) used functional metagenomics to demonstrate that cattle

manure resistomes encompass a wide range of antibiotic resistance determinants, including novel chloramphenicol ARGs. In addition, Hu et al. (2016) found that while dairy cattle waste-derived ARGs were transient, amendment resulted in the elevation of resident soil beta-lactamases for at least four months, thus highlighting that the composition of the introduced resistome is not the sole determinant of risk to the environment and human health. Lastly, cattle waste comprises the greatest proportion of animal waste applied to land in England and Wales (DEFRA, 2016).

Having highlighted the potential significance of dairy waste land-application to AMR in the UK, it is therefore surprising that there do not appear to be any UK-based temporal studies evaluating the impact of repeated dairy waste application to field soil. There are however, a number of studies investigating antibiotic usage (Brunton et al., 2012, Jones et al., 2015, Hyde et al., 2017) and the occurrence of antibiotic resistant isolates in UK dairy systems (Piddock et al., 2000, Scott et al., 2009, Ibrahim et al., 2016). While these areas are important from a husbandry perspective, they do not assess a major route by which dairy cattle waste may disseminate antibiotic resistance in the wider environment. In addition, few studies have been conducted on the application of dairy cattle waste to soil on waste collected from farms with antibiotic regulations comparable to the UK. China and the US dominate AMR publications, including those investigating dairy cattle waste amendment and represent countries which regulate agricultural antibiotic use differently from the UK and Europe as a whole (Zhang et al., 2006, Qiao et al., 2018). Nonetheless, there are exceptions, such as a study documenting ARGs on cattle and swine farmland in Finland, where antibiotic use is regulated in accordance with EU law (Muurinen et al., 2017). From a technical perspective, the majority of cattle waste-orientated studies employ qPCR, with few using metagenomic techniques, which are capable of screening thousands of ARGs across multiple databases rather than user-defined arrays limited to hundreds of genes. A review of existing literature suggests many qPCR-based investigations studying cattle waste focus on fewer than 10 ARGs (Fahrenfeld et al., 2014, Kyselková et al., 2015b, Nõlvak et al., 2016, Dungan et al.,

2019), while an increasing number employ large arrays consisting of 80 or more ARGs (Hu et al., 2016, Muurinen et al., 2017, Zhang et al., 2017b, Chen et al., 2019b).

Although the impact of cattle waste application on the soil resistome has been addressed by several microcosm-based studies (Kyselková et al., 2015b, Sandberg and LaPara, 2016, Zhang et al., 2017b, Chen et al., 2019a), a limited number carried out field-scale surveys. Field studies currently available typically provide data for soil ARGs at various points before and after cattle waste application, however, few trace the abundance of ARGs over multiple fertilisation events. Although Nõlvak et al. (2016) assessed the effects of mineral fertiliser, cattle slurry and slurry digestate on a small number of ARGs over three application events, the experimental plots used had no prior history of animal waste application and therefore do not provide an insight into the effects of cattle slurry application over a prolonged period of many years. On the other hand, a study using a large qPCR array with >300 ARGs followed the effects of cattle and swine waste application on fields with an extensive history of manure amendment, although the effects of multiple applications were not directly characterised (Muurinen et al., 2017). A survey of several field sites suggested that those subject to long-term dairy cattle waste amendment contained an increased abundance of ARGs relative to sites with no history of exposure, however only 6 ARGs were evaluated and the temporal succession of ARGs following fertilisation was not measured (Dungan et al., 2019). While another field study using experimental plots showed that application rate had a more significant impact on ARGs than application number (McKinney et al., 2018), only three ARGs were quantified and it is important to understand the effects of regular dairy cattle waste fertilisation on a working dairy farm using conventional application rates. More recently, studies from China have assessed the long-term impact of animal waste application on fields, although these typically involve anaerobic digestate or manure originating from swine farms (Xie et al., 2018a, Liu et al., 2020, Lu et al., 2020). Finally, recent reviews have underscored that a number of areas require further research, including the temporal effect of animal waste application on grassland soils, the impact of conventional storage methods on the resistome and applied studies characterising farming systems to

establish the scalability of existing laboratory-based findings (Tyrrell et al., 2019, Oliver et al., 2020a).

The present study aims to address some of the knowledge gaps previously outlined by thoroughly sampling a semi-permanent grassland soil, which frequently receives cattle slurry from an adjacent dairy farm. To provide a comprehensive picture of the resistome and microbiome, metagenomic sequencing was carried out on DNA extracted from soil samples obtained at regular intervals over a 12-month period. Sampling encompassed regular slurry applications and silage cuts. Furthermore, slurry was sampled periodically to quantify parallel temporal changes in slurry and soil. In addition, a nearby permanent grassland with no history of cattle waste application was also sampled as a comparison with the primary site, which received cattle slurry amendments. The aim of the current study was to characterise and quantify field-scale changes across time and over multiple slurry applications with reference to concurrent changes in on-site cattle slurry. Sampling multiple fields in such detail was logistically and financially impossible; however, to my knowledge, this is the first such study in the UK, and represents one of the few to provide temporal metagenomic data with intensive within-field replication.

2.2 Materials and Methods

2.2.1 Approach and Sampling

2.2.1.1 Sample Sites

Both field sites are situated in Sutton Bonington, East Midlands; a designated nitrate vulnerable zone (NVZ). The farm is a research farm, but it operates commercially and management practices are therefore largely representative of a conventional medium-sized UK farm. The main field site is a semi-permanent pasture adjoining the dairy unit, which housed >250 Holstein Friesian cattle at the time of study. The field is regularly fertilised with cattle slurry and grass is routinely cut for silage production. Slurry is

managed on site via a sump system connected to an open-topped 3M L main tank and a 9M L overflow lagoon. The slurry solid fraction (fibre) is removed by a screw-press prior to collection in the main tank. The lagoon aids storage during the NVZ closed period between 15th October and 31stJanuary (GOV.UK, 2020b). All slurry samples were obtained from the 3M L main slurry tank.

A second field site with no history of agricultural animal waste application was also utilised in this study. This site is a nearby, permanent grassland with silty clay loam soil texture. Soil Antibiotic use at the dairy unit is shown in Table 2.1.

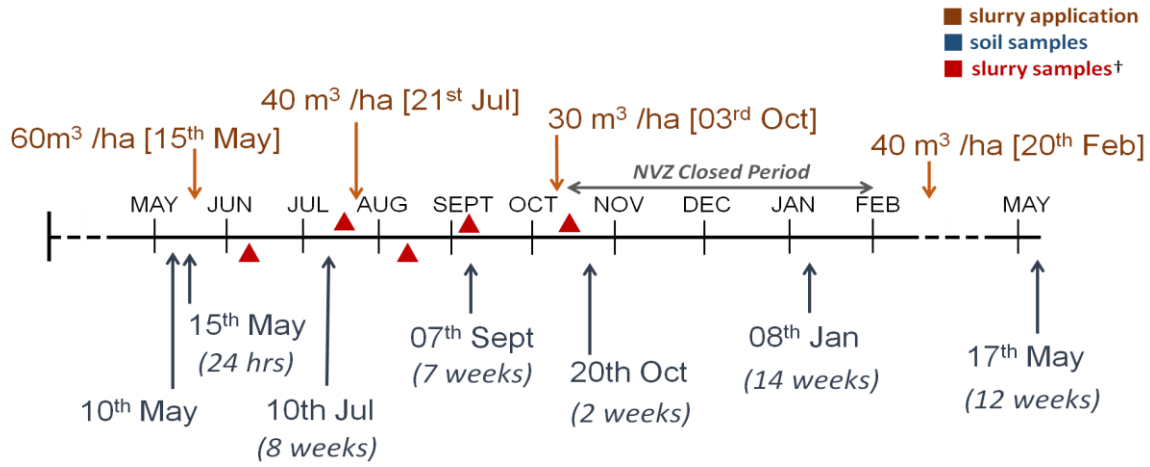
Table 2.1 On-site antibiotic use (adult herd) at the study farm dairy unit (2015-2017). Note: tulathromycin (a macrolide antibiotic) was occasionally administered to calves.

	Doses administered					
Product Name	**2015**	**2016**	**2017**	**Active Ingredient**	**Antibiotic class**	**Quantity**
Alamycin	0	39	80	Oxytetracycline Hydrochloride	Tetracycline	100mg/ml
Alamycin LA	5	109	0	Oxytetracycline Dihydrate	Tetracycline	200mg/ml
Betamox	3	24	74	Amoxicillin	Beta-lactam	150mg/ml
Bimotrim	0	19	0	Sulfadoxine	Sulfonamide	200mg/ml
				Trimethoprim	Dihydrofolate reductase inhibitor	40mg/ml
Cephaguard DC	24	0	0	Cefquinome (as sulphate)	Beta-lactam (4th generation cephalosporin)	3g/ syringe
Ceporex	5	373	65	Cefalexin sodium equivalent to Cefalexin	Beta-lactam (1st generation cephalosporin)	180mg/ml
Engemycin DD	112	0	0	Oxytetracycline (as hydrochloride)	Tetracycline	100mg/ml
Hexasol LA	0	2	0	Oxytetracycline (as dihydrate)	Tetracycline	300mg/ml
				Flunixin (as flunixin meglumin)	NA (anti-inflammatory)	20mg/ml
Naxcel	63	1	0	Ceftiofur (as crystalline free acid)	Beta-lactam (3rd generation cephalosporin)	200mg/ml
Orbenin Dry Cow	0	1	0	Cloxacillin benzathine	Beta-lactam	500mg/dose
Orbenin Extra Dry Cow	45	80	33	Cloxacillin benzathine	Beta-lactam	600mg/syringe
Tetra-Delta	64	909	261	Novobiocin Sodium equal to Novobiocin	Aminocoumarin	100mg/dose
				Neomycin Sulphate equal to Neomycin	Aminoglycoside	105mg/dose
				Procaine Penicillin	Beta-lactam	100mg/dose
				Dihydrostreptomycin Sulphate	Aminoglycoside	100mg/dose
				Prednisolone	NA (anti-inflammatory)	10mg/dose
Ubro Yellow Milking Cow	225	113	0	Penethamate Hydriodide (diethylaminoethyl ester of benzylpenicillin)	Beta-lactam	150mg/dose
				Dihydrostreptomycin Sulphate	Aminoglycoside	185mg/dose
				Framycetin Sulphate	Aminoglycoside	50mg/dose
				Prednisolone	NA (anti-inflammatory)	5mg/dose
Ultrapen LA	2	22	10	Procaine Benzylpenicillin	Beta-lactam	300mg/ml

2.2.1.2 Sampling

Soil samples were taken from the slurry-fertilised semi-permanent grassland on seven occasions from May 2017 to May 2018, corresponding to an entire season of slurry application, the winter closed period for 2017 and the beginning of slurry application in 2018. Soil from the unslurried permanent grassland was collected on two occasions

(January and May 2018). Slurry was sampled each month from June to October 2017. For a detailed timeline of the sampling scheme see Figure 2.1.



† 07/06/2017, 11/07/2017, 01/08/2017, 05/09/2017, 10/10/2017

Figure 2.1 Experimental timeline detailing soil and slurry collection, in addition to slurry application events.

Sampling of the soil and slurry was not coupled as tightly as originally intended because the farm manager prioritised commercial operations over research requirements. While this means data are representative of a typical UK dairy farm (i.e. a real world scenario), the study had to accommodate unforeseen changes to the animal waste application regimen relating to evolving management practices and weather patterns.

Soils were sampled from five points, each 20m apart, along a 'W' transect taking care to avoid field edges and isolated trees. In an attempt to capture local heterogeneity three soil cores were taken around each point on the 'W' at a depth of 10cm, using an auger sterilised with ethanol wipes between collections. Samples were deposited in sterile re-sealable bags and transported directly to laboratory facilities. Transect sub-samples were then pooled, homogenised and sieved to 3mm, giving five discrete samples per site at each sampling time. Subsamples of the homogenised soil were processed immediately for culture-based analyses whilst soil for metagenomic shotgun sequencing was frozen at -80°C until DNA extraction (Lear et al., 2018). Remaining soil was stored overnight at 4°C for physiochemical analyses.

Slurry samples were collected by lowering a sterilised bucket into the main tank and decanting the contents into sterile Duran bottles. Slurry was then homogenised with a sterile spatula and sub-sampled before being stored as described above.

2.2.2 Extraction of DNA and Sequencing

In an attempt to ensure downstream analysis reflected the diversity of recovered soils, three separate DNA extractions were carried out for each homogenised subsample (135 soil extractions). Environmental DNA was extracted from 0.25g of soil using the DNeasy PowerSoil Kit (Qiagen) as per the manufacturer's instructions. Slurry samples were processed with an equivalent kit for faecal matter; QIAamp Powerfecal Kit (Qiagen). Subsequently, the quality and yield of extracts was quantified using a spectrophotometer (NanoDrop, Thermo Scientific) and fluorometer (Qubit, Invitrogen), respectively. Low quality extracts were discarded. The three extractions per soil sample were pooled (n = 45), quality checked (260/280 = 1.83 ± 0.03, 260/230 = 1.61 ± 0.16, DNA conc. = 39.4 ng/µL ± 9.80) and refrigerated at 4°C until sequencing. Lastly, negative controls (sterile water extractions) were also sent for sequencing.

The DNeasy PowerSoil kit was selected for a several reasons. Firstly, the kit has an established precedent of use in soil microbial studies (Lear et al., 2018), and therefore may offer greater opportunity for comparison against existing research. Furthermore, widespread use has encouraged a number of reviews on DNA extraction methods to include the PowerSoil kit in comparisons (Mahmoudi et al., 2011, Vishnivetskaya et al., 2014, Zielińska et al., 2017), leading to a better understanding of biases associated with this kit. In addition, the kit appears to offer an acceptable balance between per sample cost, extraction yield and convenience of use (Zielińska et al., 2017, Lear et al., 2018).

Metagenomic shotgun sequencing of extracted soil and slurry DNA was performed and demultiplexed by Edinburgh Genomics using the Illumina NovaSeq platform (NextaraXT 150 bp paired end libraries, >30M paired reads per sample). Prior to downstream

analysis all read libraries were quality checked with FastQC (Andrews, 2015) before and after the removal of adapters and low quality bases with Trim Galore (Krueger, 2012).

2.2.3 Annotation of ARGS

ARGs were annotated using DeepARG (Arango-Argoty et al., 2018). DeepARG employs a machine learning algorithm trained on multiple ARG databases to detect ARGs and ARG-like sequences. In this respect, DeepARG differs from many ARG annotation tools currently available, which rely on homology-based searches alone (McArthur and Tsang, 2017).

A short-read and long-read pipeline is available for DeepARG. The developers of DeepARG indicate the precision and recall of metagenomic library annotations can be maximised by using the long-read algorithm, which must be run on open reading frames obtained from assembled contigs. The potential advantage of using contigs is intuitive as these longer sequences increase the query length and therefore information available for analysis (Arango-Argoty et al., 2018, Ayling et al., 2020). However, the benefit of this approach relies heavily on the quality and completeness of the assembly. Furthermore, ARGs encoded on plasmids and within genomic islands may be lost from the final assembly as it has been shown assemblers often struggle to reconstruct these stretches of sequence (Maguire et al., 2020).

The choice of DeepARG pipeline was therefore determined after evaluating the completeness and contiguity of *de novo* assemblies constructed with Megahit (Li et al., 2015b). Megahit was selected for metagenome assembly since while it exhibits greater assembly accuracy at the expense of contig length (Ayling et al., 2020) it has also been shown to outperform other leading assemblers such as MetaSPAdes when annotating genes in highly complex soil communities (Quince et al., 2017).

Despite making use of settings designed specifically for complex metagenomes (e.g. reducing k-mer step size) only 53.4 ± 1.3% of reads, on average, were properly paired

and mapped to the resulting assemblies generated from individual soil samples, in contrast to slurry samples (80.5% ± 0.59%). To maximise use of soil sequence data, ARGs were therefore annotated using unassembled reads and the short-read DeepARG pipeline, with default settings (identity ≥ 80%, probability ≥ 0.8, e-value < $1e^{-10}$).

2.2.4 Data Analysis and Statistical Methods

2.2.4.1 Determining ARG Associations by Site: Data Exploration

Initial data exploration was performed using principal component analysis (PCA), non-metric multidimensional scaling (NMDS) and t-distributed stochastic neighbour embedding (tSNE). The aforementioned methods all have particular strengths and weaknesses with regard to their interpretative power and suitability for these data. It was hoped that by using a combination of methods their individual shortcomings could be overcome, with greater confidence being placed in common patterns. PCA is often considered a routine method of data exploration, enabling researchers to carry out reproducible dimension reduction by distinguishing variables which describe the greatest variation in a dataset (principle components) (Holmes and Huber, 2018). PCA is also effective at describing global data structure (Nguyen and Holmes, 2019). However, the mathematical basis of PCA concerns linear relationships, rather than those which are non-linear (Holmes and Huber, 2018). NMDS ordination offers an alternative to PCA with improved capacity for handling polynomial interactions, albeit with an emphasis on local structure (Nguyen and Holmes, 2019). Unfortunately, NMDS of the full data set at the ARG group level failed to converge. This could be explained by the detection of many ARGs unique to slurry samples, leading to a disconnected, irresolvable NMDS plot. A similar occurrence has been reported in microbial community data elsewhere (Bell et al., 2018). Lastly, tSNE was used with the intention of finding a non-linear means of visualising the entire data-set at the ARG group level. At a high level, tSNE differs from NMDS in that it adopts a flexible, probabilistic approach which allows the user to balance local and global data structure. Optimisation of tSNE was carried out by increasing the

number of iterations until stable clustering was achieved using a suitable perplexity (between 5 and 50).

2.2.4.2 Determining ARG Associations by Site: Differential Abundance Analysis and Feature Selection

A cross-validated, model-based approach was adopted to test which ARGs were significantly associated with slurry, slurry-amended and untreated sites. Firstly, this involved the use of Corncob (count regression for correlated observations with the beta-binomial) to characterise differentially abundant ARGs (Martin et al., 2020). Corncob was selected in preference to more frequently used software such as DeSeq2 as in addition to addressing uneven sequencing depth, the former attempts to account for several other perennial statistical issues associated with sequence count data, including over-dispersion, zero-inflation, and within-group correlations (Martin et al., 2020). Corncob was also used to investigate site-specific seasonal differences in ARG abundance. The strength of ARG associations with specific sites were independently validated through use of the Boruta feature selection algorithm (Kursa and Rudnicki, 2010) and LEfSe (linear discriminant analysis effect size) biomarker identification (Segata et al., 2011).

In data science, feature selection is commonly used to determine important variables in datasets with high dimensionality and inherent noise, where the relevance of many variables for classification purposes is unknown in advance (Bolón-Canedo et al., 2015, Rudnicki et al., 2015). It is therefore unsurprising that feature selection has been used in a range of research domains including bioinformatics (Aagaard et al., 2012, Ditzler et al., 2015, Colaco et al., 2019). Boruta is a Random Forest wrapper algorithm which works by generating models which include features ordered by their estimated importance relative to the classifier outcome (e.g. site association) and duplicate 'shadow features' which have been randomly shuffled. Features are sequentially removed depending on whether they exceed the importance of their equivalent 'shadow features'. The process is

repeated until all features are confirmed as either important or unimportant (Kursa and Rudnicki, 2010, Degenhardt et al., 2019).

Most feature selection methods focus on identifying the *optimum minimum* variable set, whereas Boruta aims to establish *all-relevant* variables (Kursa and Rudnicki, 2010, Rudnicki et al., 2015). The significance of this distinction is that while the *optimum minimum* defines only strongly relevant variables (the fewest variables needed for classification), an *all-relevant* approach also includes weakly relevant variables which would otherwise be masked by their redundancy to stronger selection features (Rudnicki et al., 2015, Göpfert et al., 2018, Degenhardt et al., 2019).

LEfSe was included in analysis because it was developed specifically for the identification of biomarkers in metagenomic datasets (Segata et al., 2011) and has been widely used in publications (Looft et al., 2014, Chen et al., 2016b, Morrison et al., 2020, Tango et al., 2020). LEfSe also provides a useful comparison to Boruta, since it incorporates variable effect size into the evaluation step, thus leaning towards describing the *optimum minimum*.

2.2.4.3 Hierarchical Cluster Analysis and Heatmap

Cluster analysis was carried out on ARG groups shown to be both differentially abundant by Corncob and Boruta. Boruta was used in preference to LEfSe selections as it avoids discarding subtly relevant site descriptors. The resulting clusters were visualised and annotated using a modified heatmap.2 script in R (R Core Team, 2020).

2.2.4.4 Changes in ARG Abundance Following Slurry Application

In order to determine whether the application of slurry induced transient or persistent changes in the treated soil resistome, LEfSe biomarker identification was carried out. To this end, the resistome of background samples collected five days before the first application of slurry in 2017 were compared with samples collected within 24 hours of the first slurry application in May 2017 and 56 days after slurry application. The

described samples represent the window between the first and second application of slurry in 2017.

2.2.4.5 ARG Richness and Diversity Estimations by Site

ARG richness and alpha diversity was estimated with iNEXT (Hsieh et al., 2016). The estimates calculated by iNEXT are based on effective number of species (bootstrapped Hill numbers) and therefore attempt to account for incomplete sampling which traditional calculation methods often neglect to consider (Chao et al., 2014). Specifically, Shannon's and Inverse Simpson's Diversity measures were estimated for alpha diversity. Briefly, both Shannon's (q1) and Simpson's Diversity index (q2) take into account species richness and evenness (abundance), however Simpson's Diversity places more emphasis on dominant taxa. Richness (q0) relates to the estimated number of unique species with no weighting relating to abundance. In the context of this Chapter ARGs are the 'species' under consideration.

After checking both homogeneity of variance between site groups and within-group normality, significance testing was conducted. Briefly, if assumptions were met one-way ANOVA was performed and if significant ($p < 0.05$) pair-wise TukeyHSD tests were carried out. In cases of non-normality a Kruskall-Wallis test was performed instead. Significant Kruskall-Wallis tests were followed by a Dunn pair-wise test (implemented in the R package FSA v0.9.1). In cases where heterogeneity of variance was observed a Welch-adjusted ANOVA was carried out and if significant, followed by the Games-Howell test (implemented in R package PMCMRplus v1.9.3). Lastly, p-values were adjusted for multiple comparisons.

Beta diversity (Bray-Curtis distance estimates) of ARGs were estimated using the R package vegan v2.5-6. PERMANOVA was carried out for beta-diversity in PRIMER (Clarke and Gorley, 2006).

2.2.4.6 ARG-ARG Network Construction

To compliment differential abundance analysis intra-site ARG-ARG associations were explored with SpiecEasi; a network construction and visualisation package which aims to provide a tractable solution to the problem of network inference with compositional and underpowered datasets (Kurtz et al., 2015).

Arguments against using traditional correlation-based analyses for network construction with compositional data were outlined by Lovell et al. (2015). However it has also been shown that alternative methods like SpiecEasi do not necessarily outperform correlation-based analysis (Hirano and Takemoto, 2019). For this reason, it was decided to carry out additional testing using Spearman's Rank correlation analysis accounting for multiple testing with fdr (false discovery rate) adjusted p-values.

To limit the detection of random associations between rare ARGs, only those which occurred at least once in more than half the samples from a given site were included in analysis.

It is important to emphasise that the correlation-based networks presented in the current and subsequent chapters represent putative associations between variables, which require further validation before they can be categorically established. Indeed, in Chapter 4 specific ARG-ARG as well as ARG-taxon associations are examined without the use of correlation-based inferences using assembled contigs. Furthermore, where possible and practical, existing literature has been examined to provide possible biological explanations for specific correlations where further validation through experimentation was beyond the scope of the present work. In summary, while the included networks are a useful tool for data exploration and hypothesis formation, it is vital they are used responsibly in tandem with other methods.

2.3 Results

2.3.1 Determining ARG Associations by Site

DeepARG annotated a total of 454 ARG groups representing 23 antibiotic resistance categories. Given the number of ARG groups detected, only those which were identified as differentially abundant or are of particular clinical interest are discussed in detail. Figure 2.2 summarises the extent to which ARG groups were shared across sites, showing that 30.4% of ARG groups were only detected in a single site, with slurry-amended soil possessing the greatest number of uniquely detected groups (n = 69). However, a similar number of ARG groups were shared by all sites (32.4%).

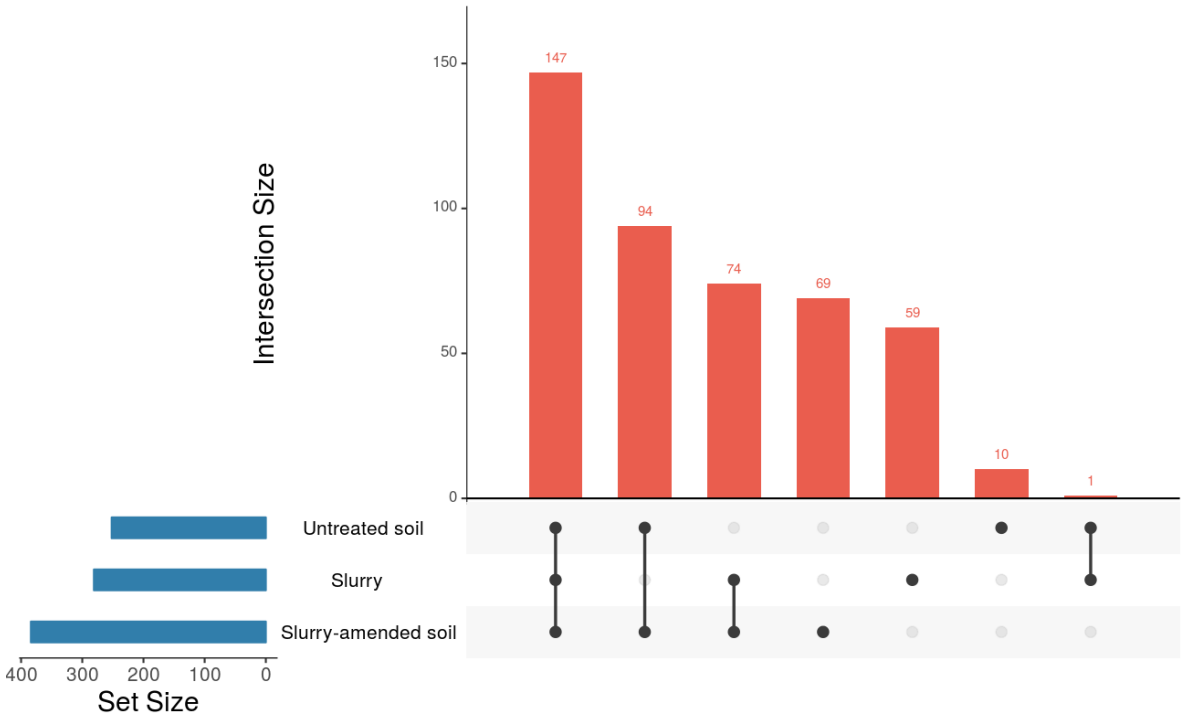


Figure 2.2 UpSet plot created using Intervene (Khan and Mathelier, 2017), illustrating the extent to which ARGs were shared across sample types (detected or undetected). Solid vertical lines connecting sites on the x-axis denote ARG group detection across multiple sites. Single dots on the x-axis denote groups which were detected in only one site. Set size relates to the total number of unique ARG groups identified within a given site. In this context, site includes the slurry tank in addition to the two field sites.

In contrast, all but three ARG categories were shared across the three sites. Specifically, oxazolidinones were consistently identified in slurry alone, sulfonamides in slurry and sporadically slurry-amended soil, while tetracenomycins were exclusively detected in soil samples (regardless of treatment history).

PCA of antibiotic category data highlighted a clear dissociation between slurry and soil samples across PC1 (57.5% variance), and a less pronounced difference between the two soil sites across PC2 (7.4% variance) as shown in Figure 2.3. Untreated soil shows greater variance than amended soil along PC2, whereas the latter is more variable along PC1. A similar pattern was observed at the ARG level.

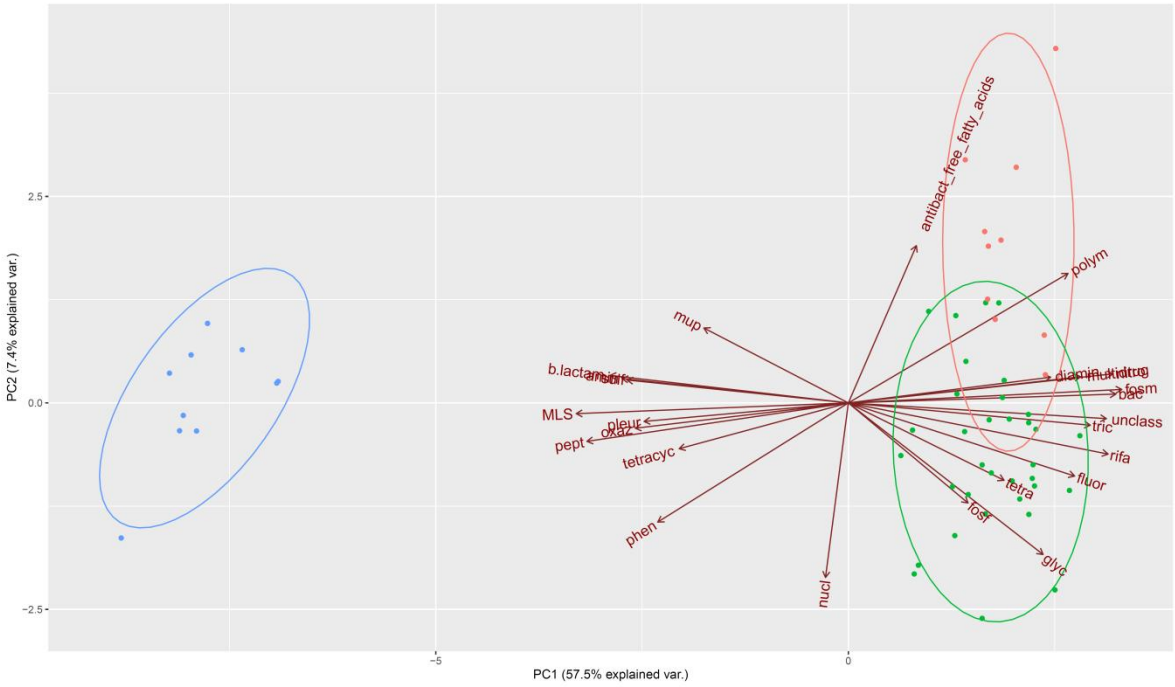


Figure 2.3 PCA biplot of 16S rRNA-normalised ARG category abundances. Colours correspond to sample origin (blue, slurry; orange, untreated soil; green, slurry-amended soil). Ellipses represent 0.9 normal probability. Data are scaled and centred. Arrows represent ARG categories. The direction of arrows reflect relative association with a principal component, while length indicates increasing values (relative abundance). The peptide category was divided on the basis of activity against Gram negative and Gram positive bacteria.

NMDS was used to further explore differences in resistome structure. At the ARG category level, NMDS was broadly similar to PCA. Initially all data were included in ARG

group analysis, however the full dataset failed to converge. Upon excluding slurry data from NMDS, and focussing on directly comparable soil data from January and May 2018, a robust analysis of this subset was achieved ($k = 2$, stress = 0.047). The resulting plot suggested seasonal variation between soil sites during these months, in addition to site differences previously described by PCA at the ARG category level (Figure 2.4). Seasonal divergence in ARG group ordination appears to increase in May relative to January. A similar trend was shown by plotting the average relative abundances of dominant ARG categories over the same period (Figure 2.5). Permutational multivariate analysis of variance (PERMANOVA) returned significant results for site and site-month interaction effects ($p = 0.001$ and $p = 0.048$, respectively). However, these inferences should be treated with caution as a test for dispersal with the betadisper function (R package, vegan) indicated heterogeneity across sample sites ($p < 0.05$), with untreated soil samples exhibiting significantly greater dispersion than slurry impacted soil samples. This is interesting, as the reduced variability seen in slurry impacted soil samples relative to samples collected from the site with no history of slurry application could reflect the impact of agricultural management practices on the resistome.

For instance, the routine application of slurry as well as other farm-associated management practices could lead to the development of a more stable resistome relative to the site with no history of large-scale organic fertilisation. The untreated site may therefore be more variable because it is governed by more random environmental effects.

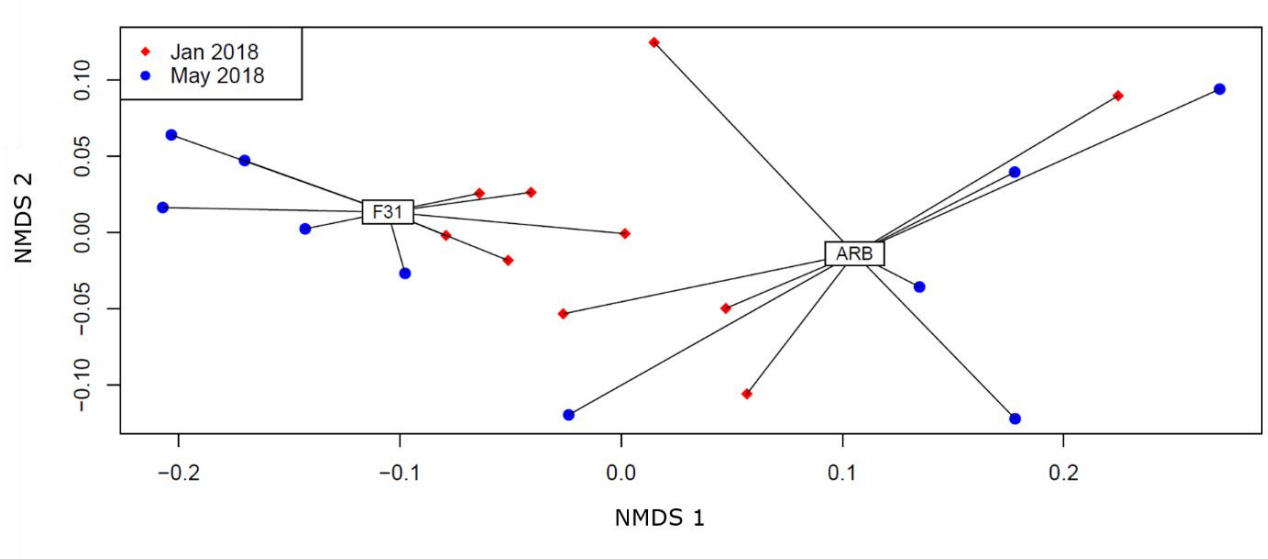


Figure 2.4 NMDS plot comparing 16S rRNA-normalised ARG group abundances in slurry-amended (F31) and untreated soils (ARB) recovered in January and May 2018 (k = 2, stress = 0.047, non-metric fit R^2 = 0.998, metric fit = 0.99). Placement of site labels reflect group centroids (averages). Point colours correspond to month (red: January, blue: May 2018).

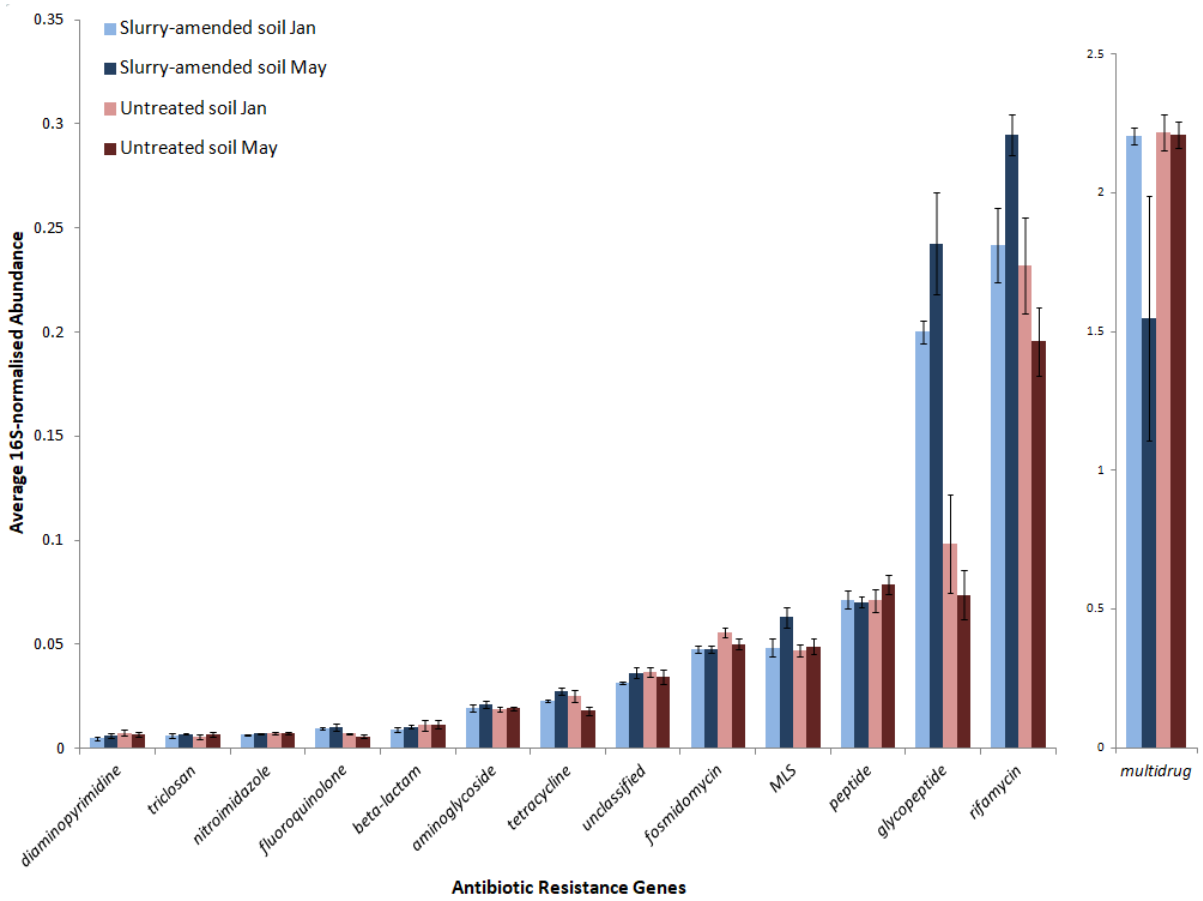


Figure 2.5 Average 16S rRNA–normalised abundance of ARGs for samples collected from slurry-impacted and untreated sites in January and May 2018. Standard error shown.

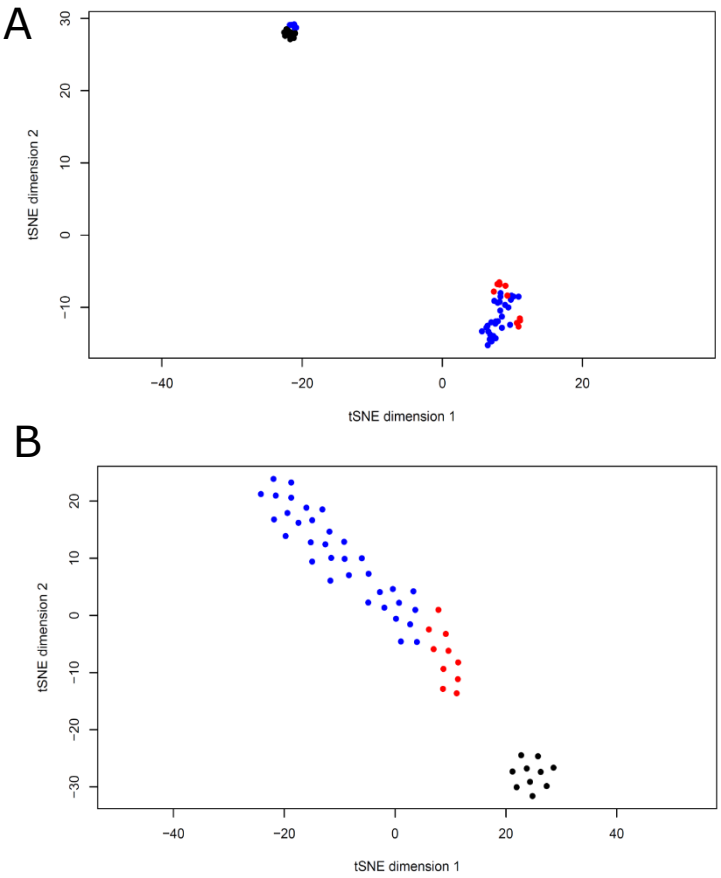


Figure 2.6 tSNE plots of ARG class (A) and group (B) data. For both plots perplexity was set to 15 and iterations 1500. Black = slurry, blue = treated soil site, red = untreated soil site.

tSNE analysis showed different clustering patterns at the ARG group and ARG category level (Figure 2.6). For ARG groups, all slurry samples clustered together, while soil samples fell into two potential clusters; one containing amended soil samples only, and the other containing all untreated soil samples together with remaining slurry-amended soil samples. However, samples within the 'mixed' cluster were clearly divided by site. The same analysis carried out at the category level yielded different results. Only two clear clusters emerged. One cluster contained most slurry-amended soil samples together with all untreated soil samples, while the other contained all slurry samples together with a small subset of slurry-amended soil samples. Slurry-amended soil samples which clustered with slurry belonged to sample sets collected in May 2017 (n = 2), September 2017 (n = 1) and May 2018 (n = 2). tSNE analysis found similar patterns to NMDS for soil data in 2018 (not shown).

2.3.2 Site Differential Abundance

Data exploration supported the hypothesis that slurry and soil resistomes were distinct, however the analyses were not suitable for statistically identifying and visualising which of the 454 ARG groups distinguished between sites. Furthermore, the degree of slurry-associated ARG enrichment in amended-soils remained unclear. Corncob was used in conjunction with cluster analysis to address these questions.

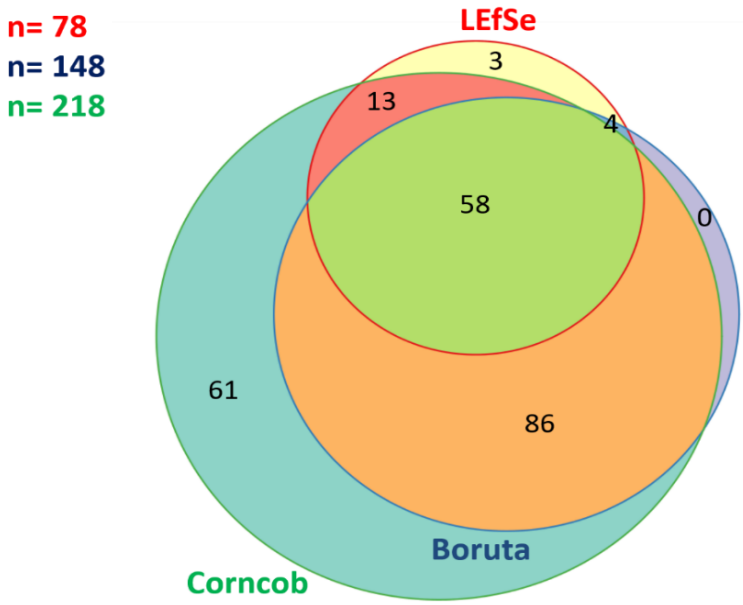


Figure 2.7 Venn diagram showing the intersection between differentially abundant ARG groups (by site) based on Corncob, Boruta and LEfSe methods.

Controlling for dispersion in site data, Corncob identified 218 ARG groups and 21 ARG categories that were differentially abundant across sites. For complete Corncob analysis plot including all differentially abundant ARG groups see Supplementary file 1. Feature selection with Boruta and LEfSe was carried out with the aim of independently identifying differentially abundant ARG groups as well as those, which discriminated between sites most effectively. Encouragingly, Corncob encompassed over 95% of ARG groups identified by selection methods. Boruta retained more than half the ARG groups shown to be differentially abundant by Corncob (Figure 2.7). As expected LEfSe was more stringent, however, it largely selected a subset of groups already confirmed by Boruta.

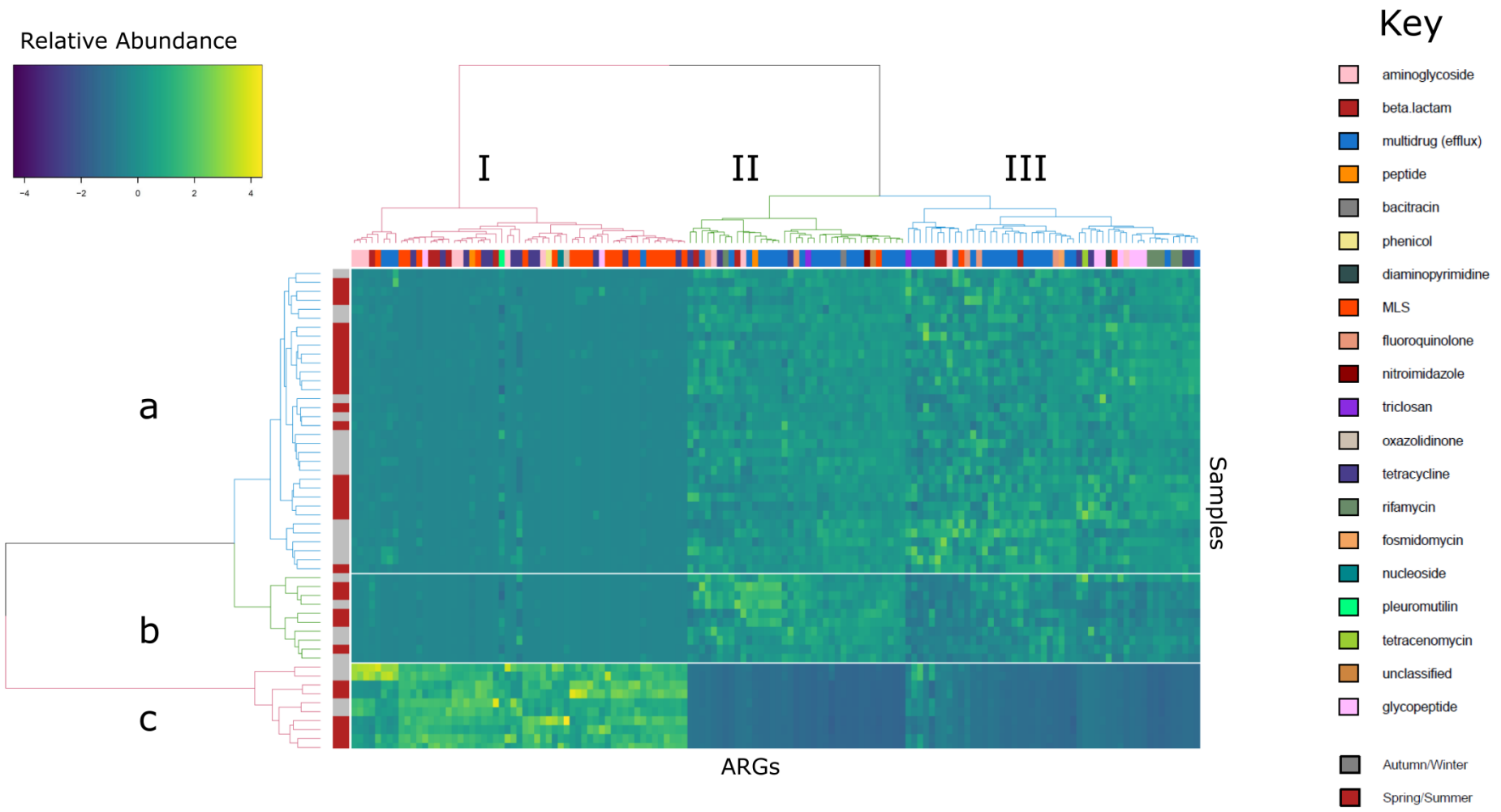


Figure 2.8 Heatmap of differentially abundant ARGs by site (based on intersection between Corncob and Boruta). Colour gradient represents increasing relative 16S rRNA-normalised abundance from blue (lower) through to yellow (higher). Each column represents a single ARG group and rows denote samples. Clusters a-c denote samples, while clusters I-III denote ARGs. The y-axis annotation bar indicates season, while the x-axis annotation bar refers to ARG category (see key for colour descriptions).

Hierarchical cluster analysis of differentially abundant ARG groups showed distinct clusters for all three sites, with slurry samples clustered at a greater distance from both soil site samples (Figure 2.8). Furthermore, specific ARG categories dominated site clusters. For instance, slurry contained an assortment of tetracycline, aminoglycoside, beta-lactam and macrolide-lincosamide (MLS) resistance genes with greater relative abundance in comparison to soil clusters. In contrast, soil clusters contained many different multidrug efflux pump gene groups. However, the detection levels of glycopeptide, rifamycin and tetracycline resistance genes in slurry-amended soil appear to distinguish amended soil from untreated soil. A small group of ARG groups were also consistently more abundant in slurry-amended soil than untreated soil (Figure 2.9).

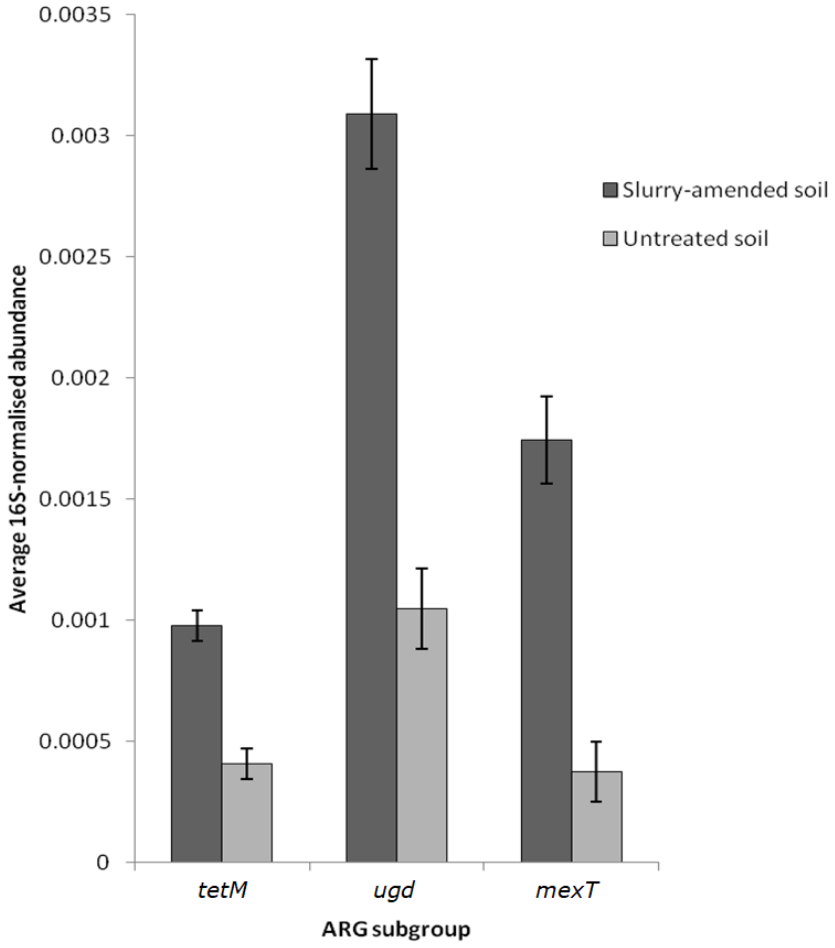


Figure 2.9 Candidate slurry associated ARG groups suggestive of long-term enrichment (standard error shown).

In summary, several analytical methods and visualisation techniques were used to identify differentially abundant ARGs which help to explain the differences between sites first shown by initial data exploration.

2.3.3 Seasonal Differences in ARG Abundance

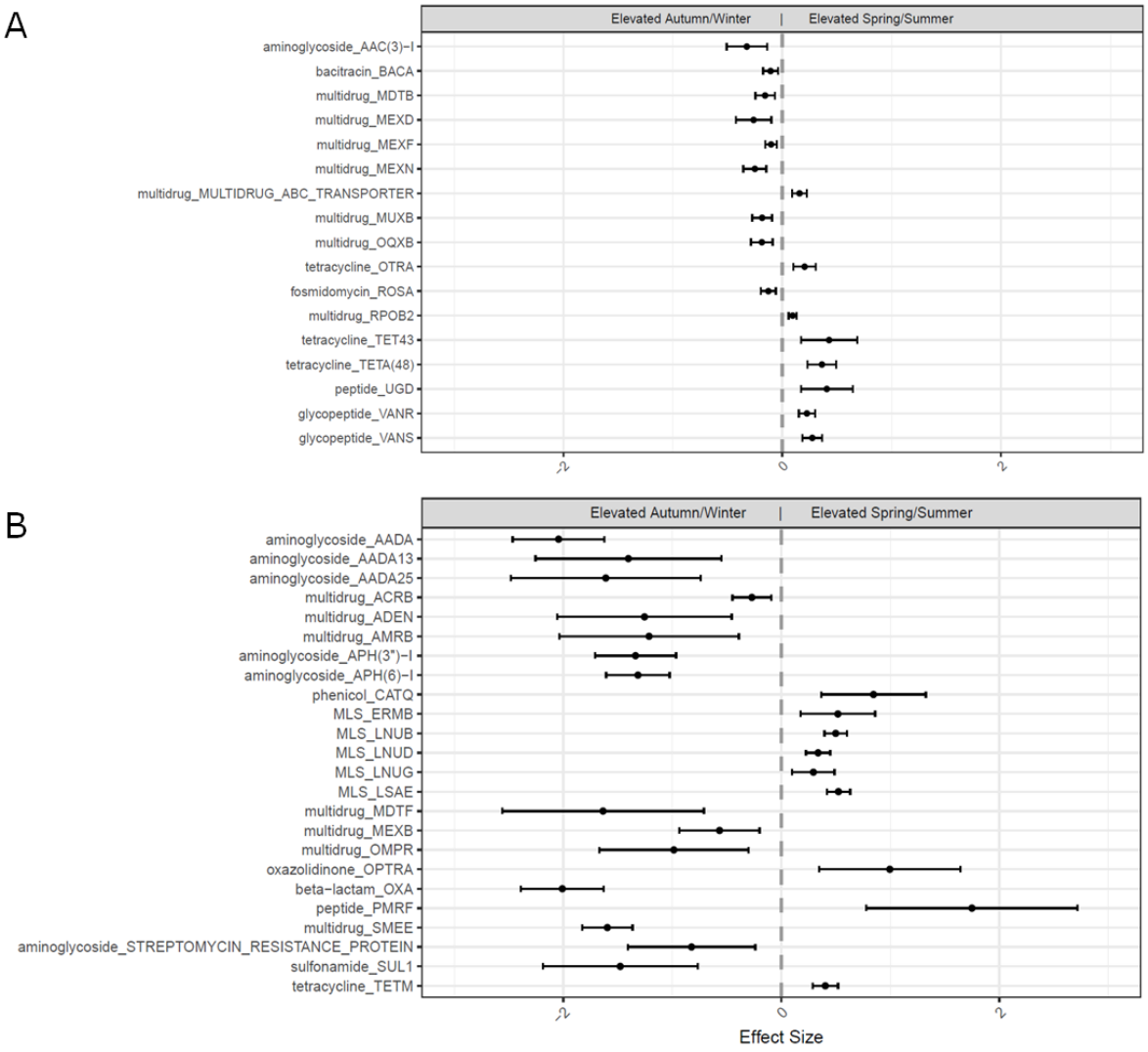


Figure 2.10 Differential abundance of ARGs by season. Slurry-impacted soil (A). Slurry (B). Points to the right of the central dotted line represent ARGs with increased abundance in spring/summer relative to autumn/winter, while those to the left are more abundant in autumn/winter relative to spring/summer. Error bars denote 95% confidence intervals.

Corncob identified a small subset of ARGs that were differentially abundant across season in slurry (n = 24) and slurry-amended soil (n = 17). In slurry, MLS resistance genes were consistently more abundant in spring/summer compared to autumn/winter (Figure 2.10B). ARGs, which were comparatively less abundant over the same period, belonged predominantly to the aminoglycoside and multidrug resistance groups. In slurry-amended soil tetracycline and glycopeptides resistance genes were more abundant in spring/summer, while multidrug resistance genes were significantly less abundant (Figure 2.10A). In contrast, no detected ARGs exhibited significant differences in abundance based on season in samples from untreated soil (analyses not shown).

2.3.4 Changes in ARG Abundance Following Slurry Application

LEfSe analysis showed the relative abundance of 23 ARGs was significantly greater in soil collected within 24 hours of slurry amendment relative to samples obtained five days prior to application (Figure 2.11A). Over half of these ARGs were consistently abundant in slurry; namely eight MLS genes and two tetracycline resistance genes. Furthermore the MLS resistance genes *lsaE*, *lnuB* and *mefB*, as well as the two tetracycline resistance genes *tetT* and *tet(36)*, were among the top 20 most abundant ARGs detected in slurry (average relative abundance). In contrast, after 56 days none of these genes were significantly increased in soil relative to pre-application samples (Figure 2.11B). In addition, only two genes out of the 18 ARGs elevated in soil 56 days after treatment were convincingly associated with slurry. These were the multidrug resistance genes *marR* and *adeC*. Irrespective of whether samples were taken less than 24 hours, or 56 days after slurry application, fewer gene groups were significantly less abundant than more abundant, and ARGs with reduced abundance were rarely considered markers for slurry.

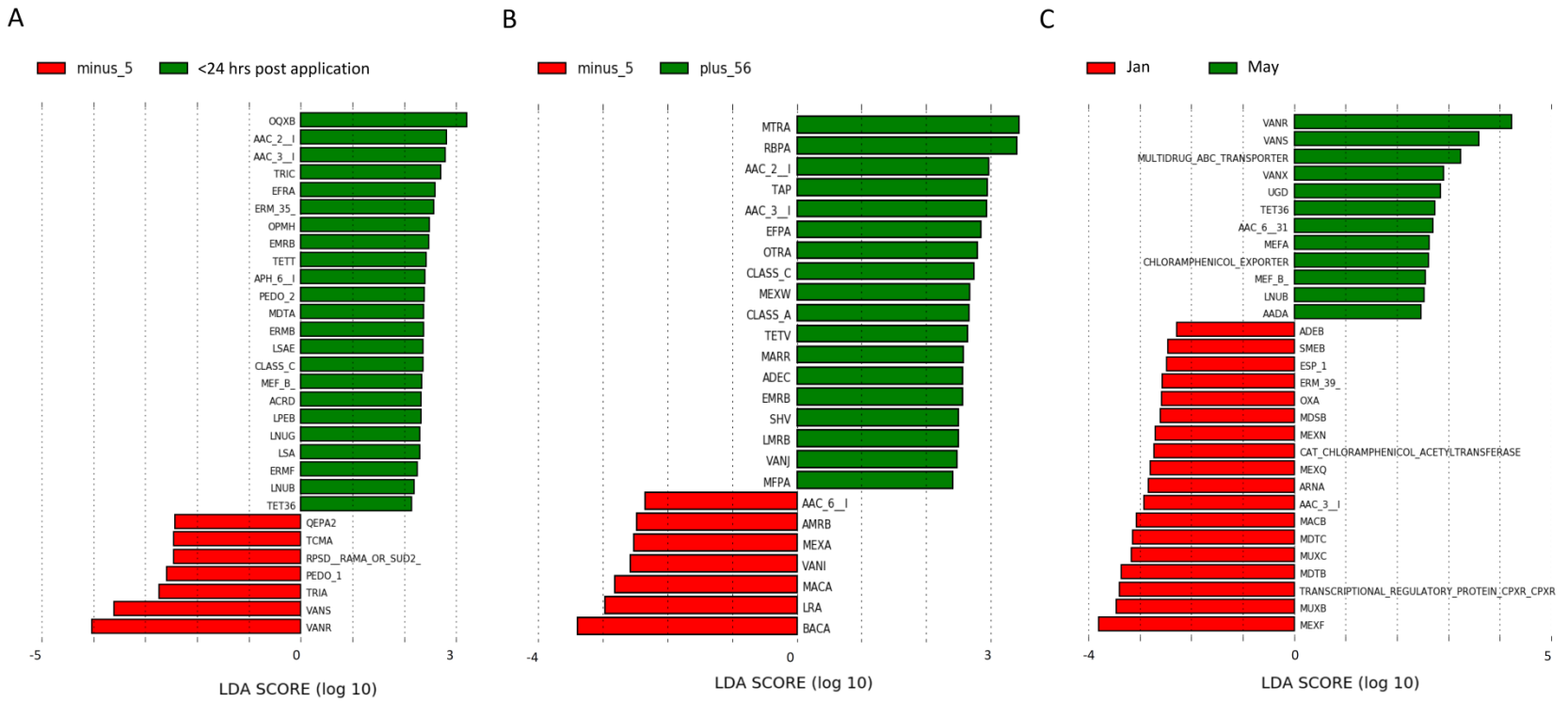


Figure 2.11 LEfSe analyses comparing 16S rRNA-normalised relative abundance of ARGs in soils before and after slurry application in May 2017 (A – B) and 2018 (C) (slurry-impacted site). Reference to 'minus' and 'plus' denotes days before and after slurry application, respectively. Green bars represent ARGs which had a greater relative abundance following slurry application, red bars those with greater relative abundance prior to slurry application (relative to the pre-application samples being assessed).

NMDS analysis (Figure 2.12) largely supported these findings indicating the composition of the resistome altered immediately after slurry application, but became more similar to the pre-treatment configuration within 56 days (PERMANOVA, time as factor: p = 0.02).

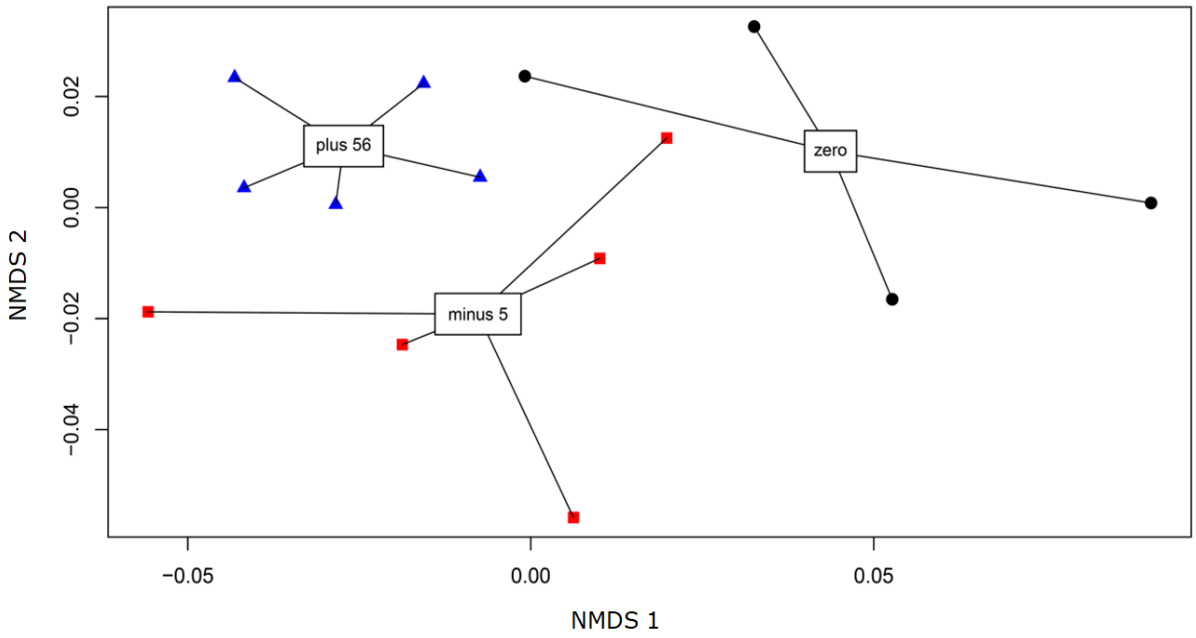


Figure 2.12 NMDS plot (k = 2, stress = 0.1) of the slurry-treated soil resistome five days before the first slurry application of 2017 (red squares), zero days after slurry treatment (black circles) and 56 days after treatment (blue triangles). PERMANOVA indicated time was a significant factor (R^2 = 0.30, p < 0.05), and assumption for homogeneity of dispersion was met (betdisper = p > 0.05). Placement of time-course labels reflect group centroids (averages).

It appeared that certain beta-lactamases were also elevated following the addition of slurry to soil, although these genes were not consistently detected in slurry. On day zero, class C beta-lactamases were more abundant, and both class A and C beta-lactamase resistance genes were more abundant in soil 56 days after slurry amendment.

2.3.5 ARG Diversity Estimations by Site

Estimates of mean richness (iNEXT q0) indicate that both slurry and slurry-impacted soil have significantly greater richness of ARG subgroups (p <0.001) relative to soil with no

history of slurry application. The elevated ARG richness estimates for slurry-impacted soil relative to untreated soil possibly indicate the introduction of slurry-associated ARGs. In contrast, both Shannon (iNEXT q1) and Inverse Simpson's Diversity (iNEXT q2) measures show that while slurry is significantly different from both soil sites, no significant difference was detected between soil sites. This suggests that slurry exhibits a greater diversity of ARG subtypes than soil samples based on moderate (q1) and low weighting (q2) of rare taxa, while soils exhibited similar levels of diversity (in particular, no significant change in dominant ARGs). The greater diversity of dominant ARG subtypes in slurry relative to soil is likely due to the intensity of exposure to various anthropogenic selection pressures within slurry (e.g. antibiotics, heavy metals and biocides).

Bray-Curtis distance estimates for beta diversity indicated that soil samples showed significant separation by site ($p < 0.001$, average distance between soil sites < 0.2), however all soil samples were also highly dissimilar from slurry at the ARG group level ($p < 0.001$, average distance from either soil site group > 0.8), in accordance with initial data exploration. Diversity measure comparisons are summarised in Table 2.2.

Table 2.2 Pair-wise comparisons of ARG group richness, alpha diversity measure estimates according to iNEXT and beta diversity (Bray-Curtis distances).

Diversity measure	Pairwise comparison (Site)	p- value
iNEXT Richness	Slurry-impacted / Untreated	< 0.001 **
	Slurry / Untreated	< 0.001 **
	Slurry-impacted / Slurry	> 0.99
iNEXT Shannon	Slurry-impacted / Untreated	0.22
	Slurry / Untreated	< 0.001 **
	Slurry-impacted / Slurry	0.002 *
iNEXT Inverse Simpson	Slurry-impacted / Untreated	0.98
	Slurry / Untreated	0.017 *
	Slurry-impacted / Slurry	0.009 *
Bray-Curtis Distance	Slurry-impacted / Untreated	< 0.001 **
	Slurry / Untreated	< 0.001 **
	Slurry-impacted / Slurry	< 0.001 **

2.3.6 ARG-ARG Network Analysis

The connectivity and composition of ARG networks differed by site (Figure 2.13). Slurry samples had the fewest isolated nodes (n = 11), followed by soil networks (n = 28). This is reflected by the degree distributions, which describe the probability of nodes sharing an increasing number of edges. It is clear that while most nodes in slurry and untreated soil networks have fewer than three connections, several nodes in slurry-amended soil have between four and eight connections, which may represent hubs.

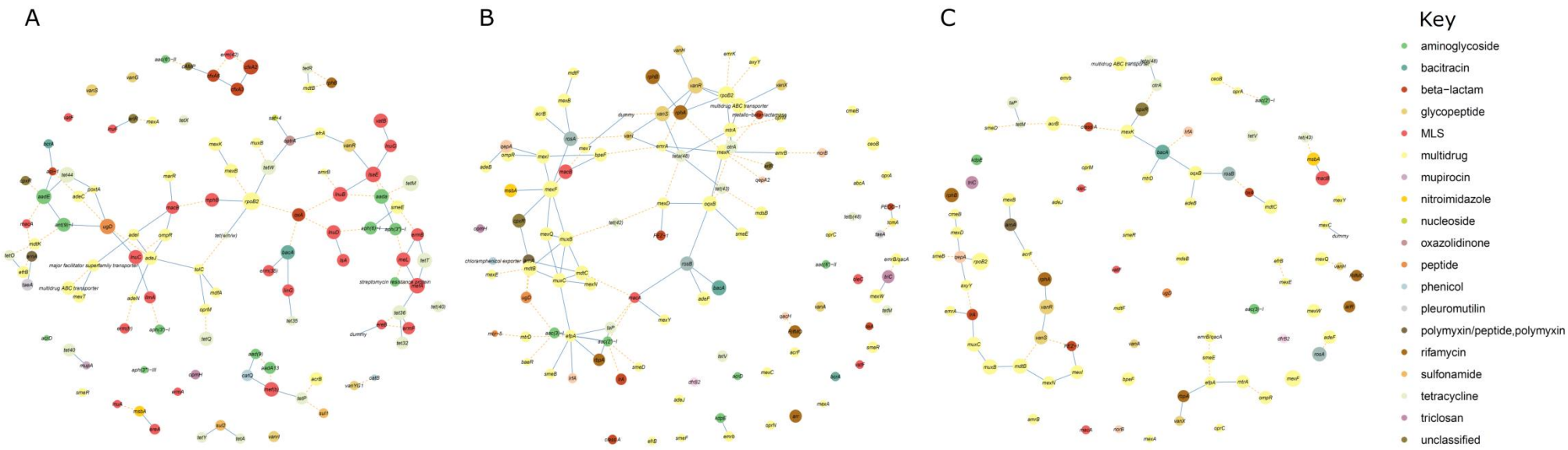


Figure 2.13 SpiecEasi network analysis of ARGs. Slurry samples (A), slurry-impacted soil (B) and untreated soil (C). Blue lines denote positive correlations, orange dashed lines negative correlations. Node size is proportional to relative abundance.

Potential hubs (≥4 positive edges) identified in slurry-amended soil include *mexF, vanR*, *muxB*, *muxC*, *rosB* and *efpA*. In the slurry network, only *adeJ* had more than three positive edges, while none was present in untreated soil.

Some concordance was present between soil sites. For example, *muxB* and *mdtB*, *muxB* and *muxC* as well as *vanR* and *rphA* are associated in both soil networks. The aforementioned associations were also cross-validated by significant Spearman's Rank correlations (fdr p-value <0.05).

2.4 Discussion

2.4.1 Slurry Resistome

With few exceptions, the core resistome of dairy cattle slurry in this study broadly reflects the findings of previous research, whether experiments considered solid manure, liquid slurry or involved cattle farming systems in different countries. For example, Kyselková et al. (2015a) identified that the tetracycline genes *tetW*, *tetQ* and *tetO* formed the core faecal resistome of dairy cattle, while *tetA*, *tetM*, *tetY* and *tetX* were also detected intermittently. In the current study, which used metagenomic methods rather than qPCR, both *tetW* and *tetQ* were in the top 25 ARGs detected in slurry based on average 16S-normalised abundances. Additionally, *tetM*, *tetA*, *tetQ* and *tetO* were also more abundant in slurry than soil irrespective of treatment. Similarly, other authors using metagenomic (Noyes et al., 2016, Zhou et al., 2016) and qPCR-based approaches (Fahrenfeld et al., 2014, Sandberg and LaPara, 2016, Wang et al., 2016a, Muurinen et al., 2017, Lin et al., 2019) identified *tetW* as one of the most abundant tetracycline resistance genes in cattle waste.

Cross-study agreement is not limited to tetracycline resistance genes; *cfxA* genes belonging to the Class A beta-lactamase resistance gene group have been frequently associated with cattle waste resistomes (Chambers et al., 2015, Zhou et al., 2016,

Muurinen et al., 2017). As has been reported elsewhere, little evidence was found in the current study to suggest many other beta-lactamases were highly prevalent in cattle waste (Muurinen et al., 2017, Zhang et al., 2017b, Qian et al., 2018), aside from *oxa* (class D beta-lactamase) which has been associated with cattle waste water in another study (Agga et al., 2015). A similar pattern was also observed in swine manure (Li et al., 2015a).

Startlingly, the high prevalence of *cfxA* genes appears to be unaffected by the restricted use of antibiotics employed by the study dairy farm and the Finnish dairy farm described by Muurinen et al. (2017). However, the extent to which increased antibiotic use impacts the resistome of cattle and farming environments remains equivocal (Oliver et al., 2011, Noyes et al., 2016, Rovira et al., 2019, Feng et al., 2020). It remains unclear whether the observed prevalence of *cfxA* genes reflects an endemic presence in the gut microbiome or implies limited antibiotic selection is required for proliferation.

With respect to sulphonamide resistance genes (namely the *sul* genes associated with mutations in dihydropteroate synthases), the results suggest they are consistently present in cattle waste, but generally occur at lower incidence than tetracycline resistance genes, with particular reference to genes encoding ribosomal protection proteins (RPPs). The fact that both Muurinen et al. (2017) and Noyes et al. (2016) found RPP tetracycline genes were typically more abundant in cattle waste relative to sulphonamide resistance genes offers support to this inference. On the other hand, while Wang et al. (2016a) similarly reported RPP tetracycline genes were more abundant than those tetracycline efflux pumps; sulphonamide resistance genes were shown to predominate over tetracycline genes in that study. It is possible these discrepancies reflect differences in antibiotic use, other aspects of animal husbandry practice, environmental factors or methodology.

Another study involving tetracycline and sulphonamide resistance genes found that *tetW* (RPP) had a greater average relative abundance in cattle slurry compared to *sul1* and *sul2*, while the occurrence of *tetG* (efflux) and *tetO* (RPP) was similar to the

sulphonamide ARGs assessed (Fahrenfeld et al., 2014). These findings are also in agreement with the results presented here. Finally, while Nõlvak et al. (2016) detected *sul1* at somewhat higher relative abundances than *tetA* (efflux) in cattle manure, the latter was the only tetracycline ARG evaluated. Likewise, in the current study, *sul1* had a slightly higher average abundance than *tetA*; with tetracycline genes as a group having a much higher average relative abundance than sulphonamides.

Macrolide, lincosamide and streptogramine (MLS) resistance genes represented the dominant antibiotic resistance category in cattle slurry metagenomes screened in the current study, followed by tetracycline resistance genes (based on average 16S-normalised abundance). MLS genes have been routinely identified among the more abundant resistance genes in cattle waste and associated run-off (Agga et al., 2015, Hu et al., 2016, Noyes et al., 2016, Zhang et al., 2017b, Gou et al., 2018). However, while one study suggests MLS resistance may be more abundant than tetracycline ARGs in cattle waste (Hu et al., 2016), others suggest *vice versa* (Zhou et al., 2016). On the other hand, other studies reported that aminoglycoside ARGs were dominant over MLS and tetracycline ARGs in cattle manure (Muurinen et al., 2017, Gou et al., 2018). Regardless, the generally high abundance of MLS, tetracycline and aminoglycoside ARGs in cattle slurry is widely evidenced, and some variability is to be expected due to both genuine site differences and experimental biases (e.g. the scope of qPCR arrays and DNA extraction kit choice).

Considering the different methods used in various studies, the overall consistency of these data suggest certain ARGs are indicative of cattle waste and therefore likely represent a resistome common to the cattle intestinal tract. It has been proposed that herd-wide core resistomes may exist irrespective of adult antibiotic treatment (Wichmann et al., 2014, Kyselková et al., 2015a) and that the core faecal resistome is largely dependent on host microbiota formed by diet during early development (Liu et al., 2019b). Further support for the concept of a core resistome was provided by Thomas et al. (2017) who demonstrated aminoglycoside ARGs were prevalent in the gastro-intestinal tracts of North American beef cattle regardless of whether animals had

received prophylactic antibiotic feed additives at the finishing stage. The characterisation and validation of the core cattle waste resistome is vital because it allows the development of gene-marker sets which can be used to robustly determine dissemination of ARGs in the environment in a standardised manner, while also enabling greater replication at reduced cost.

Indeed, despite the existence of notable parallels among previous publications it should be stressed that many of the aforementioned studies relied on qPCR arrays which were variable in size, ranging from several genes (Fahrenfeld et al., 2014, Kyselková et al., 2015a, Sandberg and LaPara, 2016) to hundreds (Muurinen *et al.*, 2017). While the validity of a targeted approach is often desirable, it is important to bear in mind that studies which assessed a very limited selection of ARGs without first conducting pilots with large arrays or alternatively, referring to studies which have done so are more likely to provide a distorted picture of what constitutes the 'core resistome'. In particular, this could have implications for which resistance genes in a given category are considered the best markers for a specific resistome. For this reason, comparisons with existing literature have focussed on the relationship between groups of resistance genes, rather than the diversity of the cattle waste resistome as a whole.

For example, although *tetW* was among the most abundant resistance gene groups detected in slurry, *tetM* was the most prevalent tetracycline subgroup, followed by *tet44*, *tetT* and *tetW*. A review of existing publications shows a number did not include some (Fahrenfeld et al., 2014, Sandberg and LaPara, 2016), or all of these highly abundant genes (Hu et al., 2016, Nõlvak et al., 2016) in their analyses. Interestingly, Wang et al. (2016a) evaluated a range of tetracycline resistance genes in cattle manure including genes encoding RPPs: *tetM*, *tetO*, *tetQ*, *tetT* and *tetW*, as well as efflux pumps: *tetA*, *tetC* and *tetG*. Not only did Wang et al. (2016a) provide further evidence that RPP tetracycline genes are typically preponderant over efflux pump tetracycline resistance genes in cattle waste, the data also illustrates that studies which only considered tetracycline efflux pumps (Agga et al., 2015, Hu et al., 2016) (*tetA* and *tetB*) would miss this nuance. Other studies have also found evidence that tetracycline RPPs are generally

more abundant in cattle waste than tetracycline resistance genes encoding efflux pumps (Noyes et al., 2016, Muurinen et al., 2017). A similar trend has been reported in soil (Wang et al., 2016a).

In this regard metagenomic or high through-put qPCR can excel, although neither of these methods is exhaustive, being dependent on the quality of selected databases and probes respectively. Ideally, it may be most appropriate to initially obtain a small collection of relevant metagenomes, and then design a bespoke qPCR array based on the most relevant genes. Finally, virtually all current molecular methods suffer from compositionality and extraction bias. It is ultimately encouraging however, that overall patterns are broadly similar between qPCR and metagenomic studies with regard to the resistome of cattle waste.

In attempting to identify potential slurry markers it is also important to acknowledge that the most abundant resistance genes in slurry were not necessarily slurry-specific. The fact that the rifamycin resistance gene *rpoB2* was the most abundant ARG in all sample types infers it is likely endemic to the wider environment and therefore a poor marker for cattle waste contamination. Furthermore, the validity of annotating *rpoB2* based on short-reads is somewhat circumspect owing to the fact that resistance of this type is conferred primarily by point mutations in wild-type *rpoB* (Goldstein, 2014). A similar pattern of occurrence was evident for the polypeptide resistance gene *bacA*. Furthermore, the rifamycin category as a whole represented only a minor contribution to the slurry resistome relative to the most widely represented ARG categories. In contrast, resistance genes such as *aadA* (aminoglycoside), *tetM* (tetracycline) and *mefA* (MLS) were in the top 10 slurry ARGs, originated from the dominant categories and were comparatively less prevalent in soil metagenomes. Importantly, these examples are supported by both the average relative abundances across samples and differential abundance modelling with corncob.

In summary, in the present study, MLS, tetracycline and aminoglycoside ARGs dominate cattle slurry at the resistance gene group and category level; this is largely in agreement

with characterisations of cattle waste resistomes described in existing literature. The relationship between these categories is described as follows: MLS>tetracycline>aminoglycoside. This is reflected by the gene groups identified in the top 20 most abundant ARGs (16S-normalised average relative abundance): MLS (n = 8), tetracycline (n = 5) and aminoglycoside (n = 3). These data demonstrate candidate biomarkers for slurry contamination in the environment include the following MLS genes: *mefA*, *lnuC*, *lsaE*, *lnuB*, *lnuD*, *mphB*, *mefB* and *vatB*. Meanwhile candidate tetracycline genes include RPP encoding genes: *tetM*, *tet44, tetT*, *tetW* and *tet36*. Aminoglycoside candidate biomarkers are represented by: *aadE*, *aadA* and *ant(9)-1*. A limited selection of beta-lactamase resistance genes were identified as strongly associated with slurry, namely *cfxA* group (ambler class A) and *oxa* (ambler class D). Lastly, the peptide resistance gene *ugd* was frequently identified among most abundant ARGs in slurry, and was considerably less abundant in soil metagenomes.

2.4.2 Persistence of Slurry ARGs in Slurry-treated Soil

Many of the ARGs shown to be highly abundant in slurry metagenomes can confer resistance to antibiotic categories listed as highly (tetracyclines and lincosamides) or critically (macrolides, aminoglycosides and beta-lactam cephalosporins) important to human health by the WHO (Collignon et al., 2016). Furthermore, similar macrolide (*mefA*), tetracycline (*tet36*), aminoglycoside (*aadA11*) and cephalosporin (*cfxA2*) ARGs have recently been shown to dominate the resistome of untreated hospital wastewater (Petrovich et al., 2020), albeit at relative abundances an order of magnitude higher than the discussed slurry metagenomes. It is therefore essential that the persistence of these ARGs in the environment is comprehensively evaluated.

In this study, the long-term application of dairy cattle slurry to soil had limited potential to permanently elevate the abundance of slurry-associated ARGs above levels found in soil with no history of cattle waste amendment. This is consistent with a study which determined the prevalence of ARGs at a site which had received dairy manure over a

period of 30 years (Peng et al., 2017). Similarly, a qPCR-based study comprising 296 probes found cattle manure obtained from 'antibiotic-free' animals had a negligible impact on ARG abundance in soil (Zhang et al., 2017b). Many slurry-associated ARGs were not detected in soil from either site, or their abundance was not significantly influenced by a history of slurry amendment (Figure 2.8). On the other hand, a small subset of slurry-associated ARGs were consistently detected with greater frequency in treated soil, regardless of proximity to slurry application date.

In most cases only transient increases in slurry-associated ARGs were detected in soil immediately after slurry treatment (Figure 2.11A), and these genes (n = 23) returned to pre-treatment abundances within 56 days (Figure 2.11B). Nine MLS resistance genes were significantly increased within 24 hours of slurry application relative to samples obtained five days previously. Of these MLS ARGs, eight were closely associated with the slurry resistome, while three (*mefB*, *lnuB* and *lsaE*) are among the dominant ARGs in slurry which have already been suggested as appropriate biomarkers for indicating slurry contamination. Accordingly, five of these MLS ARGs (including the three suggested biomarkers) were never detected in soils from the untreated site. Temporary enrichment of select MLS resistance genes in cattle-waste amended soils has been documented in previous field (Muurinen et al., 2017) and microcosm (Chen et al., 2019a) studies. One such study showed the relative abundance of *mefA* and *lnuB*, which was significantly increased one day after manure application declined towards control levels within 120 days (Chen et al., 2019a). Muurinen et al. (2017) found the abundance of *mefA*, *lnuB* and *ermB* MLS ARGs returned to background levels within 14-42 days of cattle manure amendment, suggesting rapid attenuation. Similarly, a field study which used fluorescence in situ hybridization (FISH) to detect MLS resistance in soils amended with swine manure was unable to demonstrate persistent increases in this resistance type (Zhou et al., 2010). Furthermore, significant increases in *mefA*, *mefB* and *lnuB* in soil the following year (May 2018) were observed; 84 days after the first slurry application of the season (Figure 2.11C). Although this supports the idea that the abundance of ARGs

fluctuate in a cyclical manner, it also indicates the decline of MLS ARGs can take longer than 56 days.

One MLS gene, the macrolide efflux pump *oleC* was not detected in slurry, but was an order of magnitude more abundant in treated soil relative to control site soil. However, this gene did not significantly increase after slurry application. Whether this is due to innate differences in the soil resistome of the two sites, or indicates that slurry application indirectly selects for this gene over an extended period is unclear. Interestingly, Muurinen et al. (2017) found *oleC* was also consistently abundant in soil before treatment, was not frequently detected in dairy cattle manure and was not immediately affected by manure amendment. Unfortunately, no data from a site without a history of manure application was included in the aforementioned study for comparison. In any case *oleC* is evidently an ARG indigenous to the wider soil resistome.

The relative abundance of two slurry-associated tetracycline genes also increased significantly within 24 hours of slurry application. Both *tet36* and *tetT* are RPPs and represent strong indicators of slurry contamination, as shown by the analysis of slurry metagenomes. The persistence of these two genes is reminiscent of the MLS ARGs already described, with neither being significantly elevated in treated soil 56 days after exposure in 2017. Indeed, ephemeral spikes in *tet36* and *tetT* were reported in another study already mentioned in relation to MLS ARGs (Muurinen et al., 2017). However, as with select MLS ARGs, *tet36* remained elevated above pre-treatment levels 84 days after slurry application in the following year. Again, this indicates that while enriched ARGs can decay rapidly (<56 days), it is also possible for them to persist at elevated levels for longer periods.

Despite providing evidence that specific slurry-associated MLS and tetracycline genes can remain enriched within slurry-amended soil for at least 84 days, it remains apparent these ARGs are unlikely to become entrenched within the autochthonous soil bacterial community. Critically, the aforementioned MLS and tetracycline ARGs were undetected in any soil samples obtained in January 2018, suggesting these ARGs are largely eliminated

in impacted soil prior to slurry application recommencing in spring. Indeed, Muurinen et al. (2017) inferred the elevation of ARGs was transient and that elevated ARGs would decline towards background levels by winter. It should be noted repeated reference to this work has been made since few metagenomic studies have been conducted on cattle-waste amended soil in a European country with restricted antibiotic use, and while the authors employed qPCR, one of the more comprehensive ARG arrays with over 300 probes was used. Furthermore, the Finnish winter precludes the application of fertiliser during the winter months, thus providing a useful parallel to the NVZ restrictions in place at the UK dairy farm investigated in the present work.

Various mechanisms have been postulated to explain why the soil resistome may be resilient to change following manure amendment. One prominent school of thought is that the competitive action of indigenous soil microbiota constrains the dissemination and survival of ARGs. Existing laboratory-scale studies appear to support this theory (Peng et al., 2016, Chen et al., 2017, Klümper et al., 2019, Pérez-Valera et al., 2019). On the other hand, Udikovic-Kolic et al. (2014) reported that addition of cattle manure derived from animals with no prior antibiotic treatment induced the proliferation of beta-lactamase resistant soil bacteria (*Pseudomonas* sp.) for up to 130 days, indicating indigenous bacteria can also drive soil resistome enrichment. The decay rate of antibiotics and their sorption properties in soil may also contribute to the persistence or loss of ARGs in soil (See Cycoń et al., 2019, and references therein). However, factors influencing the rate of antibiotic degradation and persistence in soil are in turn determined by soil physiochemical properties (Srinivasan and Sarmah, 2014) and local climatic conditions (Kim et al., 2011, Joy et al., 2013, Cycoń et al., 2019), suggesting some soil textures (Blau et al., 2018) may possess enhanced or reduced 'resistome resilience'.

Since the slurry application window coincides approximately with the changing of the seasons (spring/summer: open; autumn/winter: closed), it is possible the rise and fall in the discussed ARG abundances relate to changing climatic conditions rather than slurry application. However, several slurry-associated ARGs were only detected in slurry

impacted soil. Furthermore, while the resistome of the untreated and slurry-impacted sites were similar in composition during the closed period (January 2018), they diverged antagonistically following the resumption of slurry application in May 2018 (Figure 2.4, Figure 2.5). Category-wide analysis of tetracycline ARGs in 2018 corroborates this inference, as the prevalence of tetracycline resistance genes were significantly increased relative to the untreated site in May 2018, but not in January 2018. Consequently, the corncob model indicates site differences in tetracycline genes are only significant when taking into account the period of sampling (site-month interaction $p = 0.0299$, site only $p = 0.5$).

Although clear differences in slurry-associated ARGs were detected immediately before and after slurry application in May 2017, the significant changes at the category level found 84 days after slurry application in May 2018 were not apparent. One potential explanation for this could be that the first slurry application of the season in 2017 took place in May rather than February, as in 2018. It is possible the environmental changes in soil moisture content and ambient temperature are responsible for the differences in ARG persistence; i.e. a change in the receptivity of the environment to entrained ARGs. Alternatively, these differences may correspond to variation in the resistome of the applied slurry, i.e. a change in system input. Even though the resistome was largely consistent across slurry samples at the category level, certain ARGs were identified with significantly greater frequency in the autumn and winter slurry libraries than in spring and summer (typically aminoglycoside ARGs such as *aadA*; Figure 2.10). The fact that slurry sampled toward the end of the open season contained greater abundances of genes like *aadA* might explain why these genes were increased after the earlier February application in 2018, and not in May 2017. This could be a compelling explanation were it not the case that select MLS ARGS including *lnuB* (also increased in May 2018) were significantly less abundant in slurry during the same period. Unfortunately, ARG data for the slurry applied in February 2018 are lacking, and therefore clarification is not possible. On the other hand, the environmental conditions in February 2018 relative to

May 2017 may have been more conducive to enabling particular ARGs to persist; namely reduced irradiation and topsoil temperature (top 10cm).

The effect of storing animal waste on antibiotic resistance is uncertain, with some studies indicating ARGs can increase in abundance during storage (Muurinen et al., 2017), while others have shown the response of ARGs in swine slurry during storage is gene-specific (Joy et al., 2014). Unpublished modelling data based on a cattle slurry microcosm experiment using slurry from the study farm revealed ARGs generally declined over time during storage.

To date, few studies appear to have collected data from farming environments across multiple seasons. Further testing combining controlled, multi-year microcosm studies and stored slurry/manure are likely to further explain the observed differences.

The temporal pattern of the slurry-associated ARGs previously discussed provides strong evidence that slurry ARGs rarely become embedded within soil and proliferate, implying these genes would quickly disappear from the soil resistome in the absence of annual slurry application. In contrast, a few slurry-associated ARGs appear to have become fixed in treated soil at levels significantly above their abundance in soil with no history of cattle slurry application. Specifically, these ARGs include the tetracycline resistance gene *tetM*, the peptide resistance gene *ugd*, and the multidrug efflux pump *mexT* (Figure 2.9).

According to corncob modelling, all three of these genes were significantly elevated in treated soil relative to the control site, and their greater abundance occurred independent of seasonal effects, indicating the prevalence of these ARGs is not dependent on regular slurry application. Indeed, of these genes only *ugd* significantly increased in May 2018 relative to January 2018, and no immediate increase in any of these genes was identified within 24 or 56 days of slurry application in May 2017. It is therefore possible these ARGs are intrinsically more abundant at the treated site for reasons other than slurry application. Alternatively, these 'autochthonous ARGs' could have been selected for over many years through the application of slurry, proliferating

over time until they reached an elevated, albeit stable 'carrying capacity' within the soil resistome.

Perhaps the most convincing evidence of this is displayed by the RPP encoding gene group *tetM*. Firstly, several studies have demonstrated the frequent occurrence of *tetM* in the excreta of cattle (Kyselková et al., 2015a, Zhou et al., 2016, Muurinen et al., 2017) and other livestock animals (Chee-Sanford et al., 2001, Jurado-Rabadán et al., 2014, Wang et al., 2016a, Muurinen et al., 2017). Accordingly, *tetM* is one of the most commonly detected tetracycline resistance determinants in enterococci isolated from livestock (Aarestrup et al., 2000, Roberts, 2005); a genus typical of mammalian intestinal microflora (Byappanahalli et al., 2012). The *tetM* ARGs have also been identified in other faecal indicator organisms, such as *E. coli* (Bryan et al., 2004, Jurado-Rabadán et al., 2014). Furthermore, a metagenomic study revealed *tetM* genes are generally more abundant in livestock waste, human faeces and sewage treatment plant influent than in soils and watercourses (Li et al., 2015a).

However, *tetM* has a broad host range of Gram positive and Gram negative bacteria which are widespread in the environment (Roberts, 2005, Van Hoek et al., 2011). This is often attributed to the association of *tetM* with promiscuous mobile genetic elements, namely *Tn916* and *Tn916*-like transposons (Agersø et al., 2006a, Roberts and Mullany, 2011, Ciric et al., 2013). Consequently *tetM* has been found in agricultural soil with no recent history of animal waste application (Dungan et al., 2019) and garden soil (Agersø et al., 2004).

Despite its detection in undisturbed populations, various studies have shown *tetM* is increased in animal waste-impacted sites relative to untreated soils (Agersø et al., 2004, Agersø et al., 2006b, Kyselková et al., 2015a, Peng et al., 2017, Dungan et al., 2019). The studies carried out by Peng et al. (2017) and Dungan et al. (2019) are of particular relevance as they both used sites subject to long-term fertilisation with dairy cattle manure. This corresponds with the present study, which shows *tetM* genes are increased in soils with a history of repeated exposure to dairy cattle waste.

Evidence therefore suggests that although *tetM* is not uncommon in soil, application of manure can facilitate the acquisition and subsequent enrichment of these genes in the resident soil bacteria, likely through conjugal transfer from introduced bacteria as proposed by Agersø and colleagues (Agersø et al., 2004, Agersø et al., 2006b). This is further supported by studies which successfully demonstrated or indicated conjugal transfer of *tetM* between introduced donors and resident soil bacteria in microcosm experiments (Natarajan and Oriel, 1992, Andrews Jr et al., 2004). In contrast, *tet36* ARGs which are known to have a limited host range relative to *tetM* (Whittle et al., 2003, Roberts, 2005) were only sporadically detected in treated soil within the slurry application season and were undetected in samples derived from the untreated site. On the other hand, *tetW* has a reasonably broad host range but had a similar pattern of occurrence to *tet36*, implying factors other than transfer potential determine the distribution of these gene groups.

Finally, it should be apparent that there is a considerable body of literature dedicated to the prevelance of *tetM* in soil, whereas the same cannot be said for the other two slurry-associated gene groups shown to be more abundant in treated soil; *ugd* and *mexT*. Discussion of *mexT* is avoided as it is a multidrug resistance gene and the developers of DeepARG acknowledge the algorithm may conflate gene annotations where no clear antibiotic category is evident (many multidrug resistance genes share tracts of sequences with each other and other genes unrelated to antibiotic resistance function). With regard to the *ugd* polypeptide resistance gene group, there have been few studies of its prevalence in soil, manure-contaminated or otherwise. However, a metagenomic study reported the presence of *ugd* in a metagenome assembled genome (MAG) recovered from hospital waste-water influent (Petrovich et al., 2020). The extent to which animal waste application impacts the abundance of the *ugd* gene group requires further investigation.

Taken together, these findings indicate that slurry-associated ARGs can become permanently enriched within autochthonous soil bacterial communities, though this rarely occurs, and tends to involve pre-existing soil ARGs with a broad host range.

Another finding of the current study worthy of mention is the transient increase of select beta-lactamase ARGs following the application of cattle slurry in May 2017. According to LEfSe analysis (Figure 2.11A), Class C beta-lactamases were significantly elevated in soil within 24 hours of slurry application relative to samples taken five days prior to treatment. Furthermore, Class C beta-lactamases were still significantly increased 56 days afterward, in addition to Class A beta-lactamses (Figure 2.11B). These ARGs were virtually undetectable in cattle slurry, as discussed in the previous section. Interestingly, other studies have described similar increases in the abundance of beta-lactamases in soil following application of cattle manure derived from animals with no history of antibiotic treatment (Udikovic-Kolic et al., 2014, Hu et al., 2016). As in the present study, Hu et al. (2016) detected low prevalence of most beta-lactamases, except select Class D ARGs. Both aforementioned studies therefore propose the observed increases in beta-lactamase genes related to the propagation of intrinsic soil bacteria, rather than the introduction of bacteria resident in cattle manure. Some caution should be exercised with regard to this interpretation when considering the present study, since no significant increases in beta-lactamases were discernible in May 2018. Although this could be explained by the earlier application date in 2018, Muurinen et al. (2017) was also unable to demonstrate a similar response.

It also is important to acknowledge that Class D beta-lactamases were more abundant in samples from the untreated site than the treated site (corncob model: $p < 0.01$). At first glance this result was unexpected, given *oxa* genes were common in slurry and could be considered a biomarker for slurry contamination. Indeed, a q-PCR-based study showed *oxa* genes could be enriched in Danish agricultural soils receiving swine manure relative to inorganically fertilised soil (Graham et al., 2016). Existing information on *oxa* genes could offer some explanation for this result. Firstly, *oxa*-type genes are a diverse ARG subtype with between 150 - >300 variants described to date (Walther-Rasmussen and Høiby, 2006, Poirel et al., 2010, Antunes et al., 2014). In addition, these variants can be further divided into those which confer narrow or broad-spectrum (carbapenem) resistance (Walther-Rasmussen and Høiby, 2006, Poirel et al., 2010); though the precise

composition of these two groups is subject to debate (Antunes et al., 2014). The variation of these genes is of note as DeepARG appears unable to distinguish between *oxa* variants, and so the composition of slurry, treated soil and untreated soil *oxa* genes could be very different in this regard. Furthermore, *oxa*-type genes have been identified in *Actinetobacter* sp. (Reviewed by Doughari et al., 2009), *Aeromonas* sp. (Janda and Abbott, 2010) and *Pseudomonas aeruginosa* (Walther-Rasmussen and Høiby, 2006) which are known to occur in the soil environment and represent a natural reservoir of these ARGs. However, this does not necessarily explain the greater abundance of *oxa* genes in the untreated site. Further research is required to clarify the extent to which these clinically important ARGs (namely the carbapenemases) are abundant in soils devoid of livestock contamination.

Finally, many studies have reported sulfonamide ARGs are abundant in soil and/or enriched by the addition of animal waste (Fahrenfeld et al., 2014, Nõlvak et al., 2016, Zhou et al., 2016, Muurinen et al., 2017, Lin et al., 2019) In contrast, the present study showed that while sulfonamide ARGs were consistently detected in cattle slurry they were seldom detected in treated soil and were not detected in any soil samples from the site with no history of slurry application. On the other hand, another metagenomic study was unable to detect sulfonamide resistance genes in soil, despite detection in cattle manure (Noyes et al., 2016). Likewise a study conducted on Chinese arable soil with long-term swine manure exposure found negligible evidence of sulfonamide resistance genes in treated soils (Cheng et al., 2019). It is likely a multitude of factors explain these differences. Firstly, variation in soil texture can influence the sorption of sulfonamide antibiotics (Thiele-Bruhn et al., 2004) and therefore modulate selection pressure. Secondly, on-site sulfonamide use is also likely to vary between farms. Lastly, the DeepARG database contains few sulfonamide ARGs and the curators acknowledge this may impair assignment to this category (Arango-Argoty et al., 2018). Given the complex nature of the soil matrix and influencing factors, it is perhaps unsurprising that data regarding animal waste-impacted soil is less consistent than findings confined to animal waste.

2.4.3 Network Analysis

Co-occurrence networks have been used in multiple publications to infer and visualise associations between ARG groups as well as MGEs and bacterial community members (Li et al., 2015a, Hu et al., 2016, Zhang et al., 2017b, Chen et al., 2019b). Here, focus is placed on ARG-ARG group associations, as ARG-taxon associations will be explored in subsequent Chapters.

The networks for each environmental location (slurry tank, treated field site and untreated field site) were assembled individually in order that comparisons could be made between different sites. The resulting networks are sparse, although it should be stressed that many ARG groups were excluded from network construction as they were detected in fewer than half the samples representing a given location. The decision to exclude these was taken to minimise the recovery of spurious associations. Reducing the exclusion threshold can have a dramatic effect on networks. For example, if the threshold is reduced to include ARG groups which occur in >30% of samples, nodes representing *tetT* and *mefA* can be found linked together in treated soil, but not untreated soil. This is worthy of note since these ARG group associations are also common to the slurry network and represent gene groups which were sporadically increased in slurry treated soils. Consequently, the reported networks are best considered to be representations of the core resistome. With this in mind, the utility of the networks was restricted to inferring dominant ARG-ARG associations which might distinguish between the sampled sites. (Figure 2.13). For example, a putative association between the rifamycin ARG group *rphA* and the glycopeptide ARG group *vanR* was independently identified in both soil networks yet these nodes were not connected in the slurry network. In addition, an association between bacitracin ARG group *rosB* and multidrug ARG group *oqxB* was identified in both soil networks. Similarly, *muxB* and *muxC* multidrug resistance gene groups co-occurred in both soils. Interestingly, *muxB* and *muxC* form part of a resistance nodulation cell division (RND)-type efflux pump, and so their association is not lacking biological basis (Mima et al.,

2009). In addition *mdtB* co-occurred with *muxB* in both soils, which corresponds with literature indicating the aforementioned ARG groups exhibit sequence homology and similar function (Górecki and McEvoy, 2020). On the other hand, the developers of DeepARG advise caution regarding the annotation of multidrug resistance genes owing to extent of homology between genes within this category. In any case, the fact that the machine learning model distinguished between these gene groups suggests annotations in this category are not entirely unreliable (given the shared sequence homology it might be expected the two groups would be conflated). It should be acknowledged that the lower connectivity of slurry and untreated soil networks relative to amended soil may be an artefact of smaller sample size. However, the detection of the same associations in two independent soil sites support the veracity of these networks.

The network analysis of slurry ARG groups inferred the co-occurrence of tetracycline efflux and sulfonamide ARGs (*tetY*, *tetA* and *sul2* respectively). Associations between these ARG categories have been previously reported (Gow et al., 2008). Finally, *cfxA* beta-lactamase gene groups are associated with each other and their co-occurrence is further demonstrated by the heatmap where they are assigned to the same cluster.

Overall, the networks display limited complexity, although this may be largely due to insufficient sequencing depth and the assumption of network sparsity implemented by SpiecEasi. In future studies a greater and equal number of samples might be taken from sites to improve both network resolution and comparability, however this is accompanied by considerable increase in cost and computational requirements. Despite these limitations the networks imply the presence of biologically plausible associations which are consistent in soil irrespective of slurry amendment.

2.5 Conclusion

In summary, the present study has shown that dairy cattle slurry harbours a range of ARG groups conferring resistance to clinically significant antibiotics, namely MLS, aminoglycosides and tetracyclines. Of further note, these findings largely corroborate cattle waste resistomes described in previous studies which have used assorted molecular techniques (in particular qPCR) and were carried out at farm sites which vary in antibiotic treatment regimen, geographical location and waste management. The extent of cross-study agreement therefore lends support to the output of the machine-learning algorithm used for putative ARG annotation whilst also suggesting the existence of a core resistome common to cattle waste irrespective of antibiotic use. The high prevalence of MLS ARGs in slurry despite a lack of recent MLS antibiotic use on the farm, as well as an inability to detect MLS antibiotics in the slurry underscores the difficulty in determining the drivers of AMR in real-world systems. As expected, analyses demonstrated strong overall separation between slurry and soil resistomes.

With regard to the impact of slurry application on the soil resistome, it is evident that the majority of the slurry-associated ARG groups fail to become either enriched or entrenched within the soil environment, as has been found in other studies examining sites with long-term exposure to cattle waste. The current study has shown slurry ARG groups are detectable in soil immediately after slurry application, though their presence is ephemeral, with most becoming undetectable in less than 8 weeks, though some may persist for longer. The transient nature of these periodic fluxes is supported by multiple ordination methods illustrating the divergence of the treated soil resistome relative to untreated soil during the open period of slurry application. However, the role of seasonal changes in environmental factors unrelated to slurry application requires clarification, ideally through a combination of controlled microcosm studies and multi-year surveys which are currently lacking. On the other hand, results indicate a small assemblage of slurry-associated ARG groups are more abundant in treated soil than soil with no history of cattle waste exposure; however, their relative abundance did not vary according to

season. Consequently, it is proposed that introduced ARGs with high mobility may occasionally become enriched and maintained throughout the year in resident soil bacteria as a result of long-term slurry application, while introduced ARGs with a restricted host range are rapidly attenuated. In addition, indigenous soil ARGs may also become enriched following long-term slurry application. Whether the accumulation of antibiotics plays a role in the maintenance of slurry-borne ARGs is uncertain, since antibiotic concentrations in soil were not measured in this study. It should be acknowledged that the application of 'whole' cattle manure which has not undergone the separation process may produce a different effect. It has been shown that separation may reduce ARG load (Oliver et al., 2020b). Lastly, other factors such as application method (disc-incorporation and shallow injection), soil type and application rate were not explored in this study and could influence bacterial and ARG survival.

While these results largely support the findings of microcosm and 'snapshot' studies which previously analysed data from multiple farms, few metagenomic studies have combined temporal analysis of soil and on-site slurry over an entire season. In addition, this study documents the impact of UK regulations implemented in a nitrate vulnerable zone on the soil resistome. Future research would benefit considerably from dedicated efforts to validate the use of machine learning algorithms to annotate ARGs in environmental metagenomes. The use of metagenomics in conjunction with qPCR could greatly refine the use of machine learning algorithms, allowing the gamut of publicly available metagenomes to be more confidently analysed. Ultimately this would harness a largely untapped resource and make multi-season, pan-farm analyses much more feasible.

Chapter 3

Evaluating Soil Bacterial Community Shifts Following Cattle Slurry Amendment

3.1 Introduction

While the primary focus of the previous chapter was to assess the diversity and prevalence of antibiotic resistance genes (ARGs), the current chapter is concerned with bacterial composition and temporal shifts in the surveyed environmental compartments. Although the over-arching aim of the work was to evaluate the risk posed by the spread of antibiotic resistance genes through fertilisation of grassland with dairy slurry, a grasp of bacterial populations in amendments and soil over time is vital as environmental resistomes have been shown to reflect bacterial community structure in both habitat and microcosm-scale studies (Forsberg et al., 2014, Li et al., 2015a, Hu et al., 2016).

Furthermore, the dissemination of enteric opportunistic pathogens present in animal waste represents an important nexus between the study of antibiotic resistance in the environment and both human and animal health (Burgos et al., 2005, Ibrahim et al., 2016, Leclercq et al., 2016). However, given the complexity of outlining bacterial responses in over 50 metagenomes, the interactions and associations between ARGs and bacterial taxa will be discussed in the next chapter.

In the broadest sense, the impact of animal manure on soil microbial communities has been, and continues to be, the subject of extensive research (Riber et al., 2014, Leclercq et al., 2016, Zhang et al., 2018b, Podmirseg et al., 2019, Coelho et al., 2020). The vast majority of studies fail to consider short- and long-term effects at the same site. For instance, there are studies which focus exclusively on effects over the short-term (weeks/months) (Rieke et al., 2018, Xiong et al., 2018) or the long-term (months/years) (Sun et al., 2015, Wang et al., 2016b). Those that attempt to consider both can be

single ‘snapshots’ of different (and fundamentally contrasting) sites which have been sampled at varying points since a given fertilisation and thus do not represent a true temporal continuum (Liu et al., 2017). Another such study examined how differing years since the last manure exposure affected the microbial communities within field plots (Zhang et al., 2018b).

Meanwhile, a large proportion of publications incorporating temporal analyses tend to be microcosm studies that are not always representative of the farming systems they intend to emulate. Although microcosms are valuable in that they allow manipulation of discrete variables, Chen et al. (2019a) remarked that numerous experiments have spiked animal waste with antibiotics in excess of what would realistically be present as a result of typical antibiotic treatment and this is likely to exaggerate, or otherwise distort, both resistome and microbiome responses. For example, Zhang et al. (2017b) acknowledged that the concentration of tylosin in their antibiotic spiked manure-amended soil microcosms was several orders of magnitude higher than concentrations reported in field-based studies. Furthermore, the dilution effect of field-scale fertilisation on manure-borne organisms is unlikely to be replicated in particularly small microcosms. Ideally, controlled microcosm experiments would be carried out in conjunction with field studies, however the practicality of this is likely to be prohibitive.

On the other hand, only a subset of studies employ a true metagenomic approach, using techniques such as T-RFLP (Abubaker et al., 2013) and DGGE (Blau et al., 2018, Podmirseg et al., 2019) which, while certainly valuable, exclusively provide information on dominant taxa. The most common molecular technique is 16S-rRNA amplicon sequencing which typically provides data to genus level although biases can be introduced through choice of 16S region (Bukin et al., 2019). Whilst metagenomics is not without bias, it does not rely on marker regions.

In addition, work examining the persistence of manure-borne organisms in soil typically focuses on very specific indicator organisms, which will miss relevant changes in the total bacterial community and groups of understudied bacteria.

Lastly, a paucity of research considers the impact of repeated manure exposure and the consistency of their effect throughout a progression of changing climatic conditions. Related to this, only Hodgson et al. (2016) appears to have evaluated how seasonal effects interact in a nitrate vulnerable zone (NVZ) despite their abundance in the UK. Indeed, approximately 55% of land in England is currently within a designated NVZ (GOV.UK, 2020a).

In light of this, much remains to be addressed and the present experimental design represents a unique opportunity to further bolster and refine existing understanding of bacterial community dynamics in a field receiving cattle slurry amendment. Therefore, the key aim of this chapter was to characterise changes in the soil bacterial community composition in response to repeated applications of cattle slurry and describe the key abiotic and biotic drivers associated with these changes.

3.2 Materials and Methods

3.2.1 Approach, Sampling and Sequencing

Taxonomic assignment and downstream analyses were carried out on the same metagenomic libraries discussed previously (Chapter 2). Full descriptions of sample sites, sampling strategy and sequencing are given in Chapter 2.

3.2.2 Soil Physiochemical Analyses

For extractable macro- and micro-elemental analysis, 1 g of soil was suspended in 9 mL of 1M NH_4NO_3 and mixed thoroughly by agitation using a rotary shaker for 1 hour. Subsequently, samples were centrifuged and 1 mL of the resulting supernatant was diluted in 9 mL of 2% nitric acid. Finally, samples were passed through a 0.22 µm filter

before being loaded for inductively coupled plasma mass spectrometry (ICP-MS; Thermo-Fisher Scientific iCAP-Q; Thermo Fisher Scientific, Bremen, Germany). The total extractable carbon (TC), total extractable organic carbon (TOC), and nitrogen (TN) were determined using a CN analyser (Model Shimadzu TOC-VCPH, PC-controlled high-sensitivity model) after extracting the soil samples with 0.5 M K_2SO_4 solution in the ratio of 1:5 and further diluting with ultra-pure water in the ratio of 1:10.

For pH, 1 g of soil was homogenised in 2.5 mL of water, before taking readings with a pH probe (Hanna pH-209). Moisture content was calculated gravimetrically after oven drying field moist soil at 105^{o}C until constant weight was obtained (24h) (Black, 1965).

3.2.3 Taxonomic Classification

Taxonomic classification of all 150bp short-read Illumina NovaSeq libraries was performed with Kaiju v1.7.1; a protein-based classification software (Menzel et al., 2016). Kaiju was run in 'greedy mode' allowing three mismatches. The *nr_euk* database was used (version date 25/06/2019, >100M sequences), which contains all protein sequences from the NCBI BLAST database associated with Archaea, Bacteria, protists, fungi and viruses.

An already diverse array of taxonomic classification tools continues to expand as bioinformaticians respond to the increasing number and scope of metagenomic studies (Breitwieser et al., 2019). Furthermore, software selection is not trivial since reports have shown it can have clear implications for downstream interpretation of data (Sczyrba et al., 2017, Siegwald et al., 2019). With no universally superior program, optimising the choice of tools for classifying metagenomes is a substantial undertaking, which typically hinges on the environmental origin of samples, the specific research questions being investigated and the computational power available to researchers. However, a number of recent attempts have been made to independently benchmark as

many as 20 taxonomic classification tools (McIntyre et al., 2017, Gardner et al., 2019, Simon et al., 2019).

Chiefly, these reviews have aimed to determine the precision, recall and computational efficiency of software. Where recall metrics aim to reflect the extent to which classifiers capture all the unique species genuinely present within a sample, precision metrics place emphasis on the prevalence of false positives; together these metrics provide an indicator of overall classifier performance (Gardner et al., 2019, Simon et al., 2019). The relationship between these two metrics is typically antagonistic, with classifiers exhibiting high precision at the expense of recall and vice versa (Gardner et al., 2019). It should be noted that benchmarking studies rely almost exclusively on simulated datasets (McIntyre et al., 2017, Gardner et al., 2019, Simon et al., 2019), which while necessary is important to acknowledge when considering 'real world' metagenomes. Finally, the continual expansion and revision of reference databases confounds many comparisons and limits their future relevance.

Of the more recent assessments, Kraken is generally ranked favourably (Gardner et al., 2019, Simon et al., 2019). Kraken uses a k-mer-based system for classification, which is an approach widely represented among currently available classification tools (Breitwieser et al., 2019). Unlike early local alignment methods, a k-mer-based approach involves screening a reference database against sample library sequences for exact matches of a predefined nucleotide length referred to as k (default k = 35 in the case of Kraken2); this improves computational efficiency when handling large read libraries. The accuracy of abundance profiles can be further increased using Bracken, a companion program to Kraken which uses Bayesian probabilistic modelling to re-estimate relative abundance (Lu et al., 2017).

Although Kraken/Bracken has been shown to perform well in terms of recall and precision when analysing synthetic communities, reviews attempting to assess the suitability of specific classifiers for complex soil communities are lacking. The immense difficulty of simulating a metagenome for benchmarking purposes which approaches an

accurate representation of soil - a habitat which is widely regarded as both highly heterogenous and microbially diverse (Roesch et al., 2007, Myrold et al., 2014), is one possible explanation. Furthermore, soils are purported to encompass a large number of low-abundance species which often possess few or no known representatives in reference databases (Menzel et al., 2016). As such, estimated recall based on synthetic community evaluations may not translate well in practice where soil metagenomes are concerned.

Unlike Kraken, Kaiju is another classification tool, which relies on a pre-configured protein database rather than a nucleotide database. The developers of Kaiju indicate a protein-based classification system could improve recall and reduce the impact of sequencing errors. These advantages are attributed to the enhanced conservation of translated protein-coding regions (relative to non-coding regions) and the redundancy of genetic code (Menzel et al., 2016). Indeed, Menzel et al. (2016) demonstrated Kaiju could classify a considerably larger proportion (~40% vs 20%) of real soil metagenome reads compared to Kraken.

On the other hand, Simon et al. (2019) found that Kaiju typically classified fewer synthetic reads than Kraken, and reported that even though the overall performance of Kaiju largely paralleled Kraken/Bracken, the former was disproportionately susceptible to a false positive classifications below 0.01% relative abundance.

It has been suggested that low abundance (>0.1 or >0.01% relative abundance) taxa can be filtered to improve precision, however reviews have also cautioned that in soils, where many species are estimated to be present below 0.01%, such a strategy risks discarding a large number of true positives (McIntyre et al., 2017). The retention of rare taxa may also have relevance for associating ARGs with taxa entrained in soil via animal slurry application, since these bacteria may represent a miniscule proportion of the total microbiome and recovered reads.

In an attempt to address the challenge of balancing the need for high recall and precision in the current study, all metagenomic libraries were processed with both Kaiju

v1.7.1 and Kraken2 v2.0.8-beta/Bracken v2.5.0. Every effort was made to ensure the most contemporary, comprehensive and accurate databases were used with each tool. It should be noted that the RefSeq database (including 'complete' and 'representative' sequences for Archaea, Bacteria, protists, fungi, viruses and the human genome) used for Kraken (50-60M nucleotide sequences) was somewhat smaller than the protein database used for Kaiju (>100M protein sequences). A larger database including incomplete/draft genomic sequences is available for Kraken, however this was not used in favour of RefSeq-curated sequences due to limited compute resources. Perhaps unsurprisingly, Kraken only classified 9.31% ± 0.28 of slurry and 26.21% ± 0.19 of soil sample reads to species level, whereas Kaiju assigned between 30.26% ± 0.44 and 45.34% ± 0.13, respectively.

It is therefore likely that Kraken substantially underrepresented species richness (across all domains: n = 4126 ± 6, 4086 ± 3 slurry and soil respectively), while Kaiju classified a much larger contingent (n = 19811 ± 21, 18638 ± 73 slurry and soil respectively). Interestingly, a large pyrosequencing survey by Roesch et al. (2007) indicated soil microbiomes contain between 26000-53000 unique OTUs (operational taxonomic units).

Even so, without knowing the "ground truth" of species in each sample, it was decided that differential taxa reported by both Kaiju and Kraken should be compared to assess the potential impact that Kaiju's false positive rate may have on data interpretation. Comparisons of taxa shared by both databases at both genus and species level were largely in agreement, with most disparities emerging at family level due to differing database structure as opposed to estimated relative abundances. The output of Kaiju was therefore used for the final analyses in place of Kraken2/Bracken.

Future soil metagenome studies would benefit from using an equivalent database compatible with Kraken2/Bracken for evaluation. The challenge regarding false positives has also been acknowledged by the team behind Kaiju, who have since released a revised program Kaiju-Core, which aims to address the issue further (Tovo et al., 2020).

3.2.4 Data Analysis and Statistical Methods

3.2.4.1 Determining Bacterial Associations by Site: Data Exploration

To enable consistent inferences across taxonomic ranks, all reported analyses focus exclusively on reads assigned to species level and above. Preliminary analyses indicated that the exclusion of reads which did not meet this criterion had minimal impact on overall inter-sample relationships and did not alter dominant taxa. Site-based taxonomic data exploration was carried out using many of the techniques outlined previously (Chapter 2), with minor modifications. In brief, PCA was carried out on genus and phylum level centre-log ratio (CLR) transformed count data as suggested by Gloor et al. (2017). Visualisation with tSNE was also applied to CLR transformed count data, while NMDS was performed on relative abundances (taxon-specific counts as a proportion of total reads), since the Bray-Curtis dissimilarity index is unable to handle the negative values produced by CLR transformation. One limitation of the Bray-Curtis dissimilarity index is that while abundances are evaluated, the phylogenetic relatedness of taxa is not considered. However, weighted UniFrac distance takes into account both abundance and phylogenetic relatedness (Lozupone et al., 2007). To further inspect the relationship between sample groups, weighted UniFrac distances were calculated for a) all sites and b) soil samples collected in 2018 using the R package GUniFrac v1.3. (Chen and Chen, 2018). In order to calculate the UniFrac distance it is necessary to provide a phylogenetic tree containing tips representing each identified taxon. This is can be based on 16S rRNA reference sequences (Lozupone and Knight, 2005). Due to the large number of unique species identified by Kaiju, which included many uncultured organisms for which a curated 16S rRNA sequence was not readily available it was deemed impractical to construct a phylogenetic tree based on these data. Instead, the smaller Kraken classification dataset was screened against a 16S rRNA reference tree based on a sequence database downloaded from https://github.com/bowmanjeffs/paprica (accessed September 2021). Of the 4274 unique taxa identified by Kraken it was possible to match 3484 tips in the reference tree (>80%), the latter of which were used in UniFrac

calculations. Weighted UniFrac distance matrices were visualised via principal coordinate analysis (PCoA) in PRIMER v6 (Clarke and Gorley, 2006). Analysis of multivariate homogeneity of variance and PERMANOVA was also performed with the PERMANOVA+ add-on for PRIMER (Anderson et al., 2008).

An equivalent analysis was not performed for ARGs as there is no universally shared sequence region that could be used as a marker for comparisons across gene families and resistance categories.

3.2.4.2 Physiochemical Associations With Taxa

Physiochemical properties were initially explored by site with PCA. Based on site differences indicated by Corncob, relationships between key physiochemical properties and taxa (phyla-level only) were subsequently analysed with Spearman's Rank correlation testing using CLR-transformed count data (only correlations with fdr-adjusted $p < 0.05$ were considered significant).

3.2.4.3 Differential Abundance Analysis

Differential abundance analysis was performed on genus and phylum level data using Corncob as described previously (Chapter 2). Select species of interest were also individually modelled (e.g. faecal indicator organisms such as *E. coli*). Factors including site and season were considered. Feature selection was carried out as described previously for ARGs, however, the resulting matrix effectively duplicated the clustering patterns shown within the previous chapter and is therefore not shown here. Briefly, this suggests it is possible identify both ARG and taxon biomarkers which distinguish slurry, slurry-impacted soil and untreated soil from each other (see results and discussion for candidate biomarker taxa based on differential abundance analysis).

3.2.4.4 Changes in Taxon Abundance Following Slurry Application

As with the equivalent analysis of ARGs (Chapter 2), the immediate impact of slurry application on soil taxa was investigated by examining samples collected from the slurry-treated site between the first and second slurry application in 2017. However, Corncob was used in preference to LEfSe. While the transient occurrence of many ARGs during the aforementioned period prevented effective modelling of ARGs with Corncob, bacterial taxa typically exhibited consistent occurrence (but variable abundance) across samples. It was therefore possible to model changes in taxa at genus and phylum level using Corncob.

3.2.4.5 Bacterial Richness and Diversity Estimations by Site

Richness and alpha diversity estimates were calculated at the species level using iNEXT as described previously (Chapter *2*). However, for beta-diversity the weighted UniFrac distances were used in preference to Bray-Curtis distances (UniFrac method previously described in this chapter). This decision was taken because the incorporation of phylogenetic relatedness by UniFrac was deemed more biologically informative. Pair-wise comparisons for diversity measures were carried out as described in Chapter 2.

3.2.4.6 Bacterial Network Construction

Phylum level networks were constructed using SpiecEasi as previously described (Chapter 2). Genus-level network analysis could not be carried out using SpiecEasi due to compute memory requirements. Alternatively, a global (i.e. samples from all sites) genus-level network was constructed with CoNet (Faust and Raes, 2016) and visualised with Cytoscape (Smoot et al., 2011). The CoNet network was constructed with the following settings: genus minimum occurrence = 10 samples, Spearman's Rank R >0.8, fdr-adjusted p <0.05, bootstrapping = 100 iterations.

3.3 Results

3.3.1 Determining Bacterial Community Composition by Site: Data Exploration

Across the entire dataset, Kaiju classified a total of 26356 unique species belonging to 134 phyla. Of these unique species, 7.64% were assigned to a single read within the dataset. Most species were shared among all three sample sites (76.14%), while a small proportion (11.26%) were specific to a single site. Few genera (2.58%) were site-specific and no discriminatory phyla were observed (Figure 3.1). Rarefaction curves (produced with vegan) showed unique species recovery plateaued for all samples, irrespective of site (Figure 3.2). Sequencing depth was therefore deemed sufficient to characterise these environments based on the reference database used, although a large proportion of soil (<60%) and slurry reads (<75%) remained unclassified to species level (Figure 3.3A).

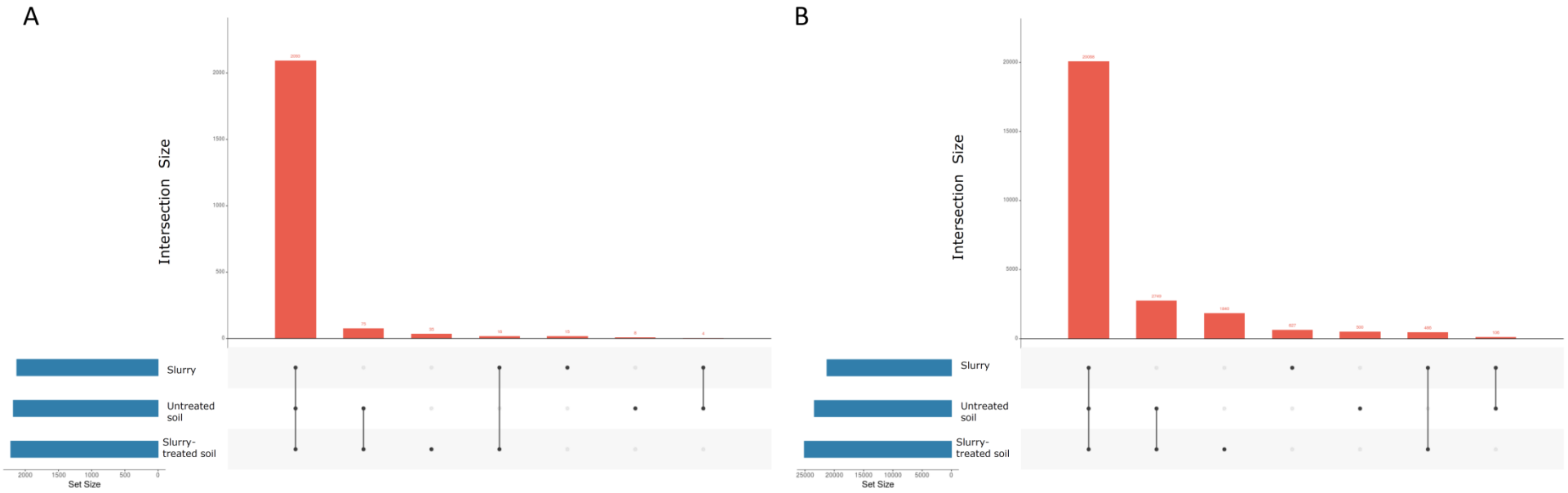


Figure 3.1 UpSet plot created using Intervene, illustrating the extent to which taxa were shared across sample types (detected or undetected). Genus (A), species (B). Solid vertical lines connecting sites on the x-axis denote taxa detection across multiple sites. Single dots on the x-axis denote taxa which were detected in only one site. Set size relates to the total number of unique taxa identified within a given site.

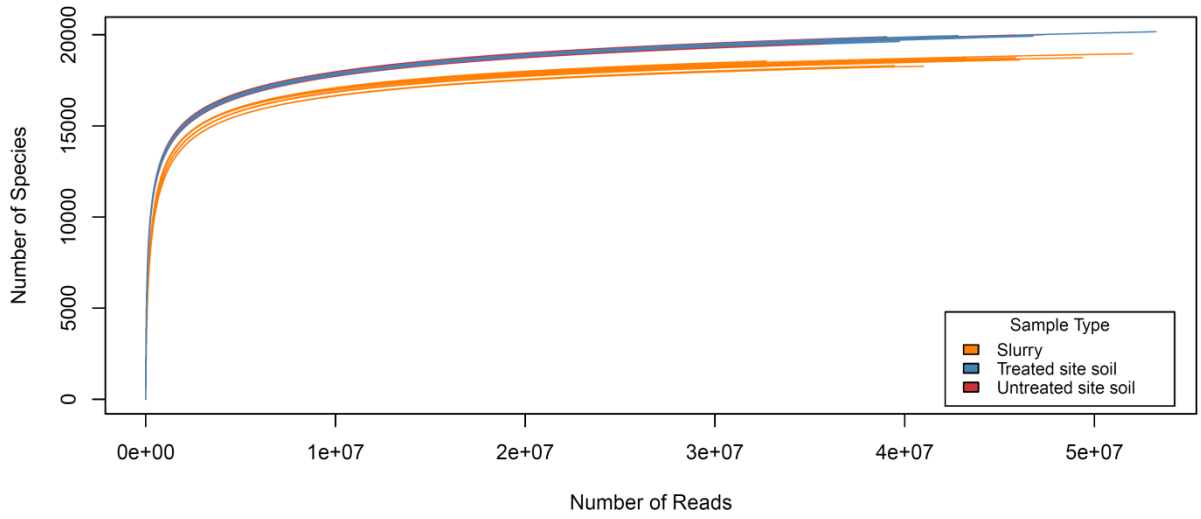


Figure 3.2 Rarefaction curves based on the number of species identified by Kaiju. Produced in R with rarefy function in vegan.

Differences in overall bacterial composition were explored by identifying dominant taxa according to site. Dominant taxa were considered as those constituting the top 10 most abundant taxa at any given taxonomic rank (dominant phyla and genera are shown in Figure 3.3B and C). Clear distinctions were demonstrated between the classified fraction of soil and slurry reads at both phylum and genus level. Specifically, the most prevalent phyla identified in slurry were Firmicutes, Bacteroidetes and Spirochaetes, respectively. In contrast, both soil sites were primarily represented by Proteobacteria, Actinobacteria and Acidobacteria. Actinobacteria formed a larger proportion of reads relative to Acidobacteria in slurry-treated soil, while the reverse was displayed in untreated soils. Untreated and slurry-treated sites were further distinguished by the ninth dominant phyla, which were Nitrospirae and Gemmatimonadetes respectively. On average, the top 10 dominant phyla comprised >90% of all reads assigned to species level irrespective of sample origin.

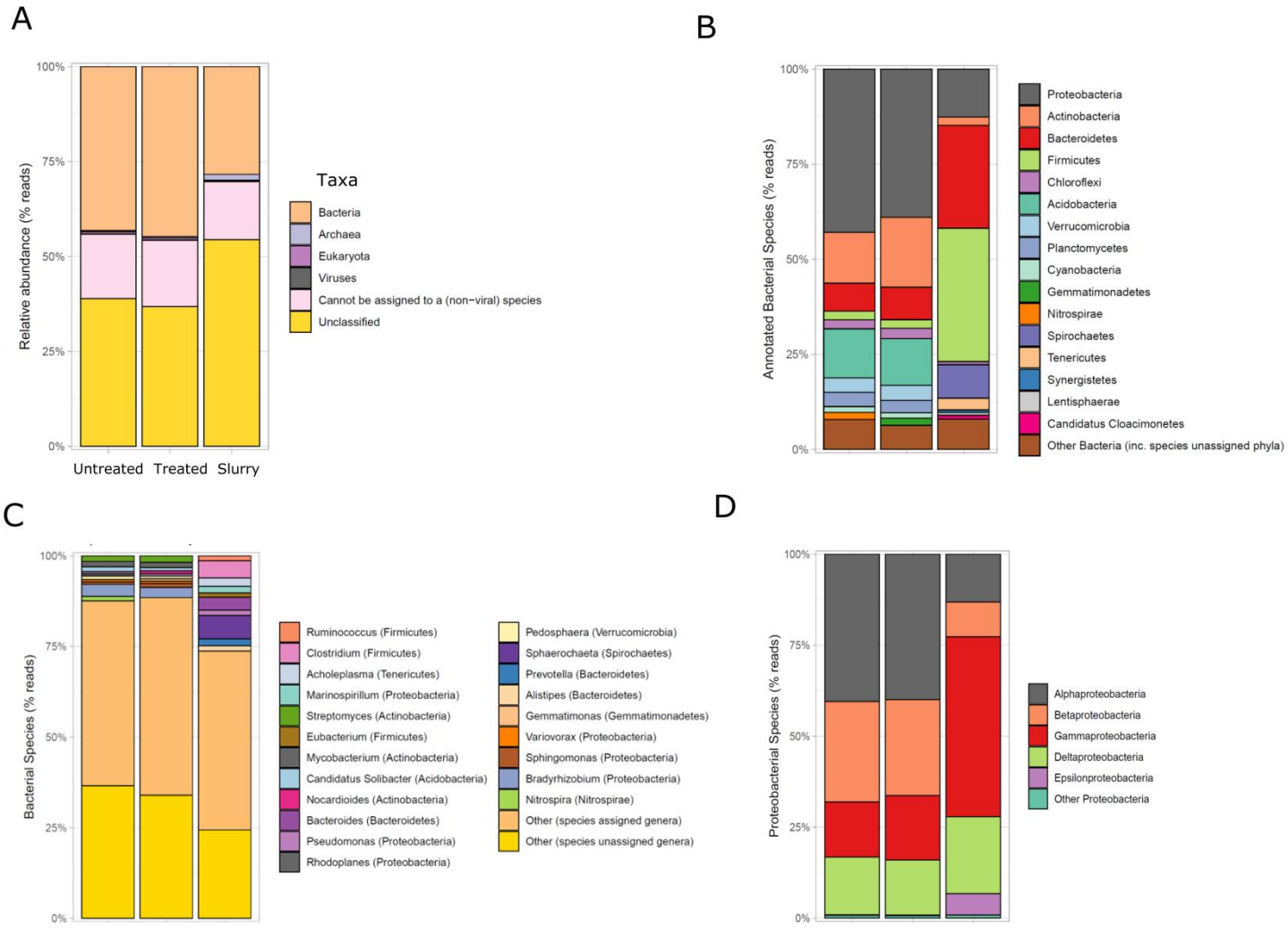


Figure 3.3 Stacked barcharts detailing taxonomic composition of reads. Untreated refers to field site soil with no history of slurry application, treated refers slurry-impacted site soil and slurry refers to slurry from the slurry tank. Breakdown of all reads by domain, including those unclassified by Kaiju (A). Top 10 phyla associated with bacterial reads assigned to species level (B). Top 10 genera associated with bacterial reads assigned to species level (C). Summary of Proteobacterial class based on reads assigned to Proteobacterial species (D).

At the genus level, the distinction between soil and slurry samples was also pronounced; only a single genus (*Pseudomonas*) was common to the top 10 dominant genera in soil and slurry samples (Figure 3.3C). Slurry samples were characterised by the preponderance of *Sphaerochaeta*, followed by *Clostridium* and *Bacteroidetes*. Meanwhile, the principal genera in soil samples included *Bradyrhizobium*, followed by *Streptomyces* and *Mycobacterium*.

The differences between dominant phyla in soil sites were also discernible at genus level. The inclusion of *Nocardioides* (Actinobacteria) and *Gemmatimonas* (Gemmatimonadetes) among dominant taxa in treated soil and their lower relative abundance in untreated soil reflects this. Likewise, *Nitrospira* (Nitrospirae) formed a larger proportion of reads in untreated soil relative to treated soil.

On average, the top 10 dominant genera contributed <12.5% of all reads assigned to species level in soil, whereas the equivalent genera formed more than 25% of these reads in slurry samples. The differing proportion of bacterial species associated with dominant genera likely alludes to a more diverse community in soils relative to slurry.

PCA of bacterial phyla and genera depicted largely similar trends regarding site. As shown in Figure 3.4A and B, data are most notably split along PC1 (75.4% and 78.8% variance for phyla and genera, respectively), with slurry clearly separated from the two soil sites. In addition, the variance along PC2 of the untreated site was consistently greater than that of the slurry-impacted site for both phylum and genus level.

In contrast, the variance of slurry along PC2 was substantially reduced for genera compared to phyla, while the opposite was apparent for treated soil samples. The separation between slurry-impacted and untreated site samples was also more pronounced at genus than phylum level (discrete clustering at 90% confidence), although the explained variance on PC2 was marginally reduced (3.9% and 7.0% variance, respectively). Lastly, the position of treated soil and slurry samples on PC2 were closer at the genus, rather than phylum level.

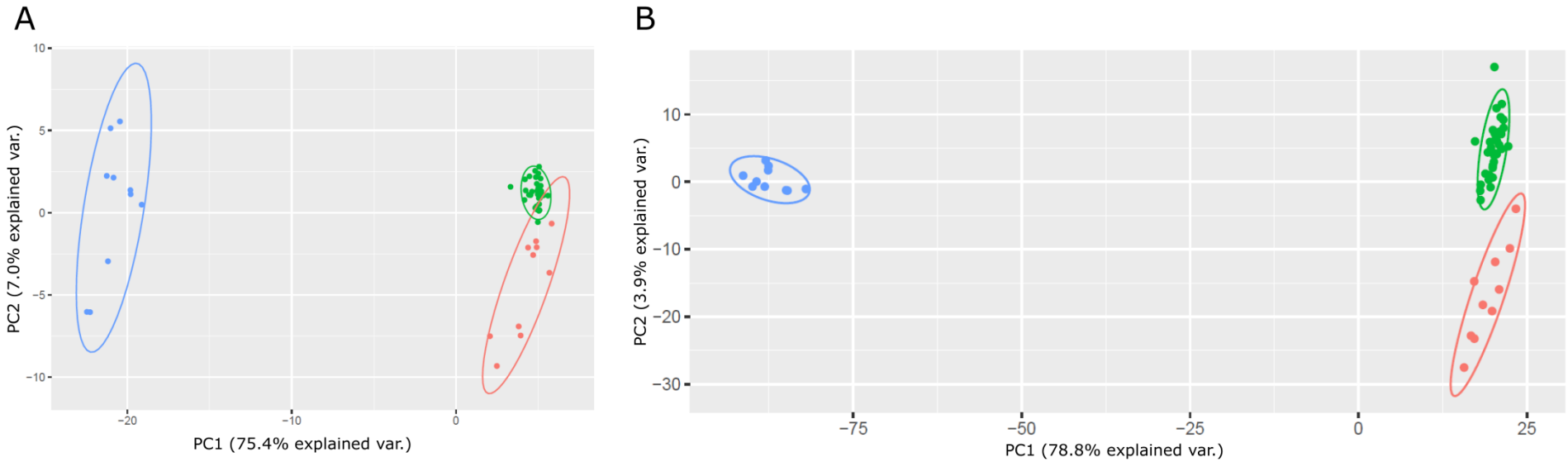


Figure 3.4 PCA of CLR (centre log ratio)-transformed data for phyla (A) and genera (B). Data are scaled and centred. Ellipses represent 0.9 probability. Colours represent samples; blue (slurry), green (slurry-treated soil) and orange (untreated soil).

NMDS of phylum and genus-level soil data for 2018 illustrates samples were clustered by site (Figure 3.5). In addition, treated soil samples were more tightly clustered than those from the untreated site (dispersion only significantly different for genus level data, $p < 0.04$), with potential site-month interactions also visible. Homogeneity of variance within site and month factors were confirmed in phylum level data only. PERMANOVA of phylum level data confirmed site and site-month interaction effects ($p < 0.01$ and 0.037, respectively). Ordination plots of bacterial composition were therefore somewhat consistent with equivalent analysis of resistome data (Chapter 2). Of particular note, Bray-Curtis distances for bacterial genera and ARG subtypes indicate significantly greater dispersion is present among soil samples collected from the untreated site relative to those from the slurry impacted site.

Overall, tSNE analysis was consistent across phyla and genera (Figure 3.6), and reflected the global structure of PCA. Accordingly, all sample types were discretely clustered, with spatial separation greatest between soil and slurry samples. This pattern was preserved at all iterations and perplexities evaluated and suggests a reliable representation of the data. In simple terms, the perplexity value controls how much emphasis is placed on preserving local or global data structure in two dimensions (Wattenberg et al., 2016). By using a range of perplexity values is it possible to establish the relative stability of the visualisations.

PCoA of UniFrac comparisons for both soil sites and slurry indicate a clear separation of soils and slurry across PCO1 (97.4% variance) and minor separation of soil groups over PCO2 (1.1% variance) (Figure 3.7A). This largely reflects patterns previously shown by PCA. Group differences were further evaluated by PERMANOVA (site, $p < 0.001$) and pair-wise tests (all site combinations $p < 0.001$). This suggests the taxa which distinguish between samples groups are phylogenetically distinct. Heterogeneity of variance was detected between slurry and both soil sites (PERMDISP, $p < 0.001$), but not between soils. Interestingly, slurry possessed the greatest within group standard error (± 0.005), which alludes to more pronounced differences in the abundance and phylogenetic relatedness of detected taxa between batches of slurry relative to soil,

where within group differences may correspond to more closely related taxa (based on the placement of reference 16S rRNA genes).

PCoA of UniFrac distances for soil data collected in January and May 2018 (Figure 3.7B) exhibit similarities to NMDS/Bray-Curtis distance plots for Kaiju data. Specifically, samples clustered by site (PERMANOVA, $p < 0.001$) and displayed possible site-month interactions (PERMANOVA, $p = 0.007$).

Heterogeneity of variance was identified between samples collected in January and May (PERMDISP, $p = 0.002$), with more variability in UniFrac distances exhibited between soil samples acquired in May relative to January. This is not surprising considering the increase in biological activity which typically accompanies the onset of spring.

Homogeneity of variance was identified across untreated and slurry-impacted soil (PERMDISP, $p > 0.2$), which contrasts with genus-based NMDS/Bray-Curtis distances for Kaiju genera (Figure 3.5B shows greater variability in untreated soil 'ARB' than slurry-impacted soil 'F31'). This may suggest the differences in dispersion based on Bray-Curtis distances correspond to taxa which are not phylogenetically distant.

However, it should be noted that the UniFrac distances presented are based on a subset of the Kraken count data and that Kaiju classified many more reads. It is not possible to determine the significance of omitted taxa and consequently comparisons with ordinations relying on the Kaiju data should be treated with caution. Nonetheless, this analysis does not deviate strongly from the PCA and NMDS/Bray-Curtis distance outputs based on the Kaiju data and therefore provides phylogeny-based evidence to further support them.

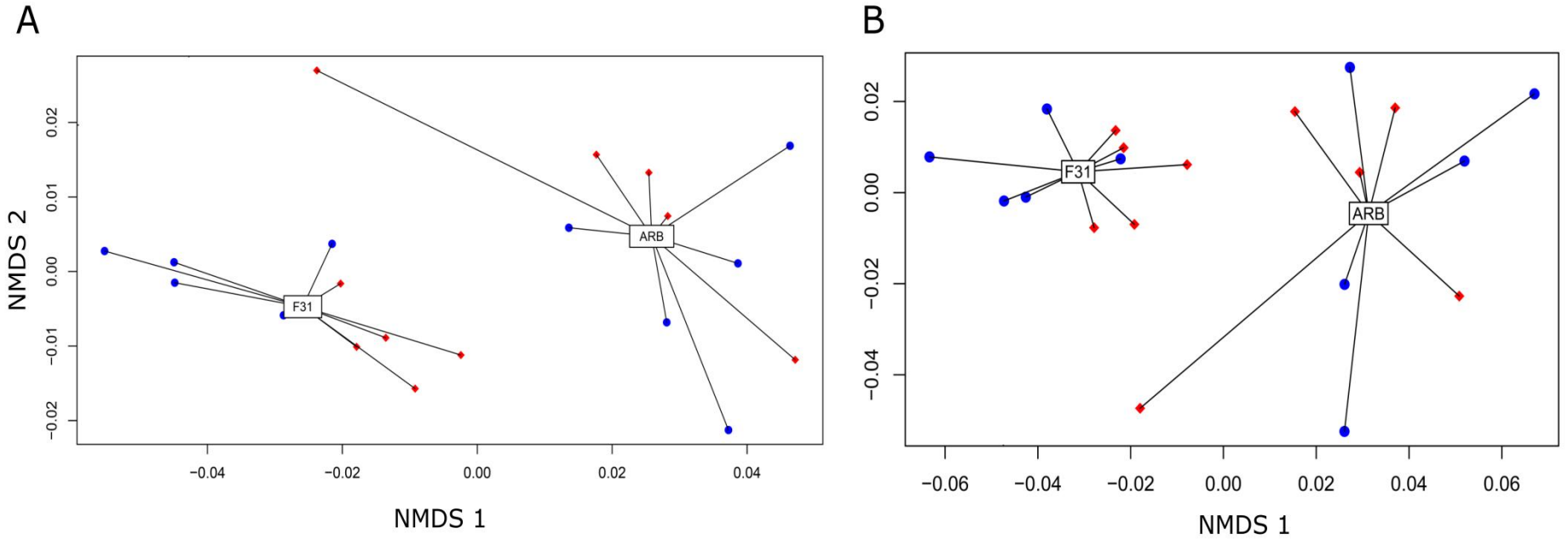


Figure 3.5 Taxa NMDS plots for slurry-impacted (F31) and untreated site (ARB) in January and May 2018. Phyla (non-metric R^2 = 0.995, linear R^2 = 0.977, stress = 0.07) (A), genera (non-metric R^2 = 0.995, linear R^2 = 0.975, stress = 0.07) (B). Colour/shape corresponds to month of sample collection; red/diamond = January, blue/circle = May.

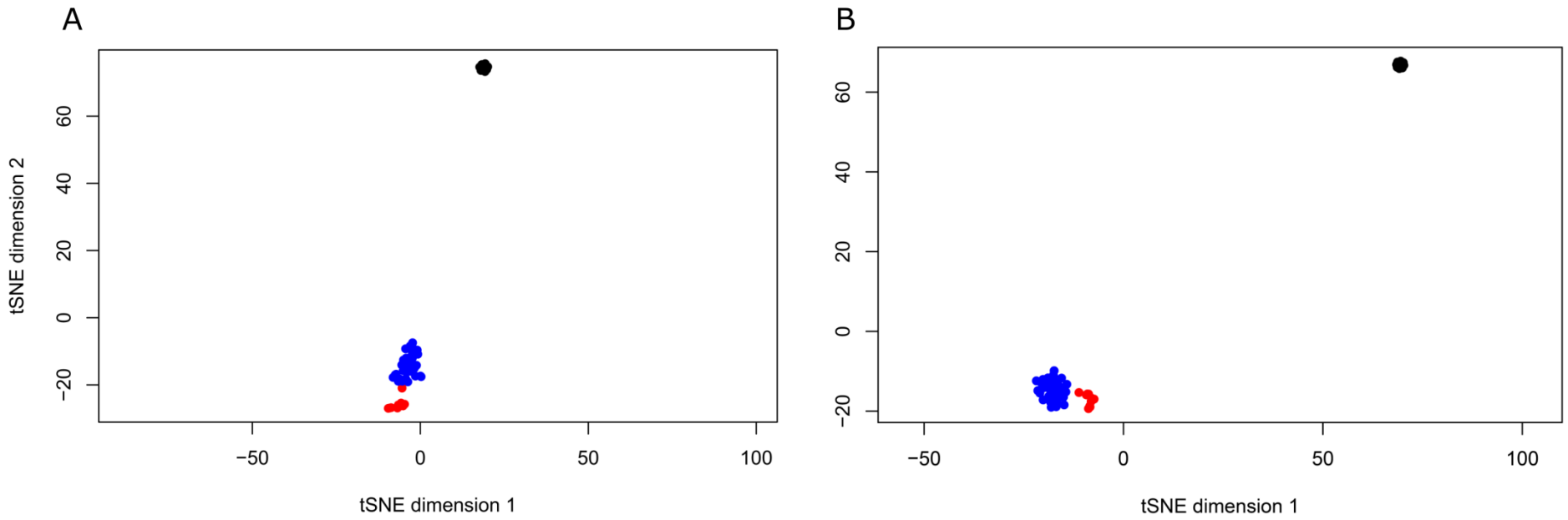


Figure 3.6 tSNE plots depicting CLR-transformed taxa data for phyla (A) and genera (B). Colours indicate sample type; blue = untreated site, red = slurry-impacted site, black = slurry). Perplexity = 10, iterations =1500.

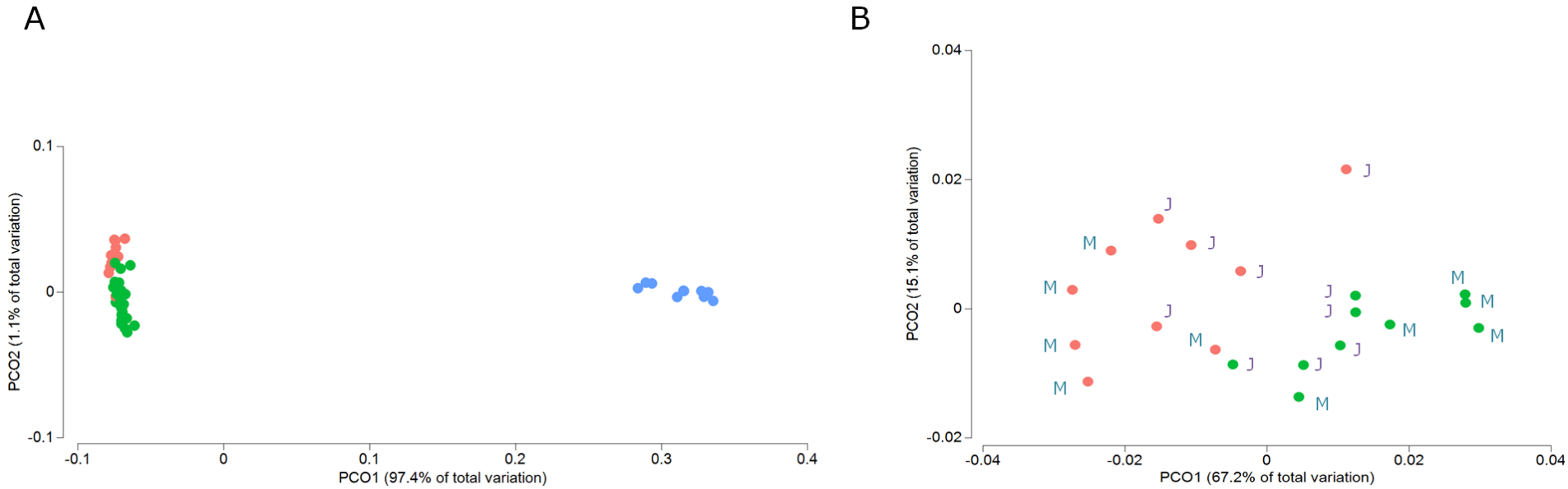


Fig 3.7 PCoA of weighted UniFrac distances across site (A), PCoA of weighted UniFrac distances for 2018 soil data (B). Phylogeny based on placement of 16S rRNA reference sequences. Where applicable, colours represent sample type; blue (slurry), green (slurry-treated soil) and orange (untreated soil). Where applicable, letters denote month of sample collection; J (January 2018) and M (May 2018).

3.3.2 Site Differential Abundance

After controlling for the effect of site on dispersion, Corncob reported 1904 genera (see Supplementary file 2) and 124 phyla were differentially abundant across sites (untreated site used as baseline). Differential taxa broadly echoed patterns inferred from the relative abundance of site-specific dominant taxa. For instance, relative to untreated soil, Actinobacteria and Gemmatimonadetes were elevated in the slurry-impacted site and significantly less abundant in slurry. Likewise, Firmicutes and Bacteroidetes were increased in slurry relative to soil.

However, Corncob additionally highlighted substantial site differences in many less abundant taxa. For example, *Candidatus* Eisenbacteria and most other differentially abundant phyla were more abundant in untreated soil relative to slurry-impacted soil (Figure 3.8). Conversely, only five other phyla were convincingly more abundant or differentially variable in treated soil relative to untreated soil. These include Tenericutes, Chlorobi, *Candidatus* Saccharibacteria, Chloroflexi and Bacteroidetes. Of these, only Bacteroidetes was among the dominant phyla. Predictably, slurry possessed many phyla which were highly abundant relative to soil (~75). In addition to the dominant phyla already described (Figure 3.3B), these include Fusobacteria, *Candidatus* Falkowbacteria and *Candidatus* Riflebacteria. Approximately 40 phyla are significantly less abundant in slurry compared to untreated soil, with the most prominent differences exhibited by *Candidatus* Eisenbacteria, Acidobacteria and *Candidatus* Rhokubacteria.

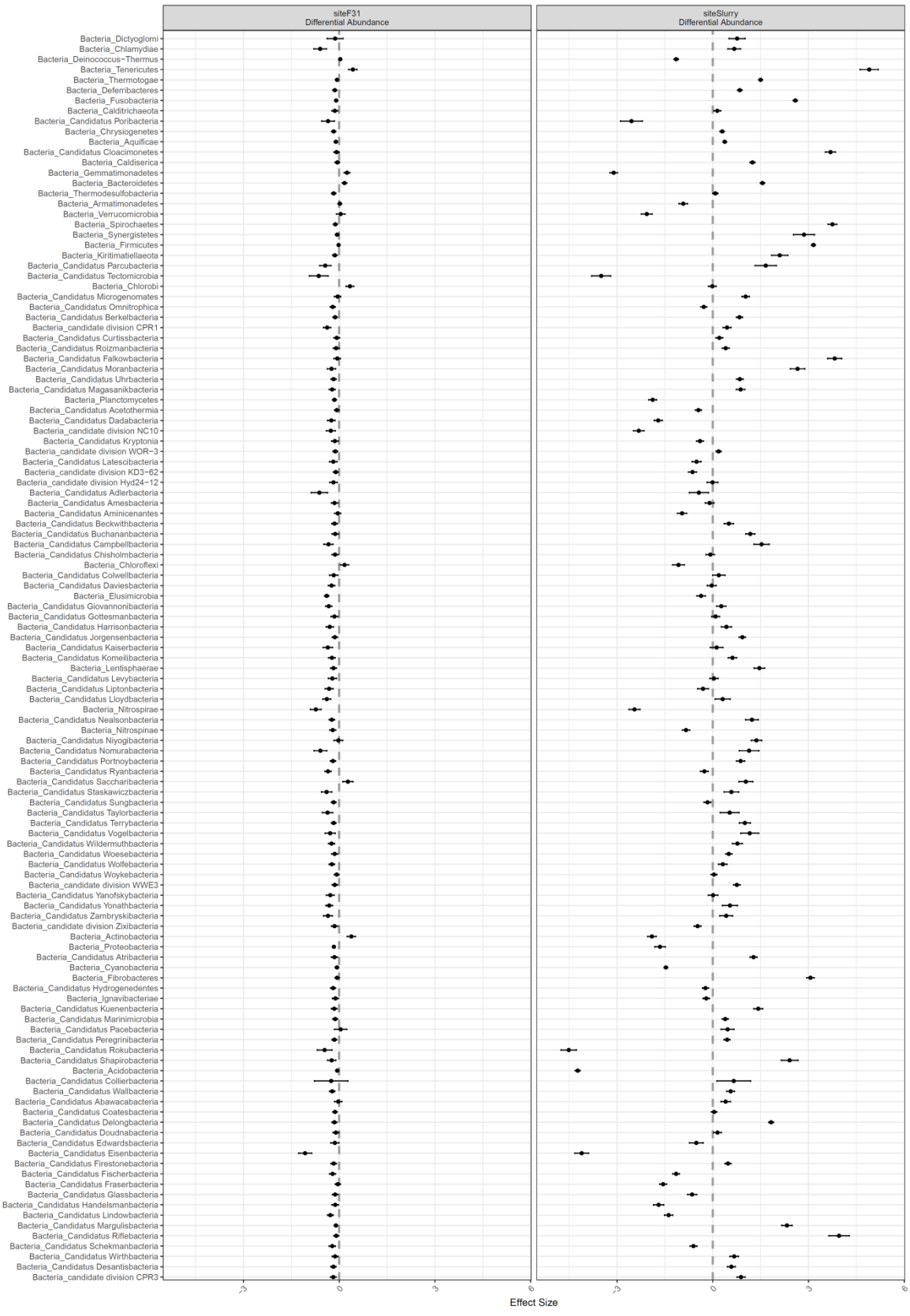


Figure 3.8 Corncob differential abundance plot comparing phyla relative abundance by sample type. Rows represent taxa. Untreated soil was used as the baseline for abundance and is represented by the central vertical dashed lines at 0. Taxa to the left of the dashed lines are less abundant relative to untreated soil, while those to the right are more abundant. Left panel compares slurry-impacted soil (F31) to untreated soil, while the right panel compares slurry to untreated soil. Bars denote 95% confidences.

Note: all taxa which were significantly different in at least one pair-wise comparison are shown, hence some confidence intervals may overlap zero where significance was not present for all comparisons.

As shown with phyla (Figure 3.8), differences between the top 10 genera in soils (Figure 3.3C) did not always signify the most differentially abundant taxa by field site. Although *Nitrospira* and *Rhodoplanes* were more abundant in untreated soil than in slurry-treated soil, *Simkania*, *Vermiphilus* and *Halobacteriovorax* were most differentially abundant; $p < 0.05$, effect size > 0.5 (note: taxa with extremely low and inconsistent read counts were excluded due to model instability). Data are not shown for logistical reasons related to size of Figure. Similarly, the relative abundance of *Nocardioides* and *Gemmatimonas* in treated soil samples compared to untreated soil was surpassed by *Terrabacter*, *Monashia* and *Intrasporangium* among others ($p < 0.05$, effect size > 1). Interestingly, the three genera associated with slurry-treated soil all belong the family Intrasporangiaceae, in turn, a member of the Actinobacteria.

The most differentially abundant genera positively associated with slurry included *Sphaerochaeta* followed by *Acholeplasma*, *Sarcina*, *Fermentimonas* and *Sedimentibacter* ($p < 0.05$, effect size >5). Unlike within the soil, dominant genera were among the most differentially abundant. However, despite the much greater prevalence of *Clostridium* in slurry compared to soil, modelling indicated less abundant genera displayed a larger difference in population size. Genera underrepresented in slurry when contrasted with soil included *Rhodoplanes*, *Pseudolbrys*, *Pseudorhodoplanes* and *Candidatus Solibacter* among others ($p < 0.05$, effect size >5). The most negatively differentially abundant genera share the order Rhizobiales (*Rhodoplanes, Pseudolbrys* and *Pseudorhodoplanes*).

Lastly, effect sizes associated with differential abundance were generally much larger when comparing slurry to soil than when contrasting soil samples. A similar pattern was observed in resistome data.

3.3.3 Changes in Taxon Abundance Following Slurry Application

Using pre-application samples as a baseline (5 days prior to amendment), shifts in the abundance of phyla and genera were observed in soil samples collected immediately after the first slurry application of 2017 (<24 hours post-application). In some instances, the immediate effects of slurry application diminished over the course of 56 days, while other taxa appeared to display a delayed response. In total, Corncob predicted 10/134 differentially abundant phyla (Figure 3.9) and 631/2232 genera (Supplementary file 3). Owing to the large number of differentially abundant taxa highlighted by Corncob, only those which exhibited the largest and most consistent effect sizes are discussed in detail.

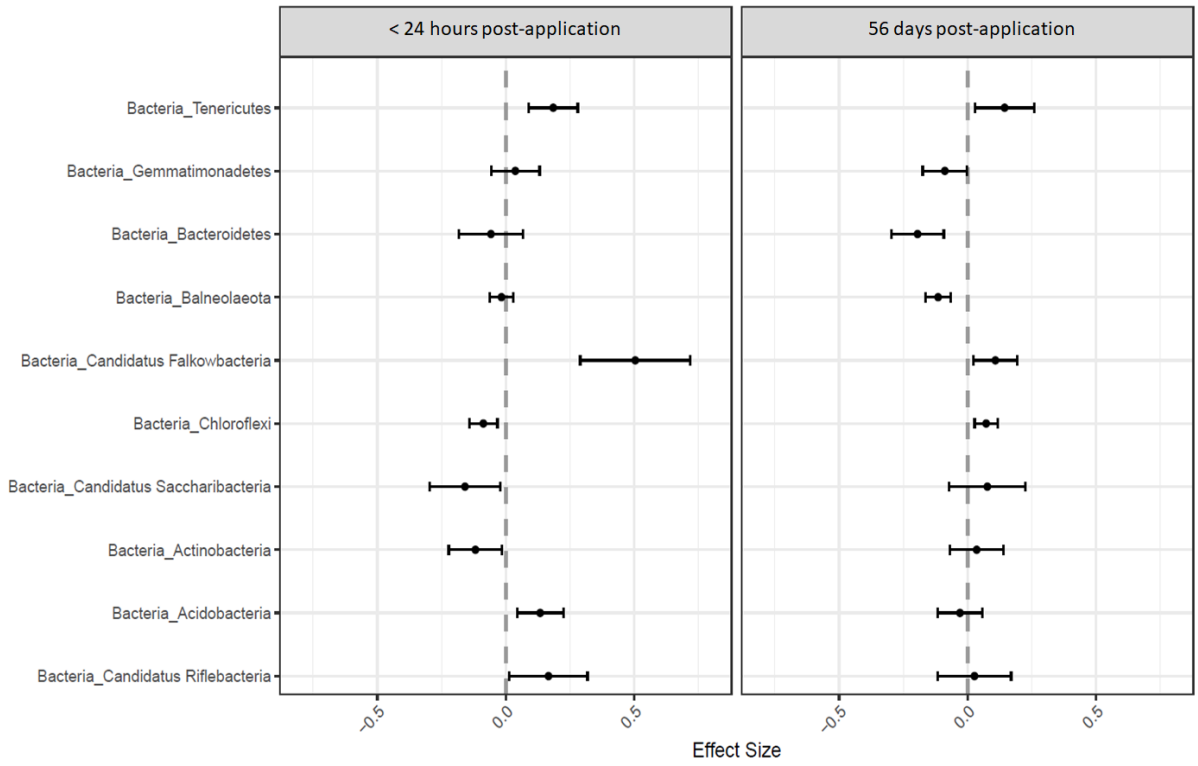


Figure 3.9 Differentially abundant phyla according to Corncob <24 hours after slurry application (left panel) and 56 days following slurry application (right panel) in May 2017. Here the base-line for effect size is represented by soil sample data obtained from the slurry-impacted site five days before the first application of slurry in 2017 (i.e., points left of the dashed lines are less abundant than baseline data, while those to the right are more abundant). Bars indicate 95% confidences. Note: all taxa which were significantly different in at least one pair-wise comparison are shown, hence some confidence intervals may overlap zero where significance was not present for all comparisons.

Candidatus Falkowbacteria was the most strongly enriched phylum both on the day of slurry application and 56 days later ($p = 0.0017$ and $p = 0.037$, respectively). However, this increase in *Candidatus* Falkowbacteria declined over 56 days. A similar, albeit less pronounced increase was observed in Tenericutes immediately after slurry application, with enrichment still evident 56 days later ($p = 0.0053$ and $p = 0.040$, respectively). The candidate phyla *Candidatus* Falkowbacteria was first proposed by Brown et al. (2015) following extensive phylogenetic analyses of metagenome assembled genomes which were shown to share unique features indicative of symbionts (these genomes collectively form the putative Candidate Phyla Radiation). Further studies have since identified *Candidatus* Falkowbacteria genomes in aquifers, suggesting its association with a specific ecological niche (Anantharaman et al., 2018). A number of phyla which were unaffected immediately after slurry application were shown to have declined in abundance 56 days later. These included Bacteroidetes ($p = 0.0059$) and Balneolaeota ($p = 0.0019$). In contrast, Chloroflexi appeared to be negatively impacted by slurry application initially ($p = 0.014$) and then recovered, exceeding the baseline abundance 56 days after fertilisation ($p = 0.014$).

Analysis indicated that a distinct assemblage of genera was increased within 24 hours of slurry application. Most prominently, these included *Fermentimonas*, *Petrimonas*, *Sedimentibacteria*, *Proteiniphilum* and *Sphaerochaeta* (listed in decreasing order of effect size). After 56 days a subset of these, namely *Fermentimonas*, *Petrimonas* and *Proteiniphilum*, were still significantly enriched relative to soil sampled before the first slurry application of the same year. Nonetheless, the abundance of these genera clearly declined over 56 days. Corncob modelling of these five genera throughout the period of study are shown in Figure 3.10. Other genera such as *Pseudolabrys* and *Bradyrhizobium* were also comparably increased 56 days after slurry application, despite a lack of enrichment evident on the day of slurry application.

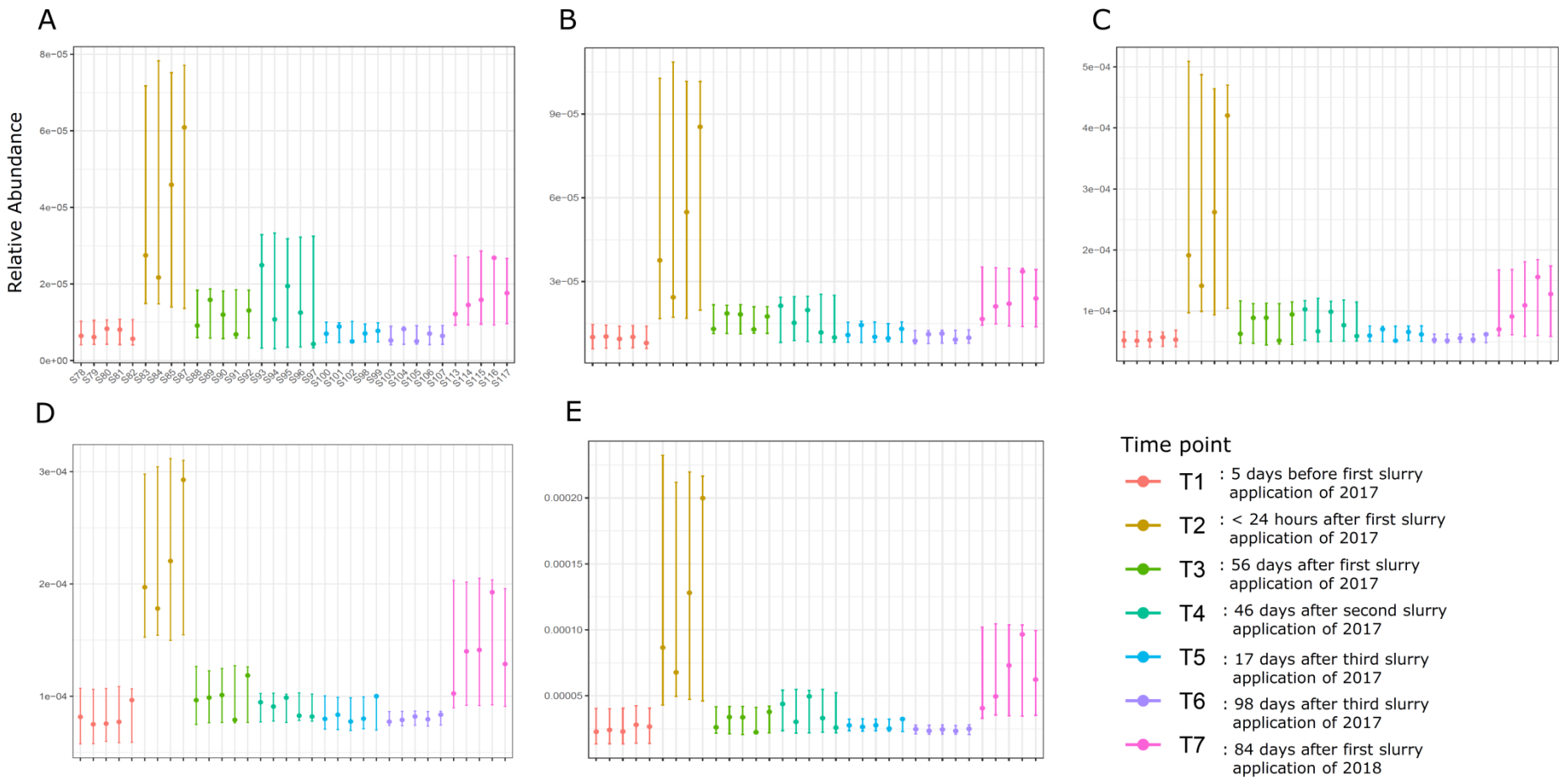


Figure 3.10 Corncob model output plots for slurry biomarker genera which increased following slurry application (slurry-impacted field site only). *Fermentimonas* sp. (A), *Petrimonas* sp. (B), *Proteiniphilum* sp. (C), *Sphaerochaeta* sp. (D) and *Sedimentibacter* sp. (E). Samples are ordered chronologically on x-axis from T1 to T7. Estimated relative abundance as modelled by Corncob displayed on y-axis. Colours denote sampling time-points (refer to legend for further details about specific time-points).

3.3.4 Seasonal Differences in Bacterial Communities by Site

Focussing on broad-scale shifts in bacterial communities according to season (spring/summer compared to autumn/winter), Corncob identified 65, 10 and 8 differentially abundant phyla in slurry, untreated soil and slurry-impacted soil, respectively (Figure 3.11). Interestingly, although the untreated site and slurry-impacted site showed limited dynamism at the phylum level, according to Corncob; 670 genera were differentially abundant by season in slurry-impacted soil compared to only 57 genera in untreated soil. This may parallel patterns described in Chapter 2 regarding seasonal changes in ARGs. Furthermore, it is apparent that microbial communities in slurry (most likely due to the limited protection from environmental conditions afforded by the open-topped tank) experience numerous phylum-level disturbances, in contrast to soil. Genus-level seasonality data are not discussed further, however Corncob plots are available (Supplementary file 4, 5 and 6). These results are included primarily to demonstrate the presence of seasonal shifts, and as the primary focus is on the dissemination of slurry-borne bacteria, such shifts will not be discussed in greater detail.

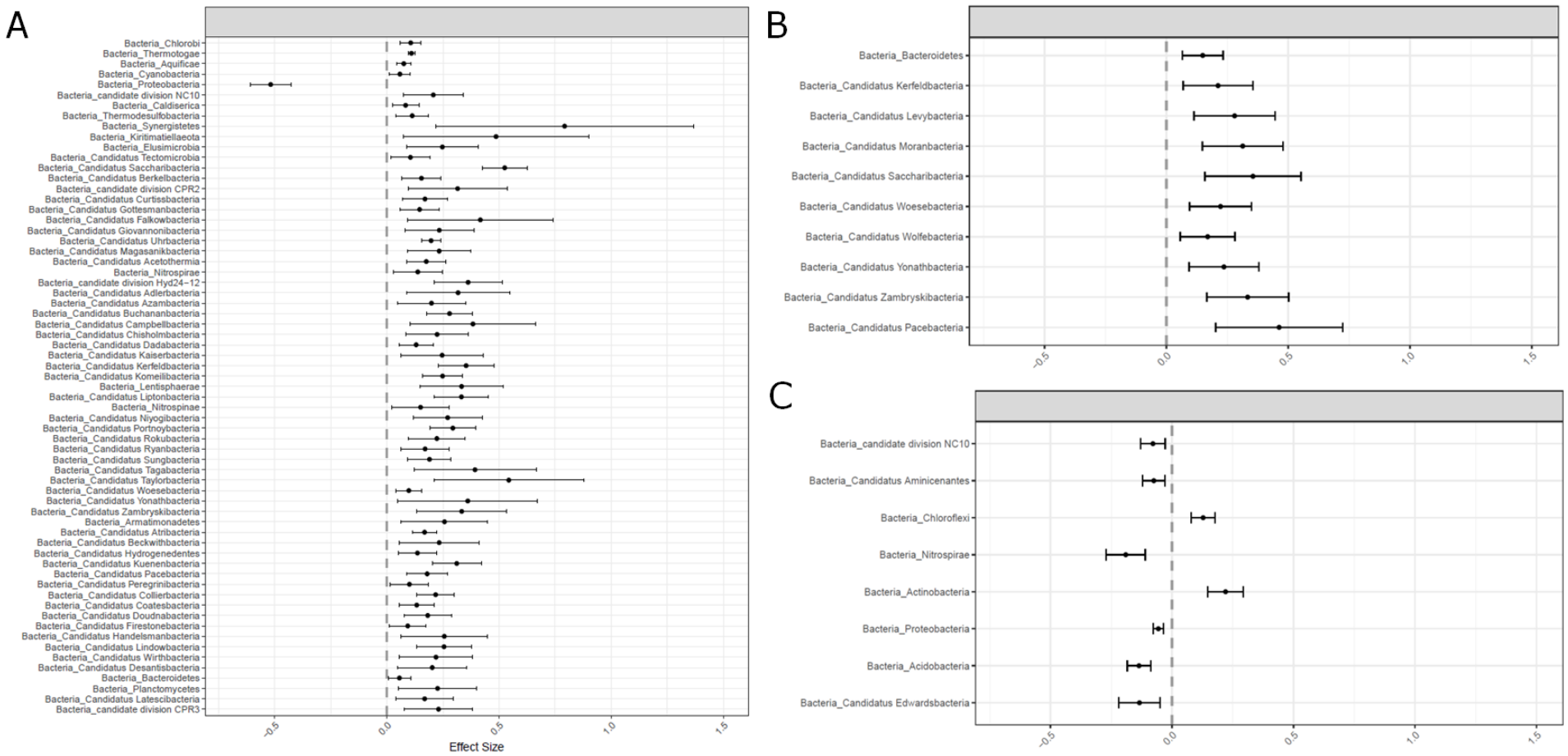


Figure 3.11 Differentially abundant phyla based on season. Slurry (A), untreated site soil (B) and slurry-treated site soil (C). Here the baseline for effect size is represented by spring/summer (i.e., points left of the dashed lines are less abundant in spring/summer, while those to the right are more abundant). Bars indicate 95% confidences.

3.3.5 Physiochemical Correlations With Taxa

Initial exploration of soil physiochemical properties suggested the presence of trends between the two sites (Figure 3.12). While the site with no history of slurry addition was associated of extractable caesium, cadmium and zinc; the treated site had a higher pH in combination with extractable calcium, magnesium and sodium. TOC, TC and barium were also distinguishing properties; however, Corncob analyses showed that pH explained the greatest proportion of variation in taxa between sites (Figure 3.13). A significant difference in pH between the two sites was further confirmed with a Welch two sample t-test ($p = 0.0053$). Lastly, correlations between pH and phyla were analysed using Spearman's Rank testing (see Figure 3.14 for selected plots). See Appendix 1 for a summary table of physiochemical data.

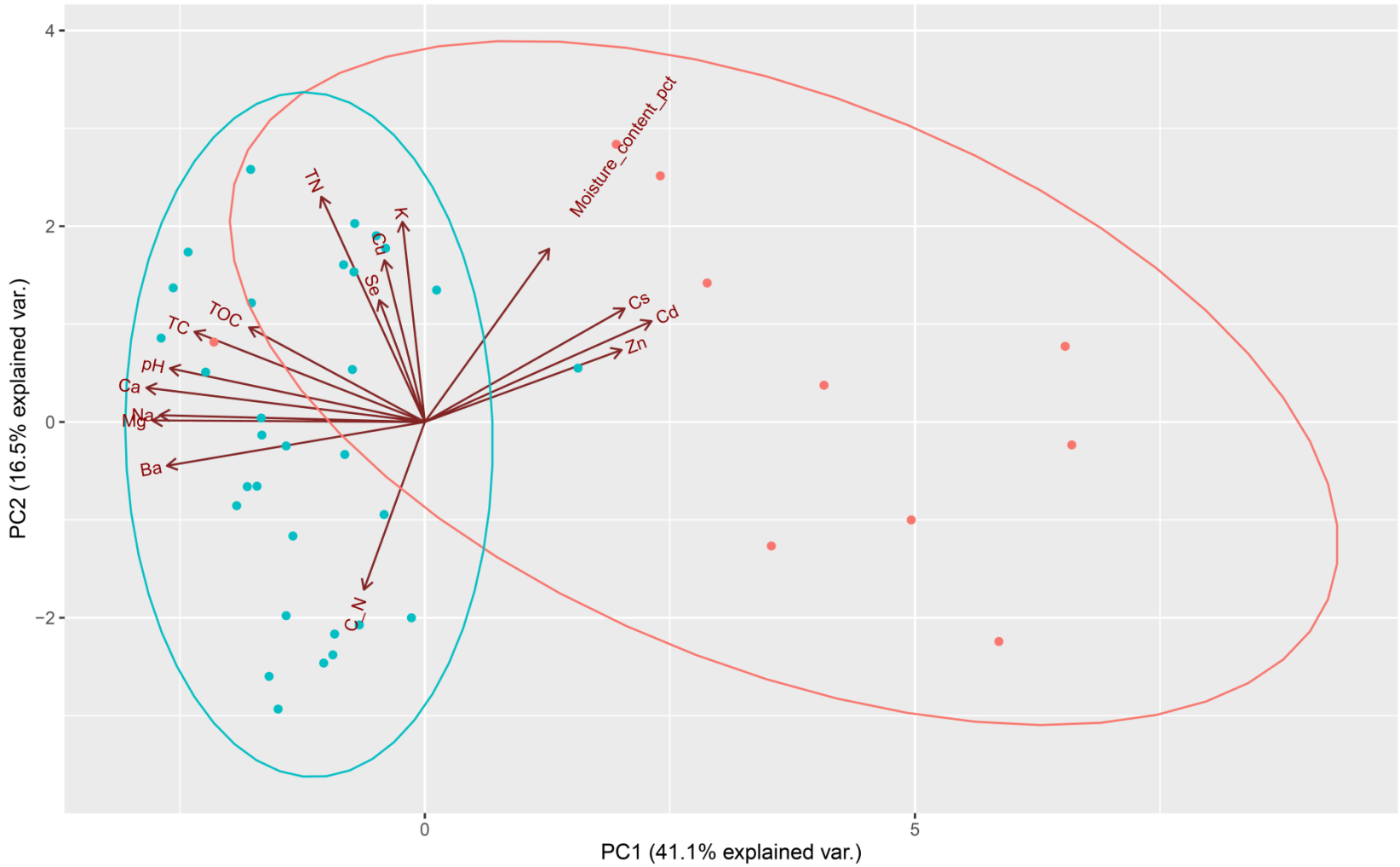


Figure 3.12 Log-normalised PCA of selected soil physiochemical properties by site (all units mg kg^{-1} except moisture which is given as percent). C_N = C/N ratio. Data are scaled and centred. Ellipses represent 0.9 probability. Arrows denote direction of greatest change, while length indicates strength of change. Colour corresponds to site; slurry-impacted soil (blue) and untreated site soil (orange).

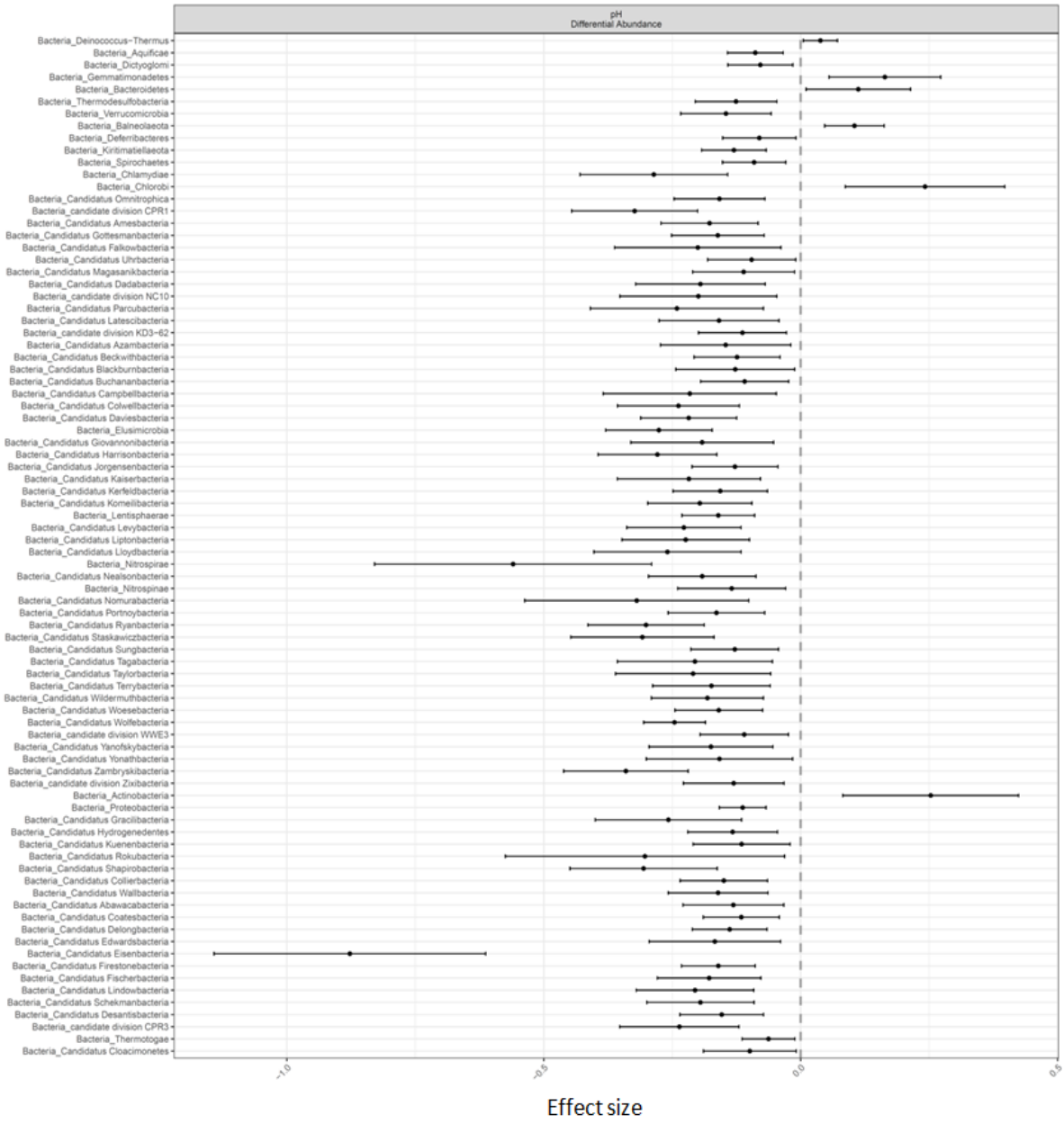


Figure 3.13 Corncob analysis output plot showing differentially abundant bacterial phyla with pH as a covariate. Only soil data are displayed. Rows represent taxa. Points left of the dotted line are associated with reducing pH, while those right of the dashed line are associated with increasing pH. Bars denote 95% confidences.

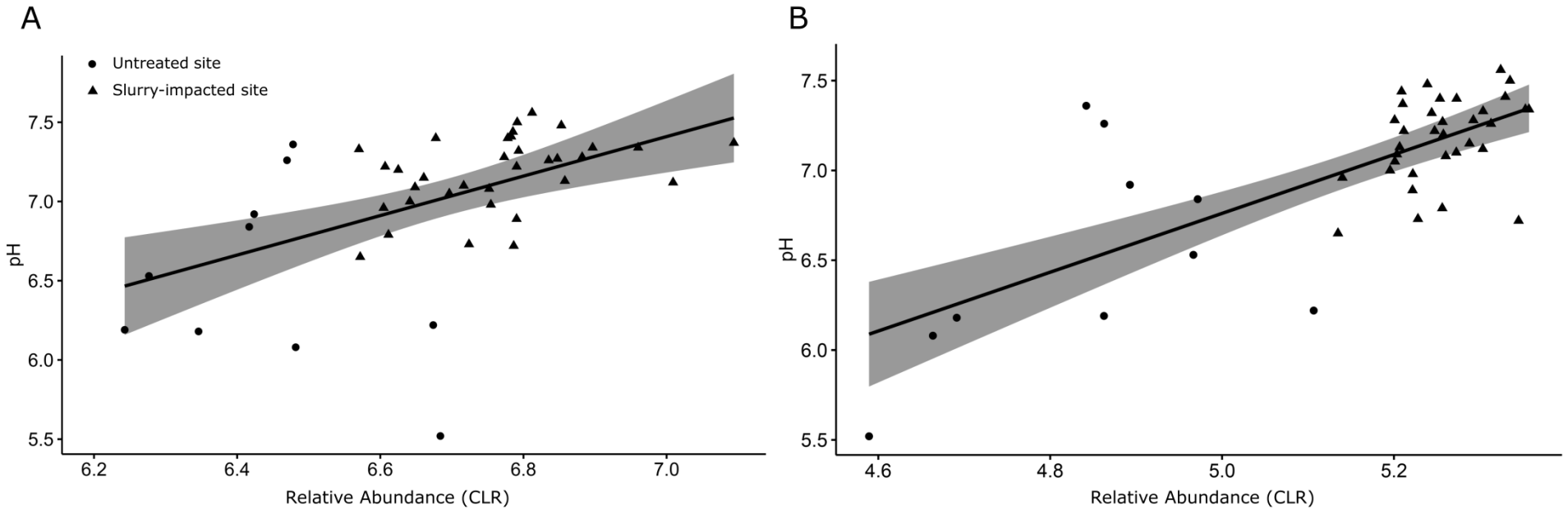


Figure 3.14 Correlations between Bacteroidetes (R^2 = 0.31, R = 0.55, p <0.01) (A), Gemmatimonadetes (R^2 = 0.32, R = 0.57, p <0.01) (B) and pH. CLR-transformed count data for both sites are displayed; untreated field site (circles), slurry-impacted field site (triangle). Bands denote 95% confidence intervals.

3.3.6 Bacterial Richness and Diversity Estimations by Site

Estimates of bacterial species richness calculated by iNEXT indicated differences between slurry and soil samples in line with raw richness estimates. On average, slurry-impacted soil and untreated soil samples had greater species richness compared to slurry, while no significant difference was observed between soil sites (see Table 3.1). Both estimated alpha-diversity measures also suggest that soils were more diverse than slurry, while no significant difference was detected between soil sites. Soil is known to be a highly structurally and biologically diverse environment with many unique low-abundance microbial taxa, whereas the slurry tank represents a comparatively hostile environment with likely fewer ecological niches to occupy. It is therefore unsurprising that soil is more diverse than slurry, even when focussing on the diversity of dominant taxa (as with q2). Beta-diversity estimates with weighted UniFrac distances further demonstrate contrasts between soil and slurry, , showing that while soil sites were indeed distinct from each other in terms of community composition ($p < 0.001$, mean distance < 0.04), slurry was considerably more divergent ($p < 0.001$, mean distance from either soil site > 0.3).

Table 3.1 Pair-wise comparisons of species richness and alpha diversity measure estimates according to iNEXT. Beta-diversity estimates based on weighted UniFrac distance .

Diversity measure	Pairwise comparison (Site)	p- value
iNEXT Richness (q0)	Slurry-impacted / Untreated	0.93
	Slurry / Untreated	< 0.001 **
	Slurry-impacted / Slurry	< 0.001 **
iNEXT Shannon (q1)	Slurry-impacted / Untreated	0.84
	Slurry / Untreated	< 0.001 **
	Slurry-impacted / Slurry	< 0.001 **
iNEXT Inverse Simpson (q2)	Slurry-impacted / Untreated	0.95
	Slurry / Untreated	< 0.001 **
	Slurry-impacted / Slurry	< 0.001 **
Weighted UniFrac Distance	Slurry-impacted / Untreated	< 0.001 **
	Slurry / Untreated	< 0.001 **
	Slurry-impacted / Slurry	< 0.001 **

3.3.7 Network Analysis

The structure, connectivity and potential hubs of bacterial networks differed by site, despite the absence of site-discriminant phyla (Figure 3.15). Connectivity corresponds to the number of correlations between nodes. Connectivity was greatest in slurry, with most nodes (phyla) connected to between three to four other nodes. Meanwhile slurry-impacted soil and untreated soil nodes were typically connected to between two to three and one to two nodes respectively. The slurry-impacted soil network possessed the node with the most positive connections; *Candidatus* Aminicenantes (n = 8). *Candidatus* Aminicenantes therefore represents a hub node in slurry-impacted soil; defined as a phylum connected (correlated) with many more phyla than average within a network. It is difficult to infer the true significance of these hub nodes without further experimental validation of the putative relationships. However, hub node phyla provide potential leads for future investigations aiming to identify ecologically important taxa in specific environments. There were many potential hub nodes in slurry, although none had more than five positive edges. In untreated soil, Ignavibacteriae was the most connected node (n = 6). While all networks contain between 73 and 80 negative edges, slurry had the most positive edges (n = 141), closely followed by slurry-impacted soil (n = 130) and untreated soil (n = 93). Core phyla ubiquitous in both field sites shared few connections, although Proteobacteria and Bacteroidetes were associated in both field sites. Interestingly, Bacteroidetes and Balneolaestona were associated in both slurry and slurry-impacted soil networks, but not in untreated soil. Additional cross-over could be found in the shared association between Kiritimatiellaeota and Lentisphaerae.

Lastly, given the difference in sampling effort between field sites caution should be exercised when comparing these networks directly and where possible, inferences should be supported with other analyses which do not rely solely on correlation.

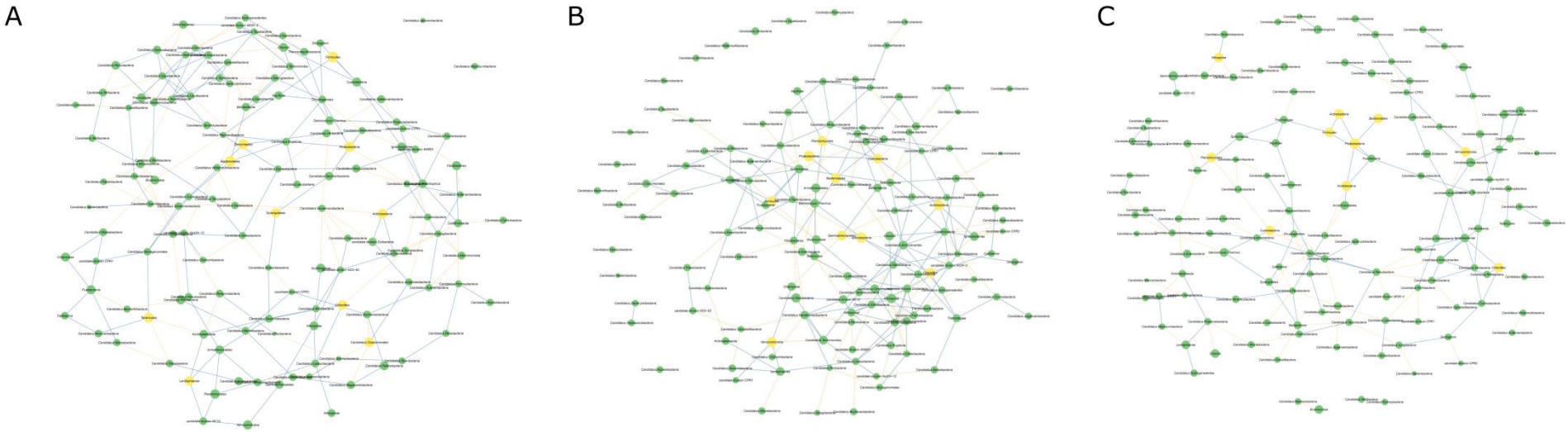


Figure 3.15 SpiecEasi networks depicting slurry (A), slurry-impacted site soil (B), and untreated site soil (C) phyla. Yellow nodes indicate top 10 most abundant phyla, all others are coloured green. Node size is proportional to relative abundance. Blue lines represent positive correlations. Orange-dashed lines indicate negative correlations. Correlations fulfil fdr-adjusted $p<0.05$.

The 'global' genus-level network constructed using CoNet suggests robust clustering at the phylum-level (Figure 3.16). In addition, the two major clusters appear to broadly separate phyla according to those dominant in specific sample types. For example, the majority of Actinobacteria and Proteobacteria (soil-dominant), are clustered separately from Firmicutes (slurry-dominant). While the former may not be surprising, it is interesting that Bacteroidetes is split into two clear clusters which appear to associate with either the soil-dominant cluster or the slurry dominant-cluster. This may suggest a functional division within this phylum in relation to the two environments sampled. Furthermore, it is interesting that alpha-Proteobacteria cluster strongly, while the beta and gamma-Proteobacteria are intermingled; perhaps alluding to interactions between beta and gamma- Proteobacteria which are not shared with alpha-Proteobacteria. Candidates for bacterial transfer from slurry to soil were isolated from the global network while preserving nearest neighbour connections. The resulting 'module' shows some of these bacteria are highly correlated with other members of this group (Figure 3.17).

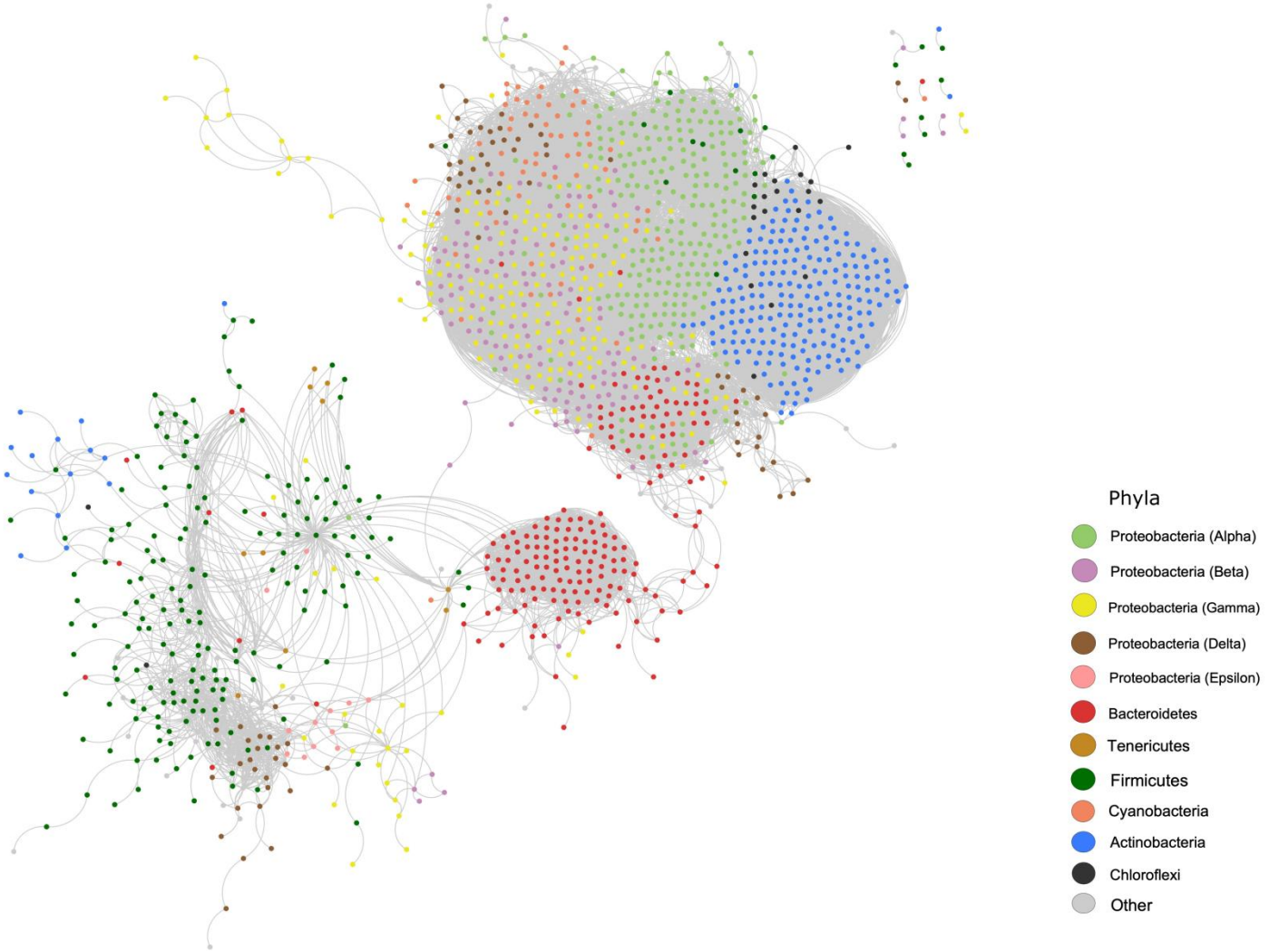


Figure 3.16 Genus-level network analysis of all samples (slurry, slurry-impacted soil and soil from the untreated site) produced with CoNet and visualised in Cytoscape. Node size is fixed. Connections (grey lines) indicate correlations; $p < 0.05$, $R > 0.8$. Colours in the key refer to the corresponding phyla of each genus.

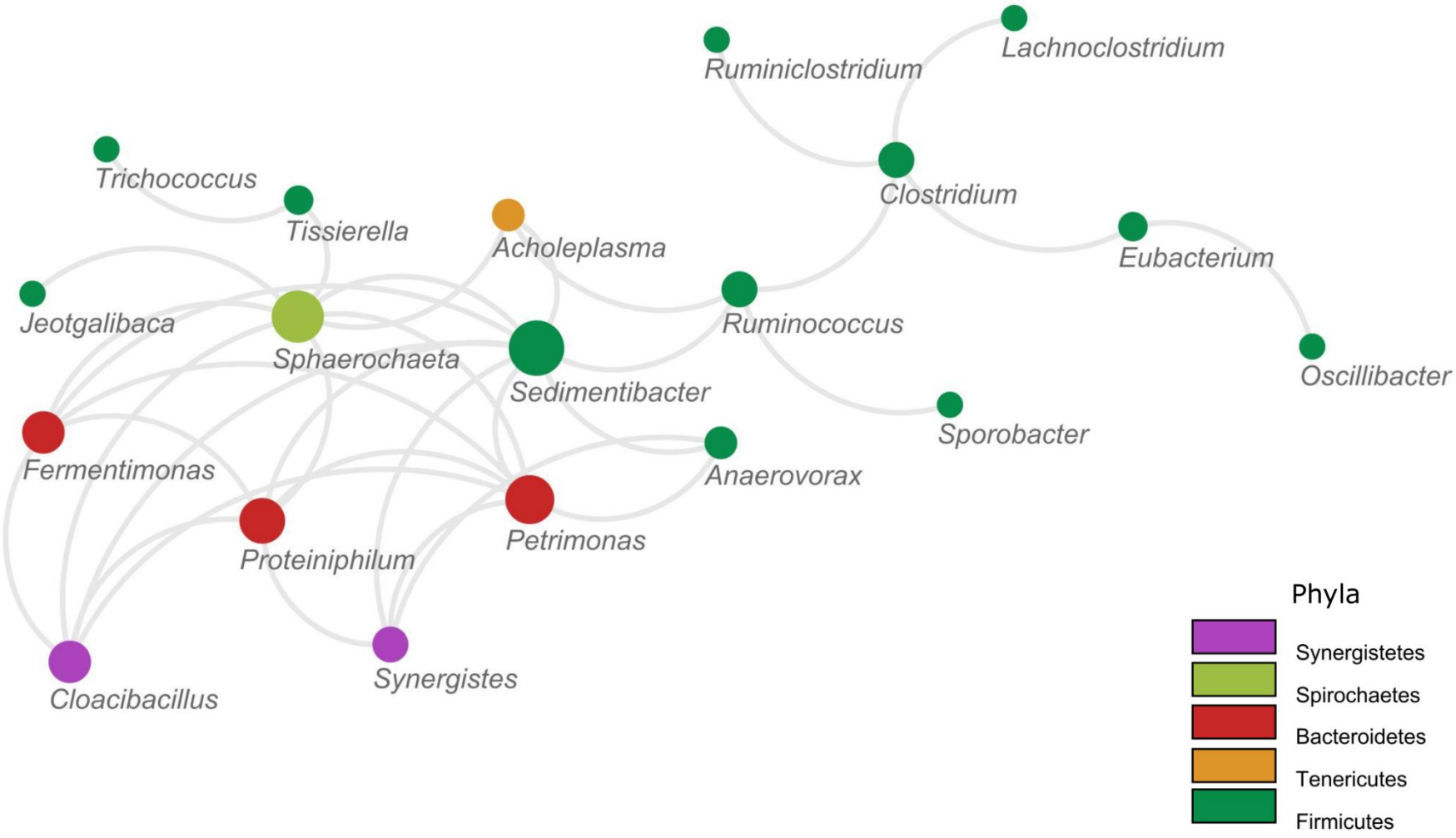


Figure 3.17 Putative slurry-borne bacterial biomarkers based on network analysis of all samples (see Figure 3.16). Node size is proportional to the number of connections. Connections indicate correlations $p < 0.05$, $R > 0.8$. Colours in the key denote the phylum of each genus.

3.4 Discussion

3.4.1 Bacterial Composition by Site

3.4.1.1 Slurry

The taxonomic composition of soil and slurry metagenomes characterised in this study broadly reflects previous research investigating these environments, irrespective of whether 16S rRNA amplicon or whole genome shotgun techniques have been implemented. For example, numerous studies determined that Firmicutes and Bacteroidetes are among the most abundant phyla in both solid cattle manure (Shanks et al., 2011, Sun et al., 2015) and liquid slurry (Li et al., 2014b, Habtewold et al., 2018). As shown in the present study, Firmicutes is typically preponderant over Bacteroidetes in cattle waste, however the reverse has also been reported (Pandey et al., 2018). Studies have also identified similar phyla (e.g. Firmicutes) are prevalent in swine faecal material, indicating these phyla inhabit core functional niches common to the gastro-intestinal tract of mammals (Sun et al., 2015, Rieke et al., 2018, Wolters et al., 2018, Lim et al., 2020). The effective use of taxa belonging to these two phyla as indicators of mammalian faecal contamination further illustrates the consistency of this association (Unno et al., 2010, Fisher et al., 2015). On the other hand, Reese and Dunn (2018) demonstrated that contrasting gut physiologies are associated with differing levels of microbial species diversity (ruminant vs. monogastric, respectively) and so are likely to contain distinct species despite sharing similar characteristics at the phylum level. It is also important to remember that slurry samples represent the herd as a population, rather than a profile of the gut contents of discrete individuals.

Nonetheless, many slurry-associated genera have also been reported in other studies evaluating swine and/or cattle faecal matter. These genera include but are not limited to *Clostridium* (Pandey et al., 2018, Wolters et al., 2018), *Bacteroides* (Fisher et al., 2015, Pandey et al., 2018), *Acholeplasma* (Wolters et al., 2016, Pandey et al., 2018), *Prevotella* (Fisher et al., 2015, Wolters et al., 2018) and *Alistipes* (Fisher et al., 2015,

Pandey et al., 2018). On the other hand, *Sphaerochaeta*, which was the most abundant genus in slurry according to the present study, was only identified among prevalent taxa in liquid cattle manure and not in solid manure according to previous research by Pandey et al. (2018). *Sphearochaeta* is unusually chimeric; representative genomes have been shown to share more unique genes with Clostridia than members of their parent phyla (Caro-Quintero et al., 2012). It has been proposed these shared genes relate to a succession of horizontal gene transfer events. It may therefore be reasonable to find *Sphaerochaeta* in locations with an abundance of Clostridia. Since the slurry tank essentially undergoes slow anaerobic digestion, the putative fermentative lifestyle of *Sphaerochaeta* inferred by Caro-Quintero et al. (2012) would be well suited to this environment. Finally, Wolters et al. (2018) found the phylum Spirochaetes to which *Sphaerochaeta* belongs, was significantly more abundant in swine waste matter compared to soil.

Overall, while a comparable range of genera were frequently recovered from faecal environments, the proportional relationship among them varies between studies. This is unsurprising given that Pandey et al. (2018) demonstrated the practice of separating the liquid and solid fraction of cattle waste could result in significantly different microbial communities at lower taxonomic ranks despite their shared influent of origin. Indeed, the microbial assemblage of animal waste has been shown to vary according to many factors including species, animal age, health status, storage method, time and ambient environmental conditions (Furet et al., 2009, Li et al., 2014b, Habtewold et al., 2018) Interestingly, while Wolters et al. (2016) found the most abundant genera in raw swine slurry differed from those in its anaerobic digestate; the dominant phyla (Firmicutes and Bacteroidetes) remained unchanged. In a later study Wolters et al. (2018) also reported differences at the phylum level, with Tenericutes exhibiting increased abundance in anaerobic digestate relative to raw slurry. Nonetheless, all faecal-derived samples clustered together in distance analyses, and were clearly discrete from soil samples at phylum and genus level (Wolters et al., 2018). In summary, the slurry-associated taxa

identified in the present study are consistent with literature assessing samples of faecal origin, although site-specific variables and species are influential factors at lower taxonomic ranks.

It is important to establish the merits of considering mammalian livestock waste in a broad sense since there are a limited number of high-throughput microbiome surveys which specifically address cattle manure application to soil and even fewer consider liquid slurry. Consequently, the discussion will draw on studies which evaluate the effects of solid manure as well as swine manure on soils. A number of differentially abundant taxa in the present work have not been identified as such in previous studies. Given that these are often putative candidate taxa (e.g. *Candidatus* Falkowbacteria) and are yet to be formally accepted (Brown et al., 2015, Anantharaman et al., 2018), the absence of these taxa from other studies may be attributed either to their deliberate exclusion or to the rapid rate of reference database revision and growth.

3.4.1.2 Soil

Regardless of management history, soils possessed a dramatically different bacterial community structure and diversity compared to slurry (Figures 3.4 and 3.6). Proteobacteria and Actinobacteria were dominant phyla across both field sampling sites and the prominence of these phyla in agricultural soil is well established (Spain et al., 2009, Aislabie et al., 2013, Sun et al., 2015, Liu et al., 2017, Zhang et al., 2017c, Zhang et al., 2018b, Wang et al., 2019). Furthermore, the proportions of Proteobacterial classes were similar in both soils, with Alpha and Beta-protobacteria forming the principal members. Interestingly, Spain et al. (2009) reviewed several publications concerning soils from a range of climates across the American continent and found that while Proteobacterial classes were variable, Alpha and Beta-proteobacteria were generally more abundant than Gamma-Proteobacteria. In contrast, slurry Proteobacteria chiefly belonged to Gamma-proteobacteria and to a lesser degree Epsilon-

proteobacteria; the latter of which are comparably scarce in soil (Spain et al., 2009, Wolters et al., 2018). Gamma-proteobacteria include the Enterobacteriaceae family such as *E. coli* (Köhler and Dobrindt, 2011) and *Enterobacter cloacae* (Keller et al., 1998) which are gut commensals as well as opportunistic pathogens; their abundance in the slurry tank is therefore expected to be greater than in soils. Likewise, Epsilon-proteobacteria are known to encompass metabolically diverse non-pathogenic bacteria (Nakagawa and Takaki, 2009), in addition to opportunistic enteric pathogens including *Campylobacter* sp. (Inglis et al., 2010), *Helicobacter* sp. (Fujimura et al., 2002) and *Arcobacter* sp. (Giacometti et al., 2015), all of which have been documented in raw dairy products or bovine manure. Meanwhile, phyla such as Acidobacteria and Bacteroidetes are also known to comprise a significant proportion of core soil bacterial communities (Aislabie et al., 2013, Zhang et al., 2018b, Podmirseg et al., 2019).

The ecological function of many bacteria associated with major soil phyla is poorly understood since a large proportion are known through culture-independent analyses alone (Janssen, 2006, Spain et al., 2009). Nonetheless, the categorisation of bacterial phyla into ecological groups based on observed 'life strategies' has been proposed on several occasions (reviewed by Ho et al., 2017), and putatively linked to fundamental soil processes such as carbon storage (Trivedi et al., 2013). One such method includes the division of phyla into copiotrophic (typified by consumption of labile carbon and rapid growth rates) and oligotrophic (characterised by the utilisation of recalcitrant carbon and slow growth) groups (Fierer et al., 2007, Trivedi et al., 2013). However, in meta-analyses, Ho et al. (2017) underscored that such broad categorisations can be misleading or inconsistent and suggest greater taxonomic resolution (i.e. to family and genus level) may be necessary to accurately discern community responses to environmental change.

In the current study, two principal groups of differentially abundant taxa emerged in soil under different management practice; those that were differentially abundant by site and those which were temporally variable within sites. While the former are likely to

represent the cumulative impact of contrasting management, the latter are taken to reflect the short-term influence of episodic slurry applications in amended soil. Although seasonal effects were detected in both soils and slurry, the data have sufficient resolution to demonstrate the proposed impact of slurry application is not merely an artefact of unrelated seasonal factors. This is most clearly evidenced by the marked increase in the relative abundance of slurry biomarker genera following slurry application in May 2017 (Figure 3.10).

3.4.2 Bacterial Differential Abundance

Firstly, phylum level differences between slurry-impacted and untreated soil will be discussed. The apparent enrichment of Bacteroidetes in slurry-impacted soil relative to untreated soil is in agreement with several publications assessing the impact of animal manure application on soil bacterial communities. These included experiments examining both short-term and long-term exposure effects of swine or cattle manure (Chaudhry et al., 2012, Wang et al., 2016b, Zhang et al., 2017c, Rieke et al., 2018, Xiong et al., 2018). Based on responses to carbon mineralisation rates, Fierer et al. (2007) concluded Bacteroidetes exhibit copiotrophic (fast growing) rather than oligotrophic (slow growing) life strategies. It may therefore be expected that the addition of cattle manure would be accompanied by an increase in the abundance of these bacteria; as shown by Wang et al. (2016b).

Considering the abundance of Bacteroidetes in cattle slurry, it could be possible their prevalence in amended soil reflects the direct addition and survival of these bacteria in soil. However, while Bacteroidetes were significantly more abundant in the slurry-impacted site compared to the untreated site overall (Figure 3.8), their abundance in the treated site was negatively affected by the first slurry application of 2017 for at least 56 days (Figure 3.9).

Surprisingly, other phyla which were ostensibly enriched in treated soil such as Actinobacter and Gemmatimonadetes also declined following slurry application, although the Actinobacterial population recovered within 56 days. Nonetheless, Gemmatomonadetes have been associated with long-term exposure to cattle manure in previous publications (Chaudhry et al., 2012, Zhang et al., 2018b), while increases in Actinobacteria have been shown to occur in soil after organic and inorganic fertilisation (Chen et al., 2016a, Dai et al., 2018, Rieke et al., 2018). On the other hand, others have reported cattle and pig manure has a negative effect on Actinobacteria (Liu et al., 2017, Xiong et al., 2018, Zhang et al., 2018b, Chen et al., 2019b). In contrast, Tenericutes was clearly enriched in impacted soil following slurry application (Figure 3.9) and also contributed a greater proportion of reads in impacted soil samples compared to untreated site soils overall (Figure 3.8).

One explanation for the immediate negative impact on certain phyla may be that introduced bacteria temporarily suppress indigenous populations through increased competition. The fact that Chen et al. (2017) found members of Bacteroidetes were enriched in soils amended with irradiated pig manure more so than raw manure potentially supports this concept. Furthermore, contaminants such as antibiotics (Hammesfahr et al., 2011, Cleary et al., 2016) and high concentrations of ammonia in cattle slurry may cause transient toxicity (Unc and Goss, 2004). Heavy metal-contaminants in cattle slurry (Nicholson et al., 2003, Franco-Uría et al., 2009) are also likely to disrupt soil bacterial communities following fertilisation (Giller et al., 2009, Xu et al., 2019). Intriguingly, no significant difference in available TOC was observed between the two sites which indicates other edaphic factors are the primary driver(s) behind differences in bacterial community structure at these locations.

Accordingly, a series of publications have reported that soil pH is a critical determinant of bacterial community structure (Lauber et al., 2009, Rousk et al., 2010, Liu et al., 2018). Moreover, Actinobacteria and Bacteroidetes are typically positively correlated with pH

(Lauber et al., 2009, Zhang et al., 2017c, Wang et al., 2019) and it is known that the addition of animal waste can increase soil pH (Hammesfahr et al., 2011, Abubaker et al., 2013, Zhang et al., 2017c). Principal component analysis of soil physiochemical properties (Figure 3.12) is consistent with the notion that pH is a key factor distinguishing the two sites. Extractable calcium and magnesium, representing major exchangeable cations in soil were correlated with pH, in addition to sodium. The latter is unsurprising since cattle manure can contain a substantial quantity of soluble salts (Hao and Chang, 2002, Unc and Goss, 2004). Finally, Spearman's Rank correlation analysis inferred significant, albeit weak positive associations between Gemmatimonadetes, Bacteroidetes and pH (Figure 3.14). While there was a trend between Actinobacteria and pH, this was not deemed significant after false discovery rate adjustment. Analysis of differential abundance based on physiochemical covariates with Corncob modelling indicated pH was the primary determinant of phylum level differences between sites (Figure 3.13). Indeed, Zhang et al. (2017c) proposed fertilisation alters soil microbial communities through pH change; a thought echoed by Liu et al. (2018) who suggested the influence of pH exceeds the impact of fertilisation-associated nutrient input in acidic soils.

Despite indications that pH and related factors explain some of the key differences in phyla in slurry-impacted and untreated soil, no significant change in pH was identified in soil shortly after slurry application events. As the average pH of treated soil (~7.1) was not dissimilar from that of the applied cattle slurry (~7.4), this potentially signifies a steady state has been reached whereby no further gains in pH are possible. However, without historic data extending back before fertilisation commenced at the site, it is impossible to confirm this supposition. It remains important to acknowledge that the distinction in pH may in fact relate to inherent differences in the soil substrate such as clay content, and bear no relation to slurry application.

According to differential abundance analysis, *Acholeplasma* was the only genus representative of Tenericutes elevated in slurry-impacted soil relative to untreated soil. Similarly, only *Gemmatimonas* and *Gemmatirosa* of Gemmatimonadetes were more prevalent in treated soil compared to untreated soil. Conversely, many Actinobacteria were apparently enriched in treated soil, although these generally belonged to Micrococcales or Propionibacteriales. Within in these orders, members of Intrasporangiaceae and Nocardiaceae, respectively were most notably increased. Meanwhile, the most differentially abundant genera belonging to Bacteroidetes included members of the family Porphyromonadacea; *Fermentimonas* and *Petrimonas* specifically.

When considering the comparative relative abundance of these genera in slurry, some interesting patterns become visible. For instance, *Acholeplasma*, *Fermenitmonas* and *Petrimonas* are all genera which are markedly more abundant in slurry, to such an extent that their increased prevalence in slurry-impacted soil could indicate bacterial transfer. Consequently, while phylum level differences can be useful to gauge broad contrasts in bacterial communities, these findings imply few genera within these phyla are largely responsible for some of the described contrasts between sites. Examination of genus level data immediately before, after and 56 days following the first application of slurry in 2017 shows slurry-borne bacteria were successfully detected in treated soil. Analyses also demonstrated the populations of all enriched bacteria declined within 56 days, however some remained above pre-treatment levels. The most compelling putative taxonomic markers of slurry application will now be discussed.

Fermentimonas, *Petrimonas*, *Sedimetibacter*, *Proteiniphilum* and *Sphaerochaeta* exhibited the largest increase in relative abundance within 24 hours of slurry application and therefore represent convincing biomarkers of slurry contamination (Figure 3.10). Furthermore, all of the aforementioned genera were highly abundant in slurry samples (Figure 3.8). Moreover, *Sphaerochaeta*, *Fermentimonas* and *Sedimentibacter* were among the most over-represented genera in slurry relative to untreated soil. A subset of

these slurry biomarker genera were recently discovered and are closely related members of Porphyromonadacea, which have been recovered from broadly comparable environments (Hahnke et al., 2016). On the one hand, *Proteiniphilum*, *Petrimonas* and *Fermentimonas* have previously been found in cattle slurry (Pandey et al., 2018, Coelho et al., 2020, Im et al., 2020), while *Fermentimonas* sp. are also known to occur in the human gut (Beye et al., 2018). These three genera were additionally identified in mesophillic biogas reactors; which often utilise pig and cattle manure as feedstock (Hahnke et al., 2016, Wolters et al., 2016, Tomazetto et al., 2018, Im et al., 2020, Prem et al., 2020). Similarly, *Sedimentibacter* were reportedly associated with biogas reactors and cattle slurry (Wolters et al., 2016, Habtewold et al., 2018, Prem et al., 2020). The occurrence of *Sphaerochaeta* in cattle slurry has already been discussed. It is therefore clear that the proposed slurry-biomarker genera are associated with the gut and faecal waste of animals rather than soil.

Having demonstrated particular slurry-borne bacteria were transferred to soil from the slurry and potentially augment existing populations, their long-term fate in soil remains to be considered. Interestingly, although the indicator organisms previously highlighted are consistently associated with animal waste products, few publications have reported their dissemination in soil following fertilisation with these substrates. Nonetheless, in a 16S rRNA amplicon study, Wolters et al. (2018) showed *Proteiniphilum* and *Sedimentibacter* were significantly enriched in field plots six days after the addition of swine manure or digestate. Meanwhile, *Petrimonas* was only elevated in manure-amended soil; presumably due to the comparatively lower abundance reported in digestate. Samples collected five months after fertilisation revealed the initial bloom of allochthonous bacteria had dissipated, with the exception of certain digestate-associated Clostridia. In this regard Wolters et al. (2018) corroborates a pervasive ecological paradigm which suggests that indigenous soil microbes typically outcompete those introduced through anthropogenic activity; thus soil microbial communities are believed to exhibit a degree of resilience to perturbation. In line with this, Podmirseg et al. (2019)

found in a PCR-DGGE-based study that soil community structure was largely unaffected by the addition of raw cattle manure or its digestate, and that manure-associated pathogens returned to pre-treatment concentrations within three months. The findings of an earlier 16S rRNA amplicon microcosm experiment parallels their results, showing the rapid loss of manure-borne organisms over two months (Leclercq et al., 2016). Likewise, recent field work by Coelho et al. (2020) showed soil fertilisation with cattle manure elicited only transient changes in bacterial metagenome libraries. In these studies, there appears to be a good level of overall cross-method agreement. Laboratory-based inoculation studies have also demonstrated competition with indigenous soil and predation by bacterivorous protozoa can attenuate invading bacterial pathogens (Acea et al., 1988, Recorbet et al., 1992, Xing et al., 2020). Studies have also shown improved survival of introduced bacteria in gamma-irradiated soils (Chen et al., 2017, Podmirseg et al., 2019, Xing et al., 2020).

In contrast, others have documented phylum-level changes in soils under long-term fertilisation (Sun et al., 2015, Zhang et al., 2018b), while contaminating enteric pathogens can survive for weeks to months (Chee-Sanford et al., 2009). Potential explanations for differing findings could relate to the fact that many studies focus specifically on faecal indicator organisms and the survival of allochthonous, non-pathogenic bacteria may have been missed. Alternatively, analyses which only consider dominant taxa may not capture subtle, yet significant changes in low abundance bacteria (Liu et al., 2017). In addition, Moynihan et al. (2015) provided evidence that while enteropathogen survival was governed by interactions with indigenous microbes; these varied markedly between site and land-use. Lastly, long-term and short-term effects need to be considered in context, since shifts in different indigenous microbial community members may occur at different rates (e.g., depending on microbial life strategies). Such responses could be further modulated by the frequency of exposure. Accordingly, greater attention should be paid to the distinction between short- and long-term effects of perturbation. Regardless, many factors are likely to play a role in

determining the persistence of allochthonous bacteria in amended soils, both directly and through modulation of their autochthonous competitors. For example, research investigating the response of soil microbial communities to cattle slurry (Abubaker et al., 2013) and sewage sludge (Zhang et al., 2018a) found the community structure of clay-based soils were more robust to change compared to sandy soils.

Furthermore, exposure to ionising solar radiation (Unc and Goss, 2004, Hodgson et al., 2016, Jang et al., 2017), temperature (Cools et al., 2001, Wang et al., 2004, Park et al., 2016), manure moisture content (Unc and Goss, 2003) and soil moisture (Wang et al., 2004, Park et al., 2016) are also known to impact the survival of introduced bacteria. It is therefore unsurprising that season is similarly important, as it represents interplay between nearly all the aforementioned factors. This also underscores the difficulty in pinpointing specific environmental drivers behind the data obtained from field studies. Nevertheless, it simultaneously establishes the fundamental need for such research. Finally, much work on survival places emphasis on model organisms such as *Escherichia coli*, which may not be representative of other manure-borne organisms.

The results of the present study suggest the vast majority of slurry-borne bacteria are incapable of establishing in soil. For instance, the relative abundance of typical faecal indicator organisms such as *E. coli* and *Enterococcus* sp. were not perceptibly elevated in amended soil even within 24 hrs of application. Moreover, *E. coli* was significantly more abundant in the untreated site soil than in the slurry-impacted site soil according to Corncob analysis ($p < 0.01$). Indeed, the confounding effect of environmental *E. coli* on the use of this species as a marker for faecal contamination is a topic of discussion (Jang et al., 2017).

The comparatively greater moisture availability in the untreated site (Figure 3.12) may also explain the disparity, as the proliferation of *E. coli* is known to benefit from increased soil moisture content (Jamieson et al., 2002, Sinton et al., 2007), although

saturation may be detrimental (Van Elsas et al., 2011). Alternatively, the substantial dilution effect concomitant with slurry application may have rendered the input of these organisms undetectable. In a review by van Veen et al. (1997), it was also suggested that a high number of viable cells would be necessary to colonise an otherwise hostile soil environment.

The finding that *Sphaerochaeta* was the dominant genus in slurry lends some credence to this idea; however, the likes of *Petrimonas* were comparably abundant to *Entrococcus* in slurry and were still clearly elevated on the day of amendment. Consequently, while it is logical to assume the abundance of bacteria in treatments plays a role in their initial detection in soil, these findings may indicate fitness potential in soil is still an important factor for detection within the first 24 hours in the present work.

Focussing on the relative abundance of the key marker genera as modelled by Corncob, it is evident that they behave in a similar fashion throughout the course of the study (Figure 3.10). The pronounced spike within 24 hours of slurry application (T2) signifies the initial input of slurry bacteria, rather than their proliferation. The variability of abundance estimates at T2 probably correspond to the uneven distribution and pooling of slurry on the soil surface. Interestingly, variability in data has been linked with dysbiosis in animal microbiota associated with disease states (Carding et al., 2015, DeGruttola et al., 2016).

A marked reduction in relative abundance was observed 56 days (T3) after slurry amendment. Nonetheless, count data modelling suggests many of the marker genera were still (albeit weakly) enriched compared to samples taken before the first application of the year. *Petrimonas* is the strongest example of this.

Forty-eight days after the second slurry application of 2017 (T4), certain marker genera such as *Proteiniphilum*, were as prevalent as in T3, which may be expected as the latter

represents a comparable passage of time after slurry application. Other genera such as *Fermentimonas* show greater variability. Critically, these data seem to show that slurry indicator organisms did not accumulate in a step-wise manner after slurry application following an approximate two month window.

Data for the third and final slurry application of 2017 (T5) indicates further decline in the relative abundance of all marker genera. At first this appears counterintuitive as samples were collected only 17 days after slurry application; however, a few factors could explain these findings. Firstly, the rate of slurry application in October 2017 was half that broadcast in May (60 m^3 $hectare^{-1}$) and somewhat less than that in July (40 m^3 h^{-1}). Accordingly, it has been shown that the application rate of swine slurry can influence the survival of introduced faecal indicator organisms (FIOs) (Rufete et al., 2006). Additionally, the level of solar radiation on the date of the third slurry application was higher than any other application date, even when taking into account reduced daylight hours. Naturally, it is difficult to quantify the direct biological impact of this exposure difference, although it seems plausible that solar radiation would exert its greatest influence on bacterial survival in the hours and days directly following application. In any case, Hodgson et al. (2016) showed the survival of FIOs could be enhanced by shallow injection, which would circumvent exposure to ionising solar radiation. A similar trend in bacterial survival was shown by Hutchison et al. (2004). Hodgson et al. (2016) also highlighted the persistence of FOIs were greatest during an October application, when mean UV levels were lower than in May. In the present study, application occurred on the day of maximal solar radiation exposure in October. Finally, other unconsidered variables may also contribute to these unexpected results.

Analyses of indicator data from samples collected towards the end of the closed period in January 2018 (T6) largely reflect the cessation of slurry applications for over 98 days. In line with expectations, the abundance of most indicator genera is within the range of samples taken prior to the first treatment of 2017. Of equal interest is perhaps the

reduced dispersion of indicator data relative to other time points, which possibly further signifies the absence of anthropogenic disturbance or more general microbial dormancy. Regardless, the abundance of indicator genera mirrored a nearby site with no known history of animal waste amendment, providing additional evidence of a return to base levels.

In contrast to 2017, the first slurry application of 2018 occurred at the end of February. This provided an opportunity to examine the effects of commencing slurry applications earlier in the year. Differential abundance analyses of samples collected from treated soil 84 days after exposure (T7) showed a relative increase in the estimated abundance of the marker genera compared to both T5 (October 2017) slurry-impacted soil and T6 (January 2018) untreated soil. Accordingly, NMDS of all genus-level data for both sites in January and May 2018 highlighted that samples were typically more dissimilar to untreated soil in May than January (Figure 3.5). Taken together, analyses assessing within site and between site differences across the two time points in 2018 underpin the characterisations of indicator genera already inferred by the slurry-impacted site data for 2017. Perhaps more interestingly, the enrichment of indicator genera appears to be sustained over an extended period of nearly three months subsequent to amendment in February 2018, whereas a marked decline was evident over a shorter time span following a higher rate of application in May the previous year.

The finding that enrichment was maintained for a prolonged period after amendment in February supports prior research highlighting improved survival of allochthonous bacteria in autumn and winter (Hodgson et al., 2016). This has been attributed to a range of climatic factors, many which have already been mentioned (e.g. increased moisture and reduced temperatures and solar radiation). Genus level network analysis of treated soil and slurry provides additional evidence of associations between the described marker genera, as well as an auxiliary group of bacteria similarly associated with slurry (primarily members of Colstridiales - see Figure 3.17). Although these bacteria were also

featured in the untreated soil network, the connections shared among the marker genera in slurry and slurry-treated soil were not featured.

It is necessary to display caution around the interpretation of less pronounced shifts in the data on account of within-field variation and the complexity of the system, however the sampling and extraction methods should provide results which capture spatial variation more faithfully than studies where one extraction is performed per homogenised soil sample as was done for by Wolters et al. (2018). Ultimately a greater sampling effort would help resolve these changes with more confidence. Another caveat pertains to the inability to distinguish extracellular DNA from that contained within living (or recently living) cells (Carini et al., 2016). As a result, increased read counts may not exclusively represent an increase in living cells. Nonetheless, the context provided by the site metadata and temporal sampling supports the interpretations drawn here. Future studies stand to benefit from metatranscriptomics, however, even this approach is not without its limitations.

3.5 Conclusion

In summary, the present work demonstrates the intensive sampling strategy was able to resolve key indicators associated with slurry in amended soil, chart their subsequent decline and propose environmental factors responsible for the observed trends. Field-scale differences between sites with differing management history were explained most effectively by pH, which may or may not have been altered by successive applications of slurry. However, these differences were apparently discrete from the short-term impacts of slurry addition which were characterised by transient pulses in indicator genera. In agreement with published work, select slurry-borne organisms were shown to be enriched in soil following fertilisation with slurry. Furthermore, the timings of fertilisation events appear likely to modulate the survival of slurry-borne bacteria. Meanwhile, the application of slurry in October underscored the difficulty of predicting the behaviour of

slurry-borne bacteria in a real-world setting and the importance of collating meta-data which are as complete as possible. Carrying out multi-year analyses with three of more field sites, including contrasting sites with no history of slurry application, would serve to confirm findings about the effects of amendments in winter and whether the indicator genera described here are more widely representative of slurry tanks. It is quite possible individual farms represent micro-ecological silos which negate the successful use of 'global' indicator organisms.

Chapter 4
ARG-Taxon Interactions, the Mobilome and Risk

4.1 Introduction

Over the last two decades, environmental resistomes and their response to perturbation events have become subjects of intensive research. This is especially evident in the context of land-application of animal waste products such as solid manure (Kyselková et al., 2015b, Hu et al., 2016, Han et al., 2018, Chen et al., 2019a), liquid slurry (Sengeløv et al., 2003, Abubaker et al., 2013, Joy et al., 2013, Cheng et al., 2019) and anaerobic digestate (Abubaker et al., 2013, Sui et al., 2016). Likewise, similar efforts have also been made to characterise shifts in soil bacterial communities following organic fertilisation (Parham et al., 2003, Hammesfahr et al., 2011, Johansen et al., 2013). With regard to animal waste, especially faecal material, emphasis has often been placed on tracking the survival of key human pathogens and faecal indicator organisms (FIOs) in soil (Fenlon et al., 2000, Pourcher et al., 2007, Hodgson et al., 2016), since these pose a clear, quantifiable risk to human health via contamination of vegetable produce and groundwater (Sharma and Reynnells, 2016). Indeed, FIOs are integrated within water quality standards throughout the world (Fewtrell and Bartram, 2001).

In contrast, a widely adopted framework for defining the risk to human health from environmental AMR is lacking (Bengtsson-Palme et al., 2018). Nonetheless, such a framework for prioritising risk has frequently been called for and discussed (Larsson et al., 2018, Ben et al., 2019). The absence of a standardised means to quantify 'AMR risk' can largely be attributed to the complexity of capturing and untangling the morass that ARG dispersal represents. In brief, horizontal gene transfer (HGT) enables ARGs to be exchanged between bacteria which may originate from different sources or exhibit contrasting tendencies for human pathogenicity. To this end, Maeusli et al. (2020) demonstrated the transfer of antibiotic resistance plasmids from lettuce-borne

Acinetobacter bayllyi (a non-pathogenic environmental bacterium) to clinical *E. coli* isolates which were subsequently able to colonise mice intestines and further disseminate the target plasmid to gut-resident *Klebsiella pneumoniae*. Moreover, rates of genetic exchange are not constant across species, and certain colonisation-potentiating traits such as capsule production, have been linked to increased rates of HGT and the acquisition of ARGs (Rendueles et al., 2018). However, there are apparent limits to HGT, since lateral genetic exchange is thought to primarily occur within phyla; with inter-phyla HGT occurring only rarely (Jiang et al., 2019). In addition, Lehtinen et al. (2020) presents evidence that the occurrence of antibiotic resistance in *Streptococcus pneumoniae* (for which capsule production is incidentally a colonisation factor) may not be driven primarily by the rate of HGT between lineages, but rather the environmental selection of variants in the host. Such host-based selection pressures include immune action, competition with co-resident microbes and exposure to antibiotic therapy.

Meanwhile, prioritising risk among ARGs is equally problematic. The range of ARGs recovered from substrates of ancient origin is continually broadening, with controversial indications that even ARGs linked to synthetic drugs could have prehistoric origins (Okubo et al., 2019). There is no doubt many ARGs are ancient and seemingly environmentally omnipresent (D'Costa et al., 2011, Perry et al., 2016). Consequently, the mere detection of ARGs in the environment does not necessarily signify a cause for concern. Moreover, Wright (2019) highlights certain multidrug efflux pumps and other putative ARGs only confer resistance when over-expressed and therefore their significance can only really be defined with access to transcription data. This is not to say environmental ARGs should be dismissed, rather that their risk to human health as resistance genes is defined by a full awareness of their context (Martínez et al., 2015, Wright, 2019). In light of this, Martínez et al. (2015) proposed that intrinsic ARGs and multidrug efflux pumps should only be considered in resistome analyses if they could be associated with mobile genetic elements. On the other hand, if the prevalence of a multidrug efflux gene were to become fixed at a higher level in the environment, this

could surely constitute an elevated risk of mobilisation in the future. Furthermore, the emergence of new ARG combinations may potentiate the clinical significance of some otherwise intrinsic resistance genotypes. Lastly, the non-specificity of some multidrug efflux pumps can have implications for cross-selection via heavy metal and biocide exposure (Pal et al., 2015, Zhou et al., 2016, Pal et al., 2017).

Consequently, calculating the risk posed by ARG-host combinations remains a formidable challenge and only serves to underscore the importance of improving the characterisation of both the phylogenetic and genetic context of ARGs.

Despite the rapid expansion of high-throughput metagenomic techniques which have opened up opportunities to study the nexus between ARGs, MRGs, MGEs and their hosts, many studies employ 16S-rRNA/HT-qPCR amplicon methods which focus on only one of these components in isolation, or rely on correlation-dependent analyses to infer associations between them (Hu et al., 2016, Zhou et al., 2016, Peng et al., 2017, Zhang et al., 2017b, Chen et al., 2019b). Although valuable, these inferences are rarely validated with culture-work or further molecular evidence. Increasing emphasis is therefore being placed upon the need to move beyond quantification and correlation towards establishing the genetic context of ARGs and MGEs as well as host range (Ma et al., 2016, Rice et al., 2020, Zhao et al., 2020). Doing so is likely necessary for the creation of a tractable risk assessment system.

However, even relatively new molecular techniques present many issues to overcome in this regard. More specifically, Slizovskiy et al. (2020) highlighted different bioinformatic pipelines can yield conflicting associations about the same metagenomic dataset. In particular, the assembly of short-reads into contigs was shown to enhance ARG recovery when compared to read-based analyses. Additionally, the quality of database curation can have a substantial impact on results. Foremost concerns are the presence of redundant sequences, gene definitions and biases in scope. Moreover, the

aforementioned considerations apply to all databases, whether they describe ARGs, MGE or taxa.

Accordingly, the present work aimed to further elucidate associations between antimicrobial resistance genes, MGEs and bacterial taxa in slurry and soil by adopting an ensemble approach to analysing short-read metagenome shotgun sequence libraries. This involved processing analyses of unassembled read data which were further supplemented by interrogating the corresponding assemblies and metagenome-assembled genomes (MAGs).

4.2 Materials and Methods

4.2.1 Approach, Sampling and Sequencing

Antibiotic resistance gene annotation (ARGs) and taxonomic assignment were carried out on the same metagenomic libraries discussed previously (Chapters 2 and 3). Full descriptions of sample sites, sampling strategy and sequencing are therefore available in the aforementioned chapters.

ARG annotation and taxonomic assignment was performed on short-read libraries as discussed in Chapter 2 and 3, respectively.

4.2.2 Exploratory Statistics and Risk Scores

Procrustes analysis was used to assess the potential relationship between ARG subgroups and the bacterial community composition of samples (Forsberg et al., 2014, Fang et al., 2019). Using vegan in R, the MetaMDS function (Bray-Curtis distance) was applied to 16S-normalised ARG subgroup abundances and bacterial genus relative

abundances. Procrustes was then performed on the two NMDS objects (999 permutations) and significance tested with the *protest* function. Only soil data from 2018 were used for this analysis due to MetaMDS non-convergence when attempting to include slurry samples.

MetaCompare (Oh et al., 2018) was used to project individual sample assemblies into a 3D hazard space representing the cumulative risk posed by the presence of mobile genetic elements (MGEs), ARGs and clinically defined pathogens on contigs. To date, MetaCompare is the only purpose-built pipeline which attempts to distil 'risk scores' from metagenomic data (Slizovskiy et al., 2020).

4.2.3 ARG-Taxonomy Network Analysis

Bipartite network analysis of short-read ARG and taxonomic data was conducted using CoNet and visualised with Cytoscape. Specifically, count data for ARG subgroups and bacterial genera were converted into relative abundances (percentage of total reads) before association mining with CoNet. CoNet and Cytoscape were used in preference to SpiecEasi (see Chapter 2) due to the improved speed and visual handling of very large networks. In an attempt to leverage maximal statistical power across the dataset, all sample types were included in the analysis (slurry-impacted soil (treated site), unamended soil (untreated site) and slurry). However, ARGs and genera which occurred in fewer than 10 samples were excluded to minimise spurious associations. Connections were only mapped if they were considered significant ($p < 0.05$) and were correlated via Spearman's Rank ($R > 0.8$). Furthermore, instances where correlations were supported by fewer than 10 non-null values were also excluded. Finally, although CoNet does not incorporate a sparsity assumption like SpiecEasi, it does employ a re-normalisation step to mitigate the effects of compositionality.

4.2.4 Metagenome Assembly

As previously discussed in Chapter 2, the assembly of short-read metagenomic libraries can enhance downstream analyses by generating longer sequences (contigs), which ostensibly contain more information (Chen et al., 2020, Pérez-Cobas et al., 2020). This can be likened to joining several individual pieces of a puzzle into larger, albeit still incomplete sections. However, the process of assembly currently exhibits limited proficiency in resolving stretches of sequences with unusual characteristics such plasmids (Maguire et al., 2020). Nonetheless, assembly is typically required to extract metagenome assembled genomes (MAGs) from short-read libraries. A plethora of assembly software is available, each with their own strengths and limitations. For *de novo* assembly, these broadly fall into two dominating heuristic paradigms; De Bruijn graph and overlap–layout–consensus (OLC) methods (Pérez-Cobas et al., 2020). The former has become favoured for short-read libraries due to the reduced computational requirements of the graphing system (Li et al., 2015b, Van der Walt et al., 2017). Even among De Bruijn graph assemblers there are a number of options; including Megahit (Li et al., 2015b), MetaSPAdes (Nurk et al., 2017) and IDBA-UD (Peng et al., 2012) among many others.

A review of literature suggests that Megahit is capable of resolving complex sequencing data such as soil metagenomes (Li et al., 2015b, Van der Walt et al., 2017), although MetaSPAdes has been shown to out-perform Megahit to varying degrees in other publications (Nurk et al., 2017, Wang et al., 2020b). Overall, Megahit was deemed to represent an appropriate trade-off between assembly integrity and computational efficiency for this study.

Megahit was therefore used to generate individual assemblies for each sample library (sample-specific contigs) and three co-assemblies of samples with shared origin (slurry-impacted soil, untreated soil and slurry, respectively). A final 'meta-soil' co-assembly including all soil samples from both slurry-impacted and the undisturbed site was

created. Potential benefits of co-assembly include increased read-depth, which may enhance recovery of low abundance MAGs (Hofmeyr et al., 2020); however, individual assemblies have also been shown to provide improved MAG recovery and strain preservation in work involving thousands of metagenomes (Nayfach et al., 2019).

Individual assemblies were computed using a dedicated preset for complex metagenomes termed 'meta-large' as defined in Megahit (Li et al., 2015b). In brief, the preset specified a kmer list ranging from 27-87 in intervals of 10 (k-step) and a minimum multiplicity for filtering of 2. Co-assemblies of treated soil, untreated soil and slurry were run with a kmer list ranging from 29-89 (k-step = 10) and a minimum multiplicity for filtering of 2. The 'meta-soil' co-assembly was run with a kmer list ranging from 31-97 (first k-step of 9 due to computational constraints; insufficient memory to graph k-27 so k-31 was selected).

4.2.5 MAG Recovery

MAGs were extracted from the separate co-assemblies of treated soil, untreated soil and slurry. Firstly, co-assemblies were filtered to remove contigs <1000bp in length. Subsequently, the quality control-passed reads used to construct each respective assembly were mapped back to the appropriate set of contigs using BWA-mem (Li, 2013). The output alignments were then converted into BAM (binary alignment map) format using SamTools (Li et al., 2009). Contig depth calculations and binning were performed with MetaBAT2 (Kang et al., 2019). The quality of bins was assessed with CheckM (Parks et al., 2015); only bins meeting an estimated quality score ≥ 50 as defined by Parks et al. (2017), were considered for further analysis. To maximise MAG recovery from untreated soil samples, MAGs were ultimately acquired by mapping to the 'meta-soil' co-assembly. This decision was made based on the analyses indicating the presence of broadly similar taxa at both sites and an average ~17% improvement in read mapping.

In an attempt to recover potential candidate MAGs for slurry-to-soil transfer, the above protocol was also carried out on BAM files produced after mapping slurry reads to the treated soil co-assembly. As a rudimentary control, the same process was performed on the untreated soil co-assembly. Lastly, all quality control passed bins were pooled and clustered with dRep (Olm et al., 2017).

4.2.6 Annotation of AMR-determinants in MAGs and Contigs

The presence of assorted antimicrobial resistance determinants in MAGs and contigs was evaluated using ABRicate (Seemann, 2020). ABRicate was coupled with the MEGARes v2.0 (Doster et al., 2020) database to assess ARGs and metal resistance genes (MRGs), while ACLAME v0.4 (Leplae et al., 2010) was used to explore associated MGEs. In an attempt to balance stringency with sensitivity, ABRicate was run with an 80% identity and 60% gene coverage cut-off.

4.2.7 Taxon Assignment of MAGs and Contigs

Taxonomic assignment was conducted on the MAG collection using GTDB-Tk (Chaumeil et al., 2020). Meanwhile, contigs produced by assembling samples individually were assigned a putative taxonomy using Kaiju. However, when supplied a file of contigs, Kaiju will only list the 'best hit' within a contig. Consequently, it is possible for a single, less representative hit to be reported over many more (potentially marginally) inferior hits. Alternatively, Ma et al. (2016) assigned contig taxonomy according to BLAST results which achieved agreement across >50% of ORFs (open reading frames) identified within a contig. In the present study it was decided combining elements of both these methods would be advantageous.

Briefly, ORFs were generated for each sample using Prodigal. Where Kaiju assigned a clear majority of ORFs to the same taxon; this taxon was considered the most appropriate assignment. When no clear winner was apparent, the best match provided by Kaiju for the entire contig was selected. In cases where only a single ORF was detected, the contig best match was also selected. Finally, when no taxon could be identified in ORFs, the contig best match was used. The decision-making process leading to final taxon assignment was summarised for each contig, enabling further assessment of assignment quality for contigs of particular interest (i.e. instances where a contig best match and majority ORF assignment are congruent, represent more confident putative identities compared to those which do not). The aforementioned voting system was implemented in R.

The challenge of confidently assigning taxonomy to contigs (or reads) with few closely related sequence matches in databases is a perennial issue associated with samples obtained from highly complex and poorly resolved environments. The incompleteness and fluidity of databases are obstacles discussed extensively by Breitwieser et al. (2017). For the purposes of assigning taxonomy to contigs, the single 'best hit' approach used by programs such as Kaiju was critiqued by von Meijenfeldt et al. (2019) who highlighted a tendency for precision loss when classifying increasingly novel sequences. Accordingly, von Meijenfeldt et al. (2019) devised an alternative, termed CAT (contig annotation tool). However, the aforementioned study did not implement a post-processing voting step for Kaiju in benchmarking, unlike the present work. Nonetheless, given the currently subjective nature of taxonomic classification, CAT was run on contigs of interest in an effort to further cross-validate results.

4.2.8 Contig Result Collation

Contigs with at least one ARG or MRG as identified by ABRicate were parsed into a single data matrix. Mobile genetic elements associated with RG-containing contigs were then

appended to the matrix. Similarly, Kaiju vote and CAT taxon assignments were paired with respective contigs. This was achieved by assigning a unique ID to each contig consistent across all analyses. To aid visualisation, the summary matrix was sorted into 80 initial clusters using the PAM (partition around medoids) function in the R package Cluster. Finally, the matrix was reviewed and manually refined where appropriate.

Note: a list of all software used thus far is provided in Appendix 2.

4.3 Results

4.3.1 Exploratory Statistics and Risk Scores

Significant similarities between bacterial taxa and resistome data in samples collected from treated and untreated soils in 2018 were indicated by Procrustes analysis (Figure 4.1; $p < 0.001$; $R = 0.95$; $M^2 = 0.094$).

Contrary to expectation, resistome risk scores determined by MetaCompare ranked both slurry-impacted soil (22.79 ± 0.06, n = 34) and soil with no history of cattle waste amendment (22.87 ± 0.16, n = 10) above that of slurry (21.61 ± 0.06, n = 10). Regardless of the ultimate risk score assigned, MetaCompare indicated a greater proportion of slurry contigs exhibited co-localisation of ARGs, MGEs and pathogens when compared to soil samples (Table 4.1). Conversely, more ARGs, MGEs and pathogens were observed per contig in soil samples, however these features consistently occurred on separate contigs. While the lack ARG/MGE co-localisation on soil contigs may relate to genuine differences between the mobilome and resistome of soil and slurry samples, it is also possible that the failure to find co-residence in soil corresponds to the fact that much more of the soil data was not contiguously assembled relative to slurry sample data as a result of its greater microbial diversity (see MetaCompare Risk Scores in

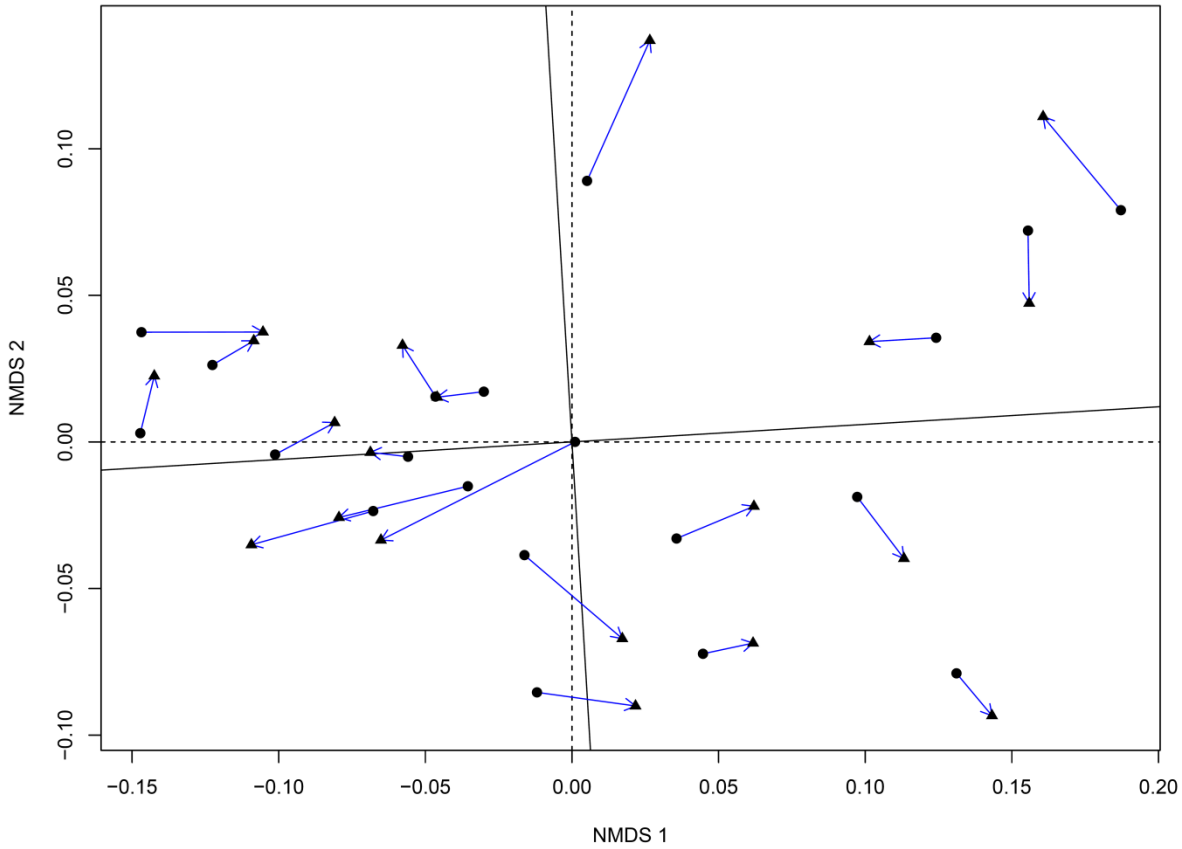


Figure 4.1 Procrustes analyses demonstrating the relationship between bacterial genera (circles) and ARG subgroups (triangles) in soil (2018 data only). Black lines reflect degree of rotation from dotted lines required to achieve best fit. Blue arrows indicate distance between bacterial and ARG data points for the same sample. 999 permutations performed. Correlation method: Spearmans'.

Discussion for further explanation). Overall, no significant difference was observed between soil risk scores, irrespective of slurry treatment history. Once again, potential explanations for this finding include both genuine biological reasons and potentially confounding methodological limitations (see MetaCompare Risk Scores in Discussion for further explanation).

Table 4.1 Metacompare results for all samples including hazard element data. Based on assembled reads, n = number, '+' signifies co-residence on contigs, '/' denotes division. ARG = antibiotic resistance gene, MGE = mobile genetic element and PAT = pathogen (as defined by MetaCompare). Heatmap colours scaled by column (green to red = low to high).

	ID	nContigs	nARG	nMGE	nPAT	nARG + MGE	nARG + MGE + PAT	nARG/nContigs	nMGE/nContigs	nPAT/nContigs	nARG + MGE/nContigs	nARG + MGE + PAT/nContigs	Risk Score
Slurry	S01	887920	1509	194	27195	15	15	0.0017	0.0002	0.0306	1.69E-05	1.69E-05	21.79
	S02	839678	1371	176	26300	15	15	0.0016	0.0002	0.0313	1.79E-05	1.79E-05	21.61
	S09	963276	1500	189	27931	18	18	0.0016	0.0002	0.0290	1.87E-05	1.87E-05	21.41
	S11	1185999	1791	263	33643	15	15	0.0015	0.0002	0.0284	1.26E-05	1.26E-05	21.25
	S17	1138837	1871	243	34180	17	17	0.0016	0.0002	0.0300	1.49E-05	1.49E-05	21.62
	S18	1308655	2122	283	38847	22	22	0.0016	0.0002	0.0297	1.68E-05	1.68E-05	21.57
	S25	1152246	1845	246	35028	20	20	0.0016	0.0002	0.0304	1.74E-05	1.74E-05	21.52
	S26	987061	1696	202	30151	16	16	0.0017	0.0002	0.0305	1.62E-05	1.62E-05	21.83
	S34	1141062	1912	215	33971	23	22	0.0017	0.0002	0.0298	2.02E-05	1.93E-05	21.74
	S35	1226904	2054	246	35769	23	23	0.0017	0.0002	0.0292	1.87E-05	1.87E-05	21.73
Slurry-impacted Site	S78	1480530	3201	424	101907	0	0	0.0022	0.0003	0.0688	0	0	22.96
	S79	1486161	3286	429	107700	0	0	0.0022	0.0003	0.0725	0	0	23.10
	S80	1651037	3601	454	110536	1	1	0.0022	0.0003	0.0669	6.06E-07	6.06E-07	23.02
	S81	1698562	3454	374	100711	2	2	0.0020	0.0002	0.0593	1.18E-06	1.18E-06	22.61
	S82	1530913	3341	363	91060	2	2	0.0022	0.0002	0.0595	1.31E-06	1.31E-06	23.03
	S83	1559208	3393	448	103737	5	5	0.0022	0.0003	0.0665	3.21E-06	3.21E-06	23.03
	S84	1648857	3497	455	106863	3	2	0.0021	0.0003	0.0648	1.82E-06	1.21E-06	22.86
	S85	1634909	3294	448	96780	2	1	0.0020	0.0003	0.0592	1.22E-06	6.12E-07	22.55
	S87	1458537	3296	410	91884	2	2	0.0023	0.0003	0.0630	1.37E-06	1.37E-06	23.26
	S88	1334112	2854	371	87692	0	0	0.0021	0.0003	0.0657	0	0	22.90
	S89	1484556	3327	415	101491	1	1	0.0022	0.0003	0.0684	6.74E-07	6.74E-07	23.20
	S90	1388995	3178	420	96688	2	2	0.0023	0.0003	0.0696	1.44E-06	1.44E-06	23.34
	S91	1614316	3561	466	112014	1	1	0.0022	0.0003	0.0694	6.19E-07	6.19E-07	23.09
	S92	1692936	3727	496	115267	1	1	0.0022	0.0003	0.0681	5.91E-07	5.91E-07	23.08
	S93	1493200	3377	474	102214	1	1	0.0023	0.0003	0.0685	6.70E-07	6.70E-07	23.26
	S94	1573390	3408	395	97888	3	3	0.0022	0.0003	0.0622	1.91E-06	1.91E-06	22.99
	S95	1492077	3267	434	100015	3	3	0.0022	0.0003	0.0670	2.01E-06	2.01E-06	23.06
	S96	1995308	3919	505	117873	3	3	0.0020	0.0003	0.0591	1.50E-06	1.50E-06	22.41
	S97	1663306	3352	434	96157	0	0	0.0020	0.0003	0.0578	0	0	22.55
	S98	2037721	4198	460	114291	1	1	0.0021	0.0002	0.0561	4.91E-07	4.91E-07	22.68
	S99	1629531	3134	415	93243	2	2	0.0019	0.0003	0.0572	1.23E-06	1.23E-06	22.30
	S100	1718414	3377	416	104784	1	1	0.0020	0.0002	0.0610	5.82E-07	5.82E-07	22.41
	S101	1688468	3662	471	108237	4	3	0.0022	0.0003	0.0641	2.37E-06	1.78E-06	23.00
	S102	1270159	2839	412	88986	1	1	0.0022	0.0003	0.0701	7.87E-07	7.87E-07	23.18
	S103	1858681	3786	502	111770	3	3	0.0020	0.0003	0.0601	1.61E-06	1.61E-06	22.62
	S104	1706788	3671	466	104039	1	1	0.0022	0.0003	0.0610	5.86E-07	5.86E-07	22.94
	S105	2016215	3871	448	105215	2	2	0.0019	0.0002	0.0522	9.92E-07	9.92E-07	22.29
	S106	1864120	3492	480	104436	3	3	0.0019	0.0003	0.0560	1.61E-06	1.61E-06	22.16
	S107	1674150	3315	447	97768	4	4	0.0020	0.0003	0.0584	2.39E-06	2.39E-06	22.46
	S113	2406115	4768	631	146816	4	4	0.0020	0.0003	0.0610	1.66E-06	1.66E-06	22.46
	S114	2097298	4183	599	133045	3	2	0.0020	0.0003	0.0634	1.43E-06	9.54E-07	22.50
	S115	1902853	3736	534	120295	3	3	0.0020	0.0003	0.0632	1.58E-06	1.58E-06	22.41
	S116	1752220	3492	448	102333	1	1	0.0020	0.0003	0.0584	5.71E-07	5.71E-07	22.49
	S117	1890925	3813	542	119598	6	4	0.0020	0.0003	0.0632	3.17E-06	2.12E-06	22.57
Untreated Site	S118	1636228	3746	541	94931	2	2	0.0023	0.0003	0.0580	1.22E-06	1.22E-06	23.34
	S119	1362420	3331	424	89246	2	2	0.0024	0.0003	0.0655	1.47E-06	1.47E-06	23.80
	S120	1509716	3201	343	82300	4	3	0.0021	0.0002	0.0545	2.65E-06	1.99E-06	22.86
	S121	1959064	3708	465	115448	1	1	0.0019	0.0002	0.0589	5.10E-07	5.10E-07	22.21
	S122	1532889	3018	364	83196	5	4	0.0020	0.0002	0.0543	3.26E-06	2.61E-06	22.44
	S108	1588893	3532	421	87115	3	3	0.0022	0.0003	0.0548	1.89E-06	1.89E-06	23.15
	S109	2006330	3845	551	122051	2	1	0.0019	0.0003	0.0608	9.97E-07	4.98E-07	22.28
	S110	1770103	3600	458	103678	5	3	0.0020	0.0003	0.0586	2.82E-06	1.69E-06	22.61
	S111	1507651	3186	387	87234	0	0	0.0021	0.0003	0.0579	0	0	22.82
	S112	1449738	3253	463	94964	1	1	0.0022	0.0003	0.0655	6.90E-07	6.90E-07	23.20

4.3.2 ARG-Taxonomy Network Analysis

Correlation-based co-occurrence analysis of putative ARG subgroups and bacterial genera resulted in a large network containing 1233 nodes and 6363 edges (Figure 4.2). Two major clusters were formed, together with several minor clusters. The largest cluster (cluster I) primarily contained a diverse range of MLS, tetracycline and aminoglycoside resistance genes. At the phylum level, cluster I was dominated by genera belonging to Firmicutes, followed by Bacteroidetes. The second most populous cluster (cluster II), was comprised almost exclusively of Actinobacteria. Cluster II contained a select group of glycopeptide, rifamycin and tetracycline ARG subgroups. Overall, Proteobacteria were linked to the largest range of unique ARGs subgroups, while Firmicutes possessed the most connections.

Partitioning ARG subtype nodes belonging to the same antibiotic class showed further phylum level trends in ARG-taxon associations. For example, over half of all edges (n = 116) linked to beta-lactam *cfxA* genes in cluster I represented Bacteroidetes genera. In contrast, Firmicutes contributed only 21.7% of edges (n = 44) connected to *cfxA* genes, despite an overall preponderance in cluster I. Similarly, 11 out of 14 edges linked to *sul* genes belonged to Proteobacterial genera. Tetracycline and MLS ARG subgroups were predominantly linked to Firmicutes (63.2% and 52.6% of edges, respectively), and Bacteroidetes (14.7% and 18.8%, respectively). However, certain subgroups within these classes were tightly associated with other phyla. Specifically, tetracycline resistance subgroups *tet-48* and *otrA* were unanimously correlated with Actinobacterial genera. Likewise, a single MLS subtype *ereB* was most frequently associated with Proteobacteria (67.2% of edges). Interestingly, Delta-proteobacterial nodes were restricted to cluster I and most (58.3%, n = 14) were connected to a single MLS ARG subtype *mefB*; alluding to potential class level trends. In contrast, Gamma-proteobacterial nodes were present in different sub-clusters and were not dominated by an affiliation with any one ARG-subtype, although MLS subgroups were the most prevalent.

Multi-drug efflux pump subgroups were largely dispersed evenly among clusters, although individual subtypes appeared to align with specific phyla. In particular, *mtrA* and *ABC transporter* subtypes were primarily linked to Actinobacterial nodes. Meanwhile, over half the nodes identified as Firmicutes (59.3%, n = 219) were associated with *efrA*. However, it is of note that *efrB*, which is required for *efraAB* efflux pump functionality was not identified in the network (Alcock et al., 2019). The reliability of inferences drawn from network analyses and the inclusion of multidrug resistance genes will be subject to later discussion.

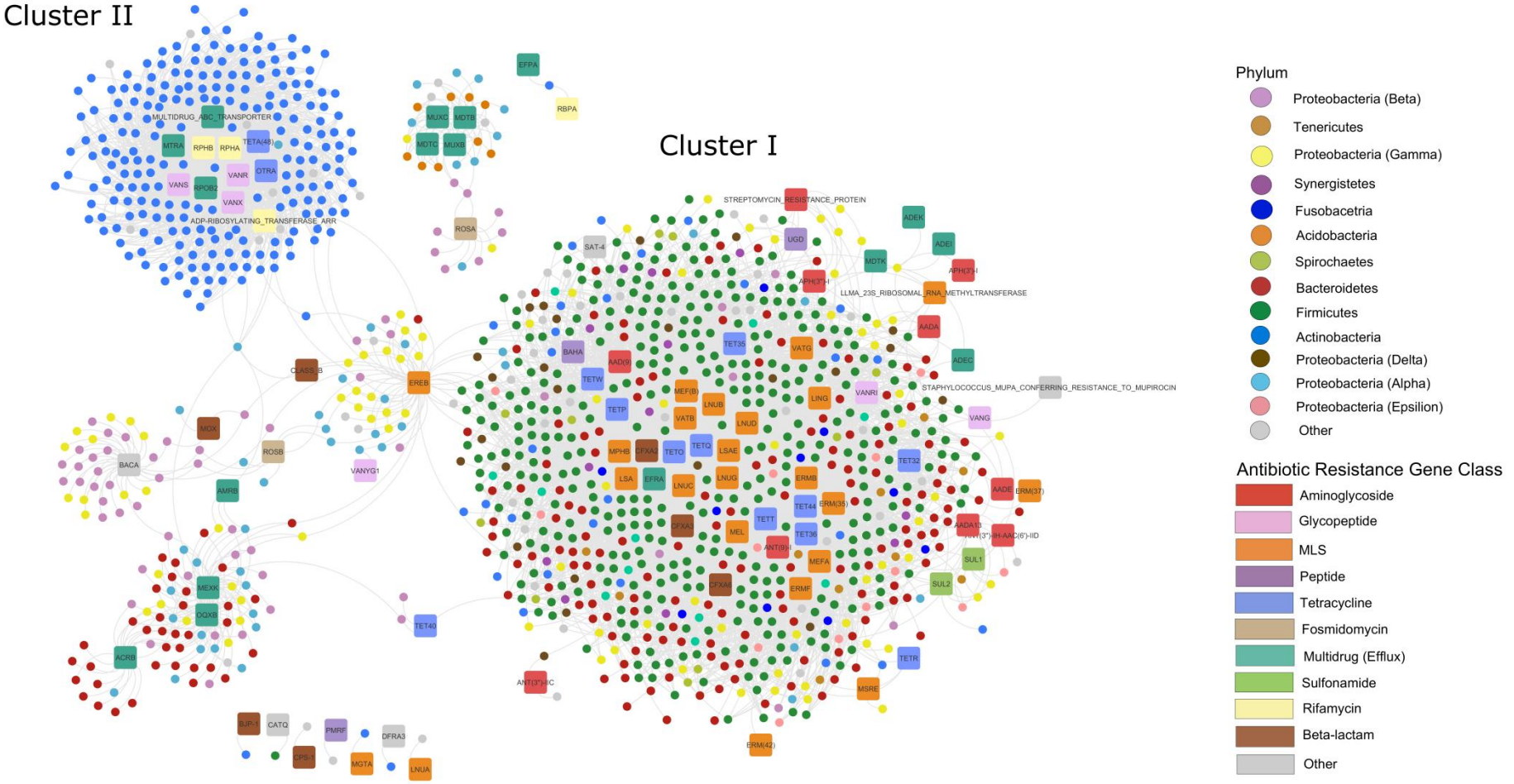


Figure 4.2 ARG-Taxon correlation network. Square nodes represent ARGs and circular nodes denote genera. See legend for colour descriptions. Spearman's Rank correlations R > 0.8, fdr-adjusted p < 0.05. Bootstrapped 100 times; unstable edges removed. Minimum occurrence of all nodes is 10 samples.

4.3.3 MAG Analyses

A total of 165 MAGs extracted from soil and slurry samples met quality control requirements. According to GTDB-Tk, the majority of these were of bacterial origin (n = 156), while a small number of MAGs were identified as Archaea (n = 9). Slurry samples yielded the largest number of quality-passed bacterial MAGs (n = 130), while soil samples (both slurry-impacted and undisturbed samples combined) produced far fewer (n = 25). The taxonomy of extracted MAGs can be found in Supplementary file 7.

Archaeal MAGs in slurry predominantly belonged to the family Methanomethylophilaceae of phylum Thermoplasmatota. In contrast, soil archaeal MAGs represented the family Nitrososphaeraceae of phylum Crenarchaeota.

The taxonomy of bacterial MAGs in slurry and soil was largely congruent with dominant phyla described by unassembled short-read data analyses (Chapter 3, Figure 3.3). Specifically, 68.5% of slurry-derived MAGs were assigned to either Firmicutes (n = 53) or Bacteroidetes (n = 36). Meanwhile, Proteobacteria accounted for just under half the MAGs recovered from soil samples irrespective of management history. Interestingly, MAGs belonging to Gemmatimonadetes and Nitrospirae were only identified in slurry-treated soil and unamended soil, respectively. On the other hand, no genera were clearly dominant as described in short-read data analyses; although the genera identified were still indicative of sample type (e.g. *Intestimonas* sp. and *Sphaerochaeta* sp. MAGs in the slurry collection).

Mapping slurry sample reads to the treated soil co-assembly resulted in the recovery of a single MAG meeting quality control requirements. GTBD-Tk classified the MAG as *Proteiniphilum* sp. and is consistent with the suggested bacterial markers of slurry application as indicated by prior differential abundance and network analysis (Chapter 3 Figures 3.10 and 3.17). Although MAGs are known to represent an agglomeration of closely related genomes, MASH-ANI (average nucleotide identity) analysis with dRep showed two MAGs clustered at >99.5% similarity (Figure 4.3A). These MAGs were

independently generated from the slurry and slurry-impacted soil co-assemblies. Furthermore, the single *Proteiniphilum* sp. MAG acquired from mapping slurry reads to the treated-soil co-assembly also fell within this cluster. The aforementioned *Proteiniphilum* sp. MAGs formed the only cluster detected in the collection (Figure 4.3B).

ABRicate failed to detect any ARGs within the MAG collection. This result remained unchanged even when stringencies were relaxed. Consequently, it was decided that focus would be placed on analysis of assembled contigs.

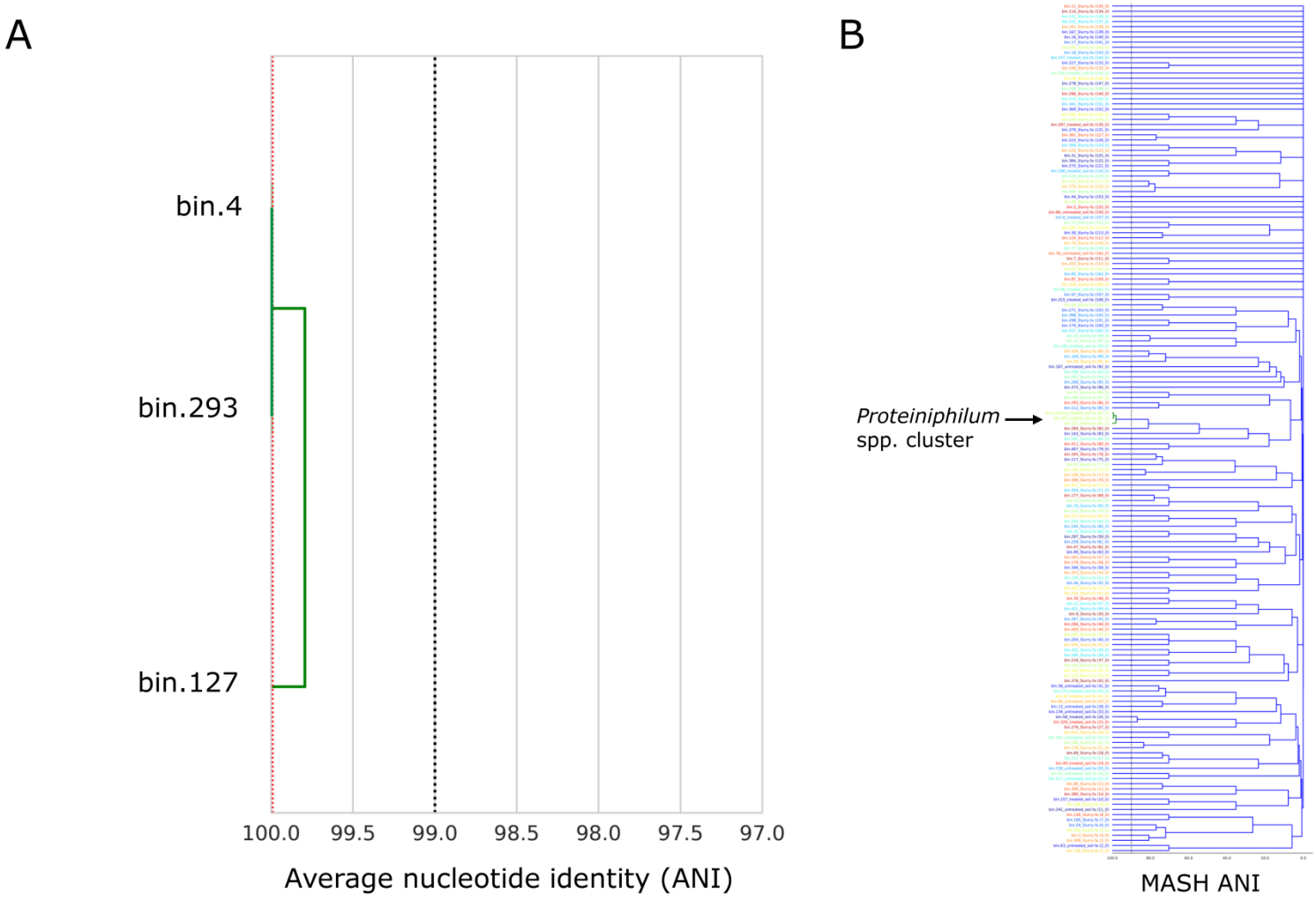


Figure 4.3 *Proteiniphilum* sp. ANI cluster plot (A); bin 4 generated from slurry reads mapped to treated soil co-assembly, bin 293 generated from slurry-impacted soil reads mapped to slurry-impacted soil co-assembly, bin 127 generated from slurry reads mapped to slurry co-assembly). Quality-control passed bacterial MAG cluster analysis based on average nucleotide identity (B).

4.3.4 Contig Analyses

Across the entire dataset only a small proportion of contigs were shown to harbour at least one resistance gene according to ABRicate screening (n = 1080; <0.001% of contigs). In addition, only 16 contigs contained multiple resistance genes. The frequency of contigs containing various combinations of ARGs, MGEs and MRGs is summarised in Table 2.2. Only contigs of particular interest will be presented in text.

Table 4.2 Summary of contigs with at least one ARG or MRG according to ABRicate. Note, MGEs are only reported here if they were detected in combination with either an ARG or MRG. ABRicate found no MRGs co-resident with ARGs.

Contig Library	**≥1 ARG**	**MRG**	**ARG + MGE**	**>1 ARG**
Slurry (10.8M)	483	24	57	16
Treated Soil (57.4M)	420	41	1	NA
Untreated Soil (16.3M)	102	10	NA	NA

4.3.4.1 Slurry

Contig screening showed the core resistome of slurry was dominated by the MLS resistance gene *lnuC* (n = 114), followed by *ant-6* (n = 39), an aminoglycoside resistance gene. A trio of ARGs occurred with lesser frequency; including the *cfxA* (n = 29) ambler class A beta-lactam resistance group, *mefA* (n = 28) MLS resistance group and *ant-3* (n = 27) aminoglycoside resistance group. The remaining genes featured on < 20 contigs. Dominant slurry ARGs were infrequently associated with MGEs. For example, *lnuC* was not identified on any contigs containing MGE genes. Nonetheless, MGE genes were shown to co-localise with less frequently detected ARGs. The co-presence of *tetM* with several genes associated with Tn*916* conjugative transposons exemplifies this.

The congruence of slurry contig taxonomic assignments harbouring dominant ARGs was variable, although Kaiju and CAT were broadly in agreement. For instance, *cfxA* was typically assigned to the order Bacteroidales by Kaiju and more conservatively, phylum Bacteroidetes by CAT. In contrast, *ant-6* was generally located on contigs which neither

Kaiju or CAT could classify beyond the Bacterial domain. Meanwhile, the most prevalent ARG on slurry contigs *lnuC* was associated with various taxa within the phylum Firmicutes. Most frequently these belonged to order Clostridiales, although on one occasion Kaiju and CAT concurred at species level with *Streptococcus uberis*. When considering *mefA,* CAT was often unable to classify contigs carrying this gene, however where Kaiju proposed *Clostridia intestinales*, CAT provided partial corroboration with order Clostridiales. Lastly, Kaiju suggested a range of Gamma-proteobacteria carried *ant-3*, however, close agreement was only found between CAT and Kaiju where *Acinetobacter* sp. were concerned.

4.3.4.2 Soil

The glycopetide resistance gene *vanRO* was the most frequently identified ARG in soil contigs, followed by rifamycin resistance in the form of *rbpA*. The multidrug resistance gene *mtrD* and trimethoprim resistance gene *dfrB* were also detected. The aforementioned genes represented the four most prevalent ARGs in both sites irrespective of slurry application history. The remaining ARGs were identified on <10 soil contigs. No robust evidence of MGE-ARG co-localisation was found in soil contigs.

CAT was unable to classify most *vanRO* contigs and often failed to agree with Kaiju, even at phylum level when an assignment was achieved. However, where CAT and Kaiju concurred, an assignment within phylum Actinobacteria was always identified. In contrast, *rbpA* was consistently found on contigs classified within Mycobacteriaceae by both CAT and Kaiju. Likewise, *mtrA* efflux genes were found on Actinobacterial contigs classified as Mycobacteriaceae and Pseudonocardiaceae. On the other hand, *dfrB* was located on contigs which CAT was often unable to classify. However, *Methylibium* sp. *CF059* was assigned by both CAT and Kaiju in several instances regardless of site origin.

4.3.4.3 Metal Resistance Genes

Contigs containing metal resistance genes (namely copper) were occasionally identified, however those found in slurry were distinct (*tcr* gene cluster) and less diverse than those in soil. Furthermore, while soil-related metal resistances were predominantly linked to Mycobacteriaceae; in slurry, copper resistance genes were associated with Firmicutes, potentially order Lactobacillales. No evidence of ARG co-localisation with MRGs was found in contigs using the selected methods.

4.4 Discussion

4.4.1 Slurry

4.4.1.1 Beta-lactamase ARGs

The detection of *cfxA* beta-lactamase genes on contigs originating from several temporally dispersed slurry samples corroborates the unassembled read-based analyses presented previously (Chapter 2). Although *cfxA* genes have been reported in cattle faecal waste by qPCR (Muurinen et al., 2017, Feng et al., 2020) and metagenomic studies (Chambers et al., 2015, Zhou et al., 2016), few attempted to explicitly demonstrate which bacterial taxa in the cattle waste carry this gene group. Contrarily, a plethora of publications have associated *cfxA* genes with bacterial hosts isolated from other environments. In particular, the *cfxA* gene group has been identified in Gram negative anaerobes recovered from human oral (Iwahara et al., 2006, Binta and Patel, 2016) and intra-abdominal (García et al., 2008) infections. Members of order Bacteroidales, especially *Prevotella* sp. and *Bacteroides* sp. predominate in this context. The occurrence of *cfxA* in Flavobacteria such as *Capnocytophaga* sp. are also reported in relation to periodontitis (gum disease) (Handal et al., 2005). In addition, *cfxA* genes have been shown to be among the most prevalent ARGs in human gut and faecal samples by pan-continental metagenomic studies (Hu et al., 2013, Li et al., 2015a). The

development of host-specific *Bacteroides-Prevotella* genetic markers for tracing sources of faecal water pollution reiterates a consistent association between documented hosts of *cfxA* and the faecal microbiome of warm-blooded animals (Okabe et al., 2007).

To recapitulate the findings of the current work, contigs bearing *cfxA* genes were typically assigned to the order Bacteroidales (Kaiju) or phylum Bacteroidetes (CAT) and were therefore largely concordant with literature hitherto discussed. Although these low-resolution taxon assignments are robust, network analysis of unassembled reads offers compelling, if tentative, indications of genus-level associations which further concur with culture-orientated literature. For instance, putative *Prevotella* sp., *Bacteroides* sp. and *Capnoytophaga* sp. reads were correlated with >1 *cfxA*-like gene. Furthermore, isolate-based evidence exists for other genera linked to *cfxA* genes in network analyses, including *Porphyromonas* sp. (Binta and Patel, 2016) and *Alloprevotella* sp. (Arredondo et al., 2020). On the other hand, there were numerous occasions where putative associations between taxa and *cfxA* in the network could not be verified within literature; this may be due to the obscurity of the organisms concerned or simply read misassignment. The *cfxA* gene group represents an intriguing nexus between cephalosporin resistance in livestock and wild animals while also possessing clinical significance for humans. The present work has outlined evidence which supports current literature regarding likely hosts in intestinal and faecal microbiomes. Limited research is available on the extent to which *cfxA* circulates in animal populations; however, understanding this may have minimal impact on managing beta-lactamase resistance in humans given this gene group is already ostensibly endemic in humans. However, Hu et al. (2013) remarked that while *cfxA* genes exhibited high prevalence in human gut samples across all nationalities analysed, copy number was highly variable between individuals. Furthermore, Duan et al. (2020) found the enrichment of *cfxA* in the gut of humans was associated with antibiotic treatment. Likewise, Chambers et al. (2015) found *cfxA* recovery was enhanced in faeces from cattle receiving cephalosporin

antibiotics, suggesting proliferation through selection. However, in a subsequent publication no significant effect on *cfxA* was recorded (Feng et al., 2020).

4.4.1.2 MLS ARGs

The *lnuC* gene group was the most frequently detected ARG in slurry contigs, and further validates the slurry biomarker list compiled based on unassembled read analyses. Macrolide and lincosamide antibiotics both target protein synthesis by binding to the 50S ribosomal subunit and are primarily geared towards treating Gram positive bacteria (Leclercq, 2002, Tenson et al., 2003). However, the *lnu* gene family confers resistance specifically to lincosamides (Leclercq, 2002). Metagenomic studies have previously identified *lnuC* in cattle faeces, though links to taxa were not explored (Feng et al., 2019, Zaheer et al., 2019).

According to network analyses *lnuC* was likely to be possessed by a broad range of potential hosts, albeit typically members of phylum Bacteroidetes. In contrast, contig-based analyses suggested taxa largely belonging to Firmicutes were the most likely hosts of this ARG-group. Limited overall agreement was therefore found between unassembled and assembled read analyses with regard to *lnuC*. Nonetheless, contig analyses produced results consistent with the literature. Firstly, two contigs harbouring *lnuC* were independently classified as *Streptococcus uberis* by both CAT and Kaiju, while a further five were classified as such by Kaiju alone. In correspondence, several isolate surveys have demonstrated the carriage of *lnu*-like genes by *Streptococcus* sp., including *S. uberis* (Achard et al., 2005, Petinaki et al., 2008, Haenni et al., 2010, Gravey et al., 2013, Zhao et al., 2014). It is also of note that *Streptococcus uberis* is well-documented in the aetiology of bovine mastitis (Leigh, 1999, Käppeli et al., 2019). Considering the licensed use of lincosamides in veterinary practice to treat mastitis (Zhao et al., 2014), it is perhaps not surprising to uncover Streptococcal contigs harbouring lincosamide ARGs However, according to available farm usage records (Table 2.1), lincosamides were not routinely administered to adult cattle from 2015-17. Regardless, at least two studies

have found negligible evidence to support the notion that lincosamide use significantly increases the occurrence of *lnuC* in the faeces of treated cattle (Feng et al., 2019, Zaheer et al., 2019). This may indicate the stable maintenance of these genes independent of anthropogenic selective pressure. To this end it is quite possible that lincosamide-producing organisms, which include Streptomyces spp. (Spížek and Řezanka, 2017) maintain these ARGs in the slurry environment. Other host contig classifications with support in the literature included *Clostridium* sp. (Saldanha et al., 2020) and *Lactobacill salvarius* (Lee et al., 2017). However, cross-validation of these assignments was not achieved and should be treated with due caution.

Phenotypically, *mefA* could be described as an equal opposite to *lnuC*, as the former confers macrolide resistance while maintaining susceptibility to lincosamides (Leclercq, 2002). In agreement with unassembled read analyses, the retrieval of *mefA* from slurry contigs suggests they are core constituents of the slurry resistome. Moreover, the prominence of *mefA* in livestock waste is amply evidenced in literature (Agga et al., 2015, Li et al., 2015a, Noyes et al., 2016, Muurinen et al., 2017, Gou et al., 2018). It should also be noted that the dairy farm unit surveyed in the present work recorded occasional use of the macrolide antibiotic tulathromycin (Draxxin); it was administered to calves but not members of the adult herd.

As already discussed in relation to other ARGs abundant in cattle slurry, few publications have examined the principal slurry-borne taxa carrying *mefA*. However, Zhang et al. (2016) used correlation network analyses to infer 11 putative hosts for *mefA* in food and effluent-derived anaerobic digestate. Five of the genera identified by Zhang et al. (2016) (*Petrimonas* sp., *Acholeplasma* sp., *Tissierella* sp., *Parabacteroides* sp. and *Sporanaerobacter* sp.) were also correlated with *mefA* in the read-based network constructed for the present work. On the other hand, a study performed by Tong et al. (2019) determined archaeal taxa were potentially key sources of several ARGs in anaerobically digested wastewater, including *mefA*. However, despite the recovery of archaeal MAGs (phyla Thermoplasmatota and Halobacterota) no archaeal slurry contigs

could be associated with ARGs in the current study. It is possible correlation analyses could reveal archaeal-ARG links, though this was not explored. Accordingly, Tong et al. (2019) suggests the archaeal resistome of anaerobic digestate represents an understudied area of research.

Thorough validation of specific taxa carrying *mefA*-like genes was not achieved as these contigs were only occasionally classified by both methods. Regardless, three contigs were conjunctively assigned to the order Clostridiales by Kaiju and CAT. Furthermore, Kaiju identified seven additional contigs as either *Clostridium* sp. or Lachnospiraceae sp. Interestingly, there appears to be little evidence to support an association between *Clostridium* and *mefA*. Even so, a clinical study found *Clostridium difficile* (re-named *Clostridioides*) isolates harbouring *mefA* within a putative MGE (Isidro et al., 2018). Meanwhile, Kaiju indicated a single contig belonged to *Streptococcus*, for which there is substantial evidence of *mefA* carriage (Clancy et al., 1996, Arpin et al., 1999, Ardanuy et al., 2005).

ARGs enabling phenotypic resistance to both macrolides and lincosamides were comparatively rare; *erm*-like genes were identified on 33 slurry contigs. The most abundant *erm* gene, *ermB* (n = 10) was typically encountered on contigs classified as the order Lactobacillales by Kaiju (CAT verification absent). A gene marker for a multi-drug resistance plasmid (pRE25) first described in *Enterococcus feacium* RE25, was co-present on all contigs containing *ermB*. This may lend further support to the placement of these contigs within Lactobacillales, or extends candidates to those known to have acquired this plasmid.

4.4.1.3 Aminoglycoside ARGs

As with the other dominant slurry ARGs previously discussed, the *ant(6)*-gene family has well-documented distribution within livestock waste, encompassing swine, cattle and poultry (Muurinen et al., 2017, Zaheer et al., 2019, Lim et al., 2020).

In the case of the study by Lim et al. (2020), *ant(6)* genes were putatively associated with *Clostridoides* sp. and *Camplylobacter* sp. In the current work, contigs displaying *ant(6)*-like genes were poorly characterised by CAT and Kaiju, although the latter indicated >20% of contigs belonged to family Erysipelotrichaceae. This family is known to occur in cattle rumen (De Menezes et al., 2011) and is also of increasing interest in human gut physiology (Kaakoush, 2015). Interestingly, only three taxa were correlated with *ant(6)*-like genes (*aade*) and of those a single genus, *Iliebacterium*, is a representative Erysipelotrichaceae. On the other hand, there appear to be no clear links between Erysipelotrichaceae and *ant(6)* genes in existing literature. Consequently, this association remains conjecture. Meanwhile, two contigs were classified as *Camplylobacter* sp*.,* which has been demonstrated to carry *ant(6)* (Hormeño et al., 2018), however these identities require similar verification.

Another aminoglycoside gene family *ant(3)*, was frequently tied to Gamma-proteobacterial contigs. Focussing on taxon for which there was genus level agreement across Kaiju and CAT (n = 4), *Acinetobacter* sp. were host organisms for these resistance genes. *Acinetobacter* sp. (namely *A. baumannii*) are of well known clinical significance in human (Van Looveren et al., 2004) and veterinary medicine (van der Kolk et al., 2019), although their pathogenicity potential is difficult to determine above species level. Concomitant with the literature, *ant(3)*-like genes are widely dispersed among *Acinetobacter* sp. (Nie et al., 2014, Zhang et al., 2017a). Indeed, Zhang et al. (2017a) suggests select *ant(3)* genes facilitate intrinsic resistance to aminoglycosides in *Acinetobacter* sp. and describes the dissemination of these genes among the genus through multiple recombination events. Of further note, *Acinetobacter* sp. have been previously implicated in the persistence of ARGs in the soil environment following manure application (Leclercq et al., 2016).

4.4.2 Soil

Contig-based analysis suggests the core soil resistome is chiefly comprised of glycopeptide (*vanRO*) and rifamycin (*rbpA*) resistance genes; in broad agreement with unassembled read analyses conducted previously. Furthermore, the dominance of these ARGs in soil with no record of livestock waste application indicates their presence in soil is intrinsic. Indeed, a recent study by Li et al. (2020) highlighted glycopeptide and rifamycin resistance genes predominate in soils collected from a 'pristine' Tibetan plateau. In addition, *vanO*-type regulators (*vanRO-vanSO*) were the principal vancomycin resistance genes in the Tibetan survey. Likewise, Zaheer et al. (2019) established that *vanO*-type regulators were dominant in soil with no history of manure application, as well as soils subject to short- and long-term fertilisation with cattle manure. Autochthonous soil microbes, especially Actinomycetes, encompass many glycopeptide and rifamycin producing organisms (Mahajan and Balachandran, 2012, Saxena et al., 2014, Chandra and Kumar, 2017); it is therefore unsurprising that corresponding resistance determinants are also abundant. Nonetheless, these antibiotics are of significant clinical importance and surveillance of environmental hosts carrying related ARGs should not be dismissed.

Interestingly, the rifamycin resistance determinant *rbpA* is thought to be uniquely active in Actinobacteria (Dey et al., 2012) and potentially restricted to Actinomycetes (Paget et al., 2001). Consequently, it is encouraging that the majority of *rbpA*-positive contigs were assigned to Actinobacterial taxa (n = 97/99). Moreover, this gene was firmly embedded within an Actinobacterial cluster in network analyses. Although *rbpA* was first reported to facilitate rifampicin resistance in *Streptomyces coelicolor* (Newell et al., 2006), this gene has also been implicated in mycobacterial rifampicin-resistant phenotypes (Verma and Chatterji, 2014). In the current work, *rbpA* was almost exclusively associated with mycobacterial contigs. Based on contig analyses it appears

that while there is an abundance of rifamycin ARGs in soil, they exhibit limited mobility and are accordingly constrained to a set of specific taxa.

CAT failed to classify the majority of contigs presenting *vanO*-type genes. However, Kaiju consistently returned *Streptomyces* sp. (34.24%) alongside assorted Actinomycetes. As with *rbpa*, the inducible *vanO* operon was first described in an environmental Actinomycete; *Rhodococcus equis* (Gudeta et al., 2014). To this end, two contigs possessing *vanO*-like genes were identified as putative *Rhodococcus* sp. by Kaiju. Gudeta et al. (2014) provided evidence that *vanO* genes were exclusively chromosomal and evolutionarily distinct from Enterococcal operons conferring vancomycin resistance. This could represent another parallel with *rbpA*, whereby *vanO* may be similarly tied to actinomycetes, however this currently lacks additional support in literature.

The *mtrA*-like genes abundant in soil purportedly relate to RND (resistance nodulation division) efflux pumps according to databases, although the provenance of this is unclear in the literature. Nonetheless, *mtrA* is thought to be involved in maintaining multidrug resistances intrinsic to *Mycobacterium tuberculosis* (Nguyen et al., 2010). On the other hand, Cervantes et al. (2020) was unable to establish any association between *mtrA* and phenotypic drug resistance in *M. tuberculosis* isolates, indicating some uncertainty about the role of *mtrA* as a *bone fide* resistance determinant. However, Piddock (2006) underscored the potentially multifarious functions of efflux pump genes. Indeed, others have demonstrated *mtrA* expression plays a pivotal role in the persistence of *M. bovis* upon entering host macrophages (Via et al., 1996, Zahrt and Deretic, 2000, Chatterjee et al., 2018). Furthermore, *mtrA* orthologues in *Streptomyces* sp. have been associated with regulation of antibiotic production pathways (Som et al., 2017, Zhu et al., 2020). In any case, *mtrA* was strictly linked to actinobacterial contigs with Mycobacteriaceae strongly represented and showing consistent agreement with existing research.

Lastly, *dfrB* is a resistance gene family encoding resistance to trimethoprim, an antibiotic commonly administered in combination with sulfamethoxazole to treat broad spectrum urinary infections in humans (WHO, 2017a) and animals (Checcucci et al., 2020). However, the synergistic effect of these drugs is being undermined by the emergence of clinical resistance to this drug combination (Eliopoulos and Huovinen, 2001). Moreover, *dfrB*-family genes are habitually embedded within MGEs, thus aiding their potential for environmental dispersal (Alonso and Gready, 2006, Sánchez-Osuna et al., 2020, Toulouse et al., 2020). Interestingly, while environmental transfer of *dfrA* genes is well documented (Zhang et al., 2009, Berglund, 2015), this is not the case for *dfrB* genes, which are genetically distinct from *dfrA*-like genes. Li et al. (2015a) also showed *dfrA* was especially abundant in human and animal waste metagenomes and less abundant in soil, which suggests they may be a marker of faecal contamination. Returning to *dfrB*, members of this resistance gene family have been identified in Enterobacteriaceae including *Escherichia coli* (Toulouse et al., 2017) and *Salmonella enterica* (Levings et al., 2006). Similarly, Ateba et al. (2020) reported *dfrB* genes in assorted coliforms recovered from water treatment systems. Occurrence in Alcaligenaceae and Aeromonadaceae has also been documented (Roberts et al., 2012). In the present study, contigs containing *dfrB*-like genes were overwhelmingly classified as *Pseudomonas putida* by Kaiju (77.27%), followed by *Methylibium* sp. Five contigs originating from slurry-treated and untreated site samples were also identified as *Methylibium* sp. by CAT, however most *dfrB*-containing contigs remained unclassified by this program. *Methylibium* sp. are Beta-proteobacterial methylotropic organisms first described for their ability to degrade methyl tert-butyl ether (MTBE), an environmentally recalcitrant gasoline-additive (Nakatsu et al., 2006). There appears to be no evidence of trimethoprim-resistance in *Methylibium* sp. in previous studies, although it is not considered pathogenic and consequently received little attention in this regard. In contrast, *Pseudomonas putida* strain T-X16B (NCBI genbank accession: KM382182.1) is known to carry a class one integron encoding *dfrB4* in addition to the chloramphenicol resistance gene *catB3*. However, no putative *P. putida* contigs were found to possess *catB*-like genes.

Furthermore, ARG-ARG network analysis failed to show significant correlations between *dfrB* and *catB*-like genes in any site (Chapter 2 Figure 2.13). It is possible that these 'missing pieces' of evidence were lost during the assembly process, although given these associations were not identified based on unassembled reads, insufficient read-depth may be an equally plausible explanation. Alternatively, determining the presently unknown origin of *dfrB* genes (Alonso and Gready, 2006, Toulouse et al., 2017) could resolve these findings more clearly.

In summary, the core soil resistome as described by metagenomic analyses is characterised by glycopeptide and vancomycin resistance genes with strong evolutionary links to actinomycetes. Furthermore, this study has shown long-term fertilisation with dairy cattle manure is unlikely to dramatically alter the core resistome, or associated hosts. On the other hand, transient shifts are detectable in unassembled reads, however these are likely lost in assembled data due to insufficient sequencing depth.

4.4.3 MetaCompare Risk Scores

Quantifying the relative risk to human health posed by the environmental resistome remains an established priority within AMR research and related policy development (Ashbolt et al., 2013, Bengtsson-Palme and Larsson, 2015, Martínez et al., 2015, Larsson et al., 2018, Ben et al., 2019). In particular, Martínez et al. (2015) expressed the need to rank relative risk potential across different environments. However, it appears MetaCompare remains the only software to date which offers a standardised approach to calculating 'resistome risk scores' from assembled metagenomic data (Slizovskiy et al., 2020). Here, MetaCompare will be discussed within the specific context of the metagenomic data collected for the current study, meanwhile wider considerations/implications for the continued improvement of 'risk scoring' systems will be discussed further in the following chapter.

When comparing data for the same period, MetaCompare indicated the resistome risk potential of soils were statistically similar whether samples originated from the slurry-impacted field or the untreated site. Furthermore, samples collected within 24 hours of slurry application in May 2017 exhibit no clear elevation in risk potential compared to samples collected from the same site five days previously. Considering soil was collected within 24 hours of slurry application in May 2017, it is perhaps surprising that no significant difference was detected. In contrast, Chen et al. (2019a) reported soil risk scores (as calculated by MetaCompare) were significantly increased one day after exposure to cattle manure in a replicated microcosm study. The same study also showed that risk scores declined to control levels within 120 days. There are several possible explanations for why the risk scores calculated here do not correspond with the study performed by Chen et al. (2019a); these encompass both biological and technical reasons.

Firstly, the present study was conducted at the field-scale in an uncontrolled environment which concomitantly introduced a range of factors typically absent in a controlled laboratory-scale experiment (e.g. dilution effects); a point acknowledged by Chen et al. (2019a). Secondly, the cattle waste collected by Chen and colleagues was obtained directly from cattle and underwent neither solid-liquid separation or slow anaerobic digestion as occurs in a slurry tank. These processes change the physiochemical properties of the waste material while also impacting the microbial community (Pandey et al., 2018) and the resistome (Wallace et al., 2018, Tong et al., 2019). Furthermore, the slurry in the present work was broadcast, and not incorporated as in the study by Chen et al. (2019a); the latter practice has been shown to positively influence FIO survival (Hutchison et al., 2004, Hodgson et al., 2016). Finally, the average risk score associated with slurry in the present study (21.61 ± 0.06, n = 10) was more comparable to that assigned to wastewater treatment plant effluent (20.92 ± 1.3, n = 3) than to dairy lagoons (25.52 ± 1.1, n = 5) by Oh et al. (2018). This could

imply that slurry from the study dairy farm confers a genuinely reduced risk to human health relative to similar or equivalent waste products generated on other farms.

On the other hand, the risk scores may have been confounded by technical decisions and limitations inherent to the current configuration of MetaCompare. For instance, MetaCompare only normalises the occurrence of 'hazard' features against the total number of contigs and does not incorporate a normalisation step to account for individual contig length. Although this is acknowledged by Oh et al. (2018) in their paper introducing MetaCompare, this could have substantial implications for its use in certain scenarios.

To this end, consider two contigs, one 500bp in length and another 2000bp in length. In this case, the probability of locating any individual or combination of 'hazard' features is likely to be greater on the longer contig, despite the fact contig length is an artefact of assembly success and unrelated to biological risk potential. The direct comparison of slurry and soil samples in the current study exemplifies this issue. Specifically, although the original read libraries of both slurry and soil samples contain a similar number of reads (~42M and ~40M reads, respectively), Megahit was typically able to assemble fewer, larger contigs (~1.1M) for slurry than for soil (~1.7M). Indeed, on average each slurry sample contained 18 contigs with at least one ARG, MGE and pathogen marker, while soil samples had <3 on average. Although slurry would be expected to contain more contigs with all three of these 'hazard' features (as evidenced by unassembled read data), it is not clear to what degree their co-localisation on contigs is due to a greater contig length.

The developers of MetaCompare provide options to adjust parameters according to assembly quality; however, integrated contig-length normalisation would potentially negate the need for this, maintain standardisation and enhance accessibility. Alternatively, if one were to filter short contigs (e.g. <1kb) the normalising factor would

be manipulated. Doing this may cause further distortion since the distribution of 'hazard' features appearing separately is more likely to be skewed towards shorter contigs and so these features would be disproportionately lost in lower quality assemblies. Slizovskiy et al. (2020) also underscores the absence of a normalisation factor for gene length can lead to longer ARGs or MGEs being artificially penalised by a range of ARG-screening software. Accordingly, one might expect the risk scores of soil to be significantly smaller relative to slurry scores. However, the opposite is the case, with soil consistently assigned a greater risk score than slurry. This can be partially explained by the sheer number of soil contigs, which carry a single hazard feature type. Moreover, no additional weighting appears to be applied to contigs which possess all three 'hazard' features; therefore even though MetaCompare identifies that slurry samples have many more 'high risk' contigs than soil, this information is not incorporated into the final risk score. In addition, even soil subject to limited anthropogenic impact has been shown to be replete with generic efflux pumps (Van Goethem et al., 2018) which are not necessarily meaningful for ascribing risk potential (Martínez et al., 2015). While Oh et al. (2018) mention the removal of such genes could improve MetaCompare risk scores, their inclusion in the current version could explain the unintuitive results reported here.

Nonetheless, MetaCompare attempts to address a key knowledge gap that few have sought to tackle in practice and attests to the difficulty of doing so. Also of note, the intermediary figures calculated by MetaCompare correspond well with the relative number of contigs containing MGEs and ARGs as detected by ABRicate. The additional function of identifying the frequency with which these features occur on contigs carrying pathogen markers is also of interest. For example, according to MetaCompare these data show nearly all contigs carrying ARGs and MGEs are also likely to originate from pathogens.

Finally, aside from Chen et al. (2019a) and Slizovskiy et al. (2020) (although the latter did not make use of the risk score function), no other publications to date have used MetaCompare; further use will no doubt improve future versions and inspire entirely new programs.

4.4.4 The Mobilome

Many publications investigating the composition and perturbation of assorted environmental resistomes consider both ARGs and MGEs. However, studies concerning land-application of animal waste products often rely on correlation-based analyses to infer associations between these elements (Cleary et al., 2016, Muurinen et al., 2017, Zhang et al., 2017b, Chen et al., 2019b, Zhao et al., 2019, Wang et al., 2020a); this can generally be ascribed to the use of qPCR-based techniques which quantify target genes rather than their genetic context. In comparison, MGE-borne ARGs are more frequently confirmed and their genetic structures established in wastewater treatment environments (Che et al., 2019, Ju et al., 2019, Yin et al., 2019, Zhao et al., 2020). Nonetheless, select studies have assessed the structure of ARG-bearing MGEs in animal waste (Ma et al., 2016) and surrounding farm environments (Mencía-Ares et al., 2020). Moreover, Yang et al. (2020) identified swine slurry as a hotspot for ARGs and MGEs. Together with existing network analyses these provide a summary of MGE candidates likely to be involved in ARG mobilisation within waste-amended soils.

In the current work, contig-based analyses characterised the co-localisation of specific slurry-associated ARGs and MGEs. For example, three slurry contigs contained the complete recombination module of the Tn*916* conjugative transposon adjacent *tetM*, a core member of the slurry resistome (Figure 4.4).

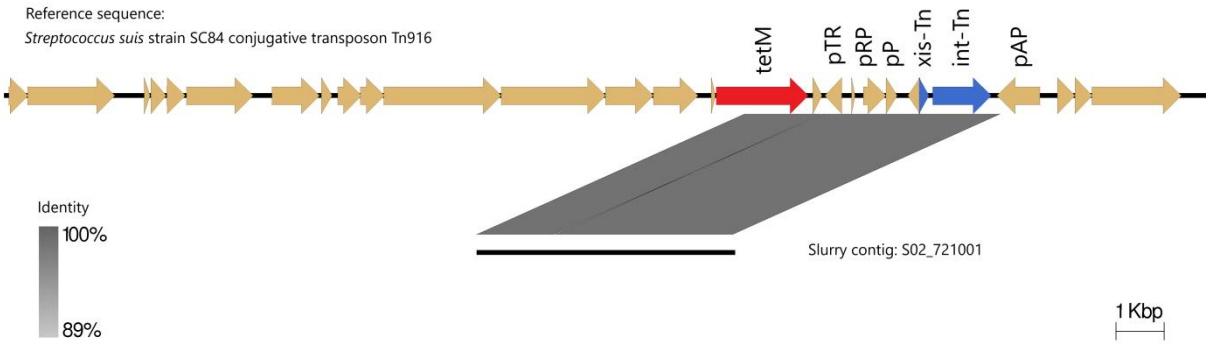


Figure 4.4 Schematic of *tetM* embedded within Tn*916* transposon. pTR = putative transcriptional regulator, pRP = putative regulatory protein, pP = putative polyribonucleotide, pAP = putative abi-alpha protein, xis-Tn = excisionase, int-Tn = integrase. Example contig from the current study is shown as a black line. Alignment region shown in grey. Gradient reflects percentage identity. Colours indicate function: resistance genes (red), mobilisation (blue) and other (beige).

Tn*916*-like elements typically encode *tetM* and are known to circulate within a broad range of hosts, including Gram positive and Gram negative bacteria (Roberts, 2005, Hegstad et al., 2010, Ciric et al., 2013). Potential hosts also encompass pathogens of clinical importance to human and veterinary medicine (Roberts and Mullany, 2011, Fischer et al., 2013, Pinto et al., 2014). Furthermore, Tn*916*-like elements can carry a diverse repertoire of other resistance determinants, including various combinations of macrolide (*erm*, *mef*), tetracycline (*tet*), aminoglycoside (*aphA*) and occasionally mercury (*mer*) resistance genes (Roberts and Mullany, 2011, Ciric et al., 2013, Pinto et al., 2014). Moreover, Pinto et al. (2014) demonstrated *in vitro* Tn*916*-mediated transfer of macrolide and tetracycline resistance genes from a bovine strain of *Streptococcus agalactiae* to a susceptible recipient of human origin. In the present study Tn*916*-like elements were only co-localised with one other ARG; *tetO* (n = 4). However, this association was limited to a Tn*916*-family transposase and a complete Tn*916*-like recombination module was not found within these assembled contigs. On the other hand, Lu et al. (2020) found the relative abundance of Tn*916* in livestock slurry was positively correlated with ARGs belonging to various antibiotic classes (tetracycline, macrolide, aminoglycoside and sulfonamide ARGs).

Interestingly, while both laboratory and plot-scale experiments have reported the persistence of *Tn916* in animal waste-amended soil (Natarajan and Oriel, 1992, Andrews Jr et al., 2004, Lu et al., 2020), a similar trend could not be established based on the contigs generated in the present study; it is possible deeper sequencing efforts could indicate otherwise. The Class A beta-lactamase gene *cfxA* was also occasionally co-localised with MGE markers on contigs (n = 3). Specifically, *cfxA* genes were found alongside an NBU1-like region (Figure 4.5). Indeed, NBUs (non-replicating *Bacteroides* units) are well-characterised integrative elements and have been shown to carry *cfxA*-like genes (Li et al., 1993). However, NBUs are not self-transmissible and require the presence of a suitable conjugative transposon (CTnDOT-family) to facilitate further dispersal (Shoemaker et al., 1996, Rajeev et al., 2006). These transposons were not identified among slurry contigs in this study, however lack of detection does not prove their absence. Furthermore, CTnDOT transposons carry *ermF* and *tetQ* (Waters and Salyers, 2013); genes which are at least individually present on slurry contigs.

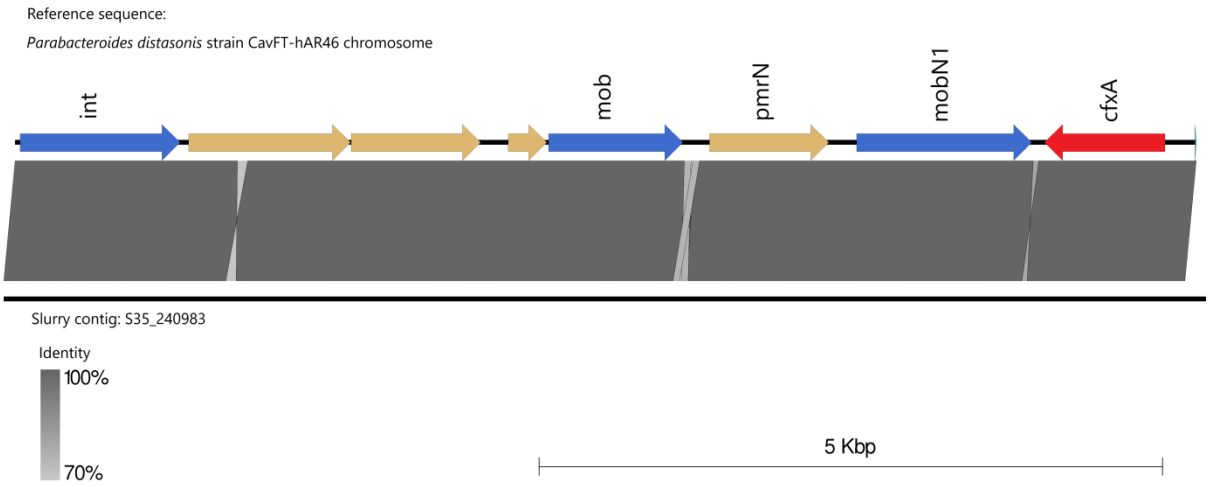


Figure 4.5 Schematic of *cfxA* and NBU-1-like region. int = integrase, mob = mobilisation protein, pmrN = putative primase, mobN1 = mobilisation protein. Example contig from the current study is shown as a black line. Alignment is shown in grey. Gradient reflects percentage identity. Colours correspond to function: resistance genes (red), mobilisation (blue) and other (beige).

To this end, mobilisation of NBUs could explain why other studies have shown *cfxA* is prevalent within livestock waste (Zhou et al., 2016) and human faeces (Li et al., 2015a). Of additional note, the conjugal transfer of CTnDOT-like transposons can be enhanced by exposure to low levels of tetracycline (Waters and Salyers, 2013). Speculatively, the prominence of *cfxA* in slurry may therefore relate to tetracycline use and exposure to sub-inhibitory residues. Consequently, horizontal gene transfer between *Bacteroides* sp. represents a plausible route for the propagation of *cfxA* genes within slurry.

A subset of ten contigs contained the *ermB* MLS resistance gene in close proximity to a Tn*3* family transposase and recombinase (Figure 4.6). More specifically, the configuration resembles (identity = 99.96%, coverage = 90%) Tn*917* first described in *Enterococcus faecalis* by Shaw and Clewell (1985). This is also consistent with the putative assignment of order Lactobacillales by Kaiju. However, although Enterococci are documented nosocomial human pathogens, emergent vancomycin rather than erythromycin resistance is of greater clinical concern (Werner et al., 2020).

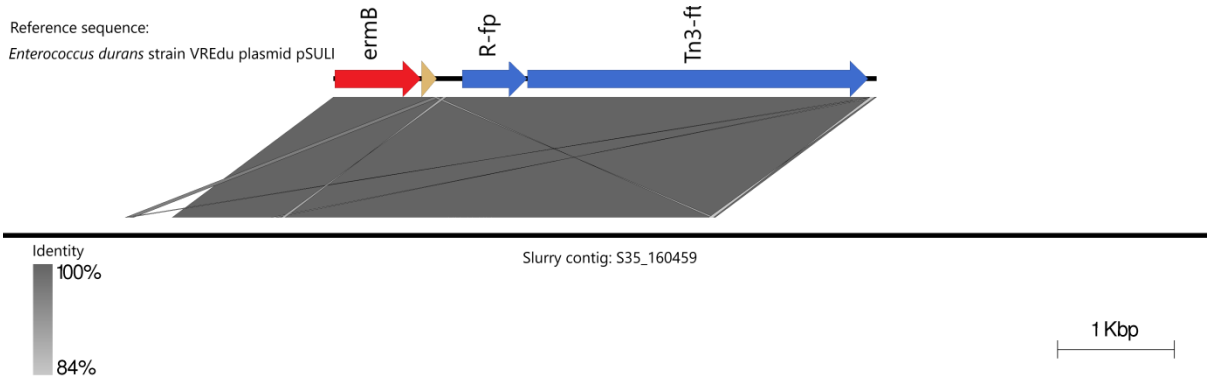


Figure 4.6 Schematic of *ermB* and Tn3-like region. R-fp = recombinase family protein, Tn3-ft = Tn3 family transposase. Example contig from the current study is shown as a black line. Alignment is shown in grey. Gradient reflects percentage identity. Colours correspond to function: resistance genes (red), mobilisation (blue) and other (beige).

Previous research has established animal faecal material is often associated with aminoglycoside and sulfonamide resistance genes (Nõlvak et al., 2016, Noyes et al., 2016, Muurinen et al., 2017, Checcucci et al., 2020). In the current work, 15 contigs contained *aph(3")* and *aph(6)* streptomycin resistance genes, while an additional contig

included these in combination with *sul2* and an MGE marker gene (Figure 4.7). The *aph(3")* and *aph(6)* genes (synomymous with *strA-strB*), commonly co-occur with *sul2* on plasmids recovered from isolates of human, animal and environmental origin (Sundin and Bender, 1996, Anantham and Hall, 2012). Furthermore, the *sul2-strA-strB* motif has recently been reported in previously undisturbed Antarctic ice cores, suggesting its emergence pre-dates antibiotic use (Okubo et al., 2019). Despite the apparent global distribution of *strA-strB* motifs, other studies have indicated their environmental prominence responds to anthropogenic activity. In particular, Ludvigsen et al. (2018) showed the incidence of *strA-strB* was higher in the gut of honeybees in the US relative to Norway where agricultural use of streptomycin is comparatively limited. Lastly, Zhao et al. (2020) also highlighted a variety of *strA-strB* configurations in antibiotic production wastewater. These also included different accessory genes such as *bla* and *sul* ARGs. In the current work, a single contig encoding *strA-strB* was also co-localised with *tetY*, although no MGE markers were additionally detected.

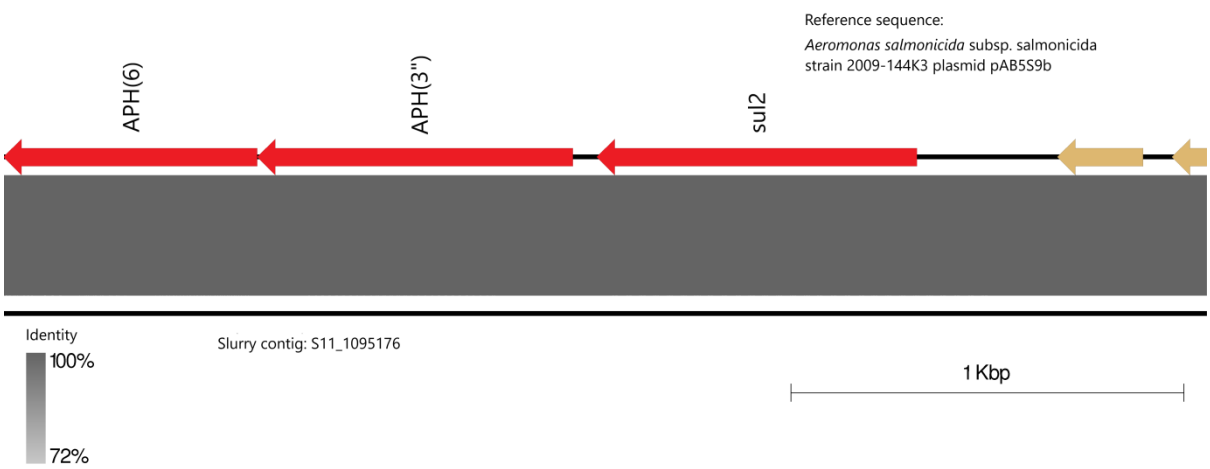


Figure 4.7 Schematic of *aph(6), aph(3")* and *sul2* containing region. Example contig from the current study is shown as a black line. Alignment is shown in grey. Gradient reflects percentage identity. Colours correspond to function: resistance genes (red), mobilisation (blue) and other (beige).

A different aminoglycoside resistance gene cluster was identified on another contig in the form of *ant(6)-1a* (synonymous with *aadE*) and an *ant(9)*-subgroup gene (Figure 4.8). Analysis with NCBI-BLAST showed the contig concerned had 100% query coverage and identity with several conjugative transposons and chromosomal sequences belonging to

various members of Firmicutes. Furthermore, network analysis on unassembled read data also alluded to an association between the aforementioned genes (see Chapter 2, Figure 2.13A).

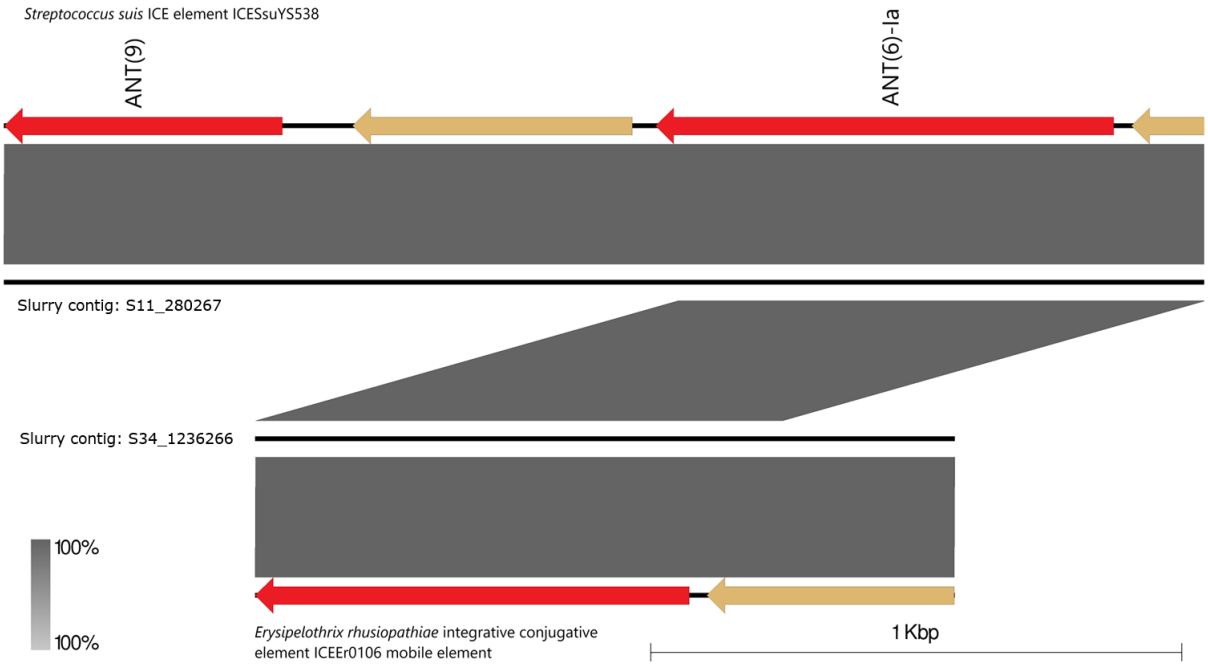


Figure 4.8 Schematic of *ant(9), ant(6)-1a* containing region. Example contig from the current study is shown as a black line. Alignment is shown in grey. Gradient reflects percentage identity. Colours correspond to function: resistance genes (red), mobilisation (blue) and other (beige).

Although limited discussion about this specific ARG pairing is evident in publications, the dissemination of multiple co-resident aminoglycoside resistances genes does not appear uncommon, as already outlined with regard to *strA-strB*. Accordingly, van Overbeek et al. (2002) postulated multiple aminoglycoside ARGs may be maintained on MGEs due to host-specific expression of certain ARGs. In this way, ostensibly redundant genes could maximise the range of hosts to which an MGE confers a selective advantage.

In contrast to slurry, evidence of ARG-MGE co-occurrence could not be established in soil contigs (based on ABRicate screening), this was likely due to a combination of insufficient sequencing depth, reduced average contig length (assembly efficiency) and the dilution of slurry-associated MGE sequences in soil.

4.4.5 Tracking Slurry-borne MAGs

To the author's knowledge no previously published studies have traced MAGs in slurry amendments directly to slurry amended soil. The recovery of a slurry MAG with high average nucleotide identity (>99.5%) to a MAG derived from amended soil is indicative of transfer. Although the detection of a genetically similar MAG in slurry and slurry-treated soil may not seem especially significant, this belies its importance from the perspective of validation. In particular, both differential abundance and network analysis conducted on unassembled reads had already highlighted *Proteiniphilum* sp. as key indicators of slurry application. Mapping back the *Proteiniphilum* sp. MAG recovered from slurry to the read libraries of each soil sample further corroborates, albeit crudely, the transient increase in slurry-associated *Proteiniphillum* sp. within treated soil (e.g. 22.7 ± 1.2% increase in mapped reads <24hrs after slurry application in May 2017). Lastly, at least one other study has also found evidence of manure-mediated increases in the relative abundance of *Proteiniphilum* sp. within soil (Wolters et al., 2018). In addition, metagenomics was capable of detecting changes in these populations despite their relatively low abundance in the soil studied.

Meanwhile, *Proteiniphilum* sp. are obligately anaerobic (Hahnke et al., 2016) and therefore not likely to be inherently populous within well-drained soils. Consequently, their elevated detection in treated soil underscores a wider point that there are many potential markers of animal waste contamination which could be utilised when using metagenomics rather than classical culture techniques. Likewise, the present study showed traditional markers of faecal contamination, namely *E. coli*, proved comparatively poor indicators, with greater relative abundances often detected in soil devoid of systemic livestock waste amendment (discussed in Chapter 3). This likely occurred due to the presence of environmental *E. coli* with no association with slurry.

The ultimate purpose of recovering MAGs was to establish and validate the specific hosts of ARGs as indicated by network analysis of unassembled reads. Unfortunately, this was not possible as ARGs could not be detected within the MAGs generated (including those belonging to *Proteiniphilum* sp.) using the bioinformatic methods employed. As mentioned previously, this may correspond to assembly failing to resolve genomic islands and plasmids bearing ARGs (Maguire et al., 2020), insufficient sequencing depth or the genuine absence of these genes. However, contig-based analysis suggests the latter is not the case. In contrast, several other publications have successfully identified resistance determinants within MAGs (Kantor et al., 2019, Stamps and Spear, 2020, Tan et al., 2021).

4.5 Conclusion

In summary, the current chapter successfully contextualised various links between ARGs, putative associations with bacterial taxonomy and provided additional insight into the composition and structure of the cattle slurry mobilome. This was achieved and partially validated through using a combination of methods which interrogated the sequence data at different levels, namely the unassembled reads and assembled contigs. Interestingly, while the experimental data presented here clearly indicates bacterial community composition is a key determinant of the resistome, analysis of the mobilome underscored MGEs are integral to the dispersal of abundant slurry ARGs, such as *tetM* which were not correlated with specific taxa. Furthermore, this chapter indicates members of Bacteroidale*s* and Lactobacillales may be especially important for the propagation and maintenance of ARG-harbouring MGEs in slurry.

On the other hand, metagenome assembly was unable to provide information about the composition and genetic structure of ARG-harbouring MGEs in soil samples. In addition, the episodic and transient increase in slurry-associated ARGs detected within soil as evidenced by unassembled read data in previous chapters, could not be validated by

assembled contigs. Nonetheless, agreement between read-based and contig-based analysis could be demonstrated for at least one taxon enriched in soil following slurry application. Carrying out deeper sequencing to better characterise the spread of animal waste-borne ARGs in soil, their respective hosts and related MGE complements is vital to establishing appropriate risk management strategies. Complementary to this, the continued development of tractable risk scores which can be derived directly from metagenomic data should be an area of further research, especially as the software currently available was unable to handle the combination of different environmental compartments evaluated here.

Chapter 5
Final Discussion

5.1 Summary of Key Findings

The primary objective of the study was to characterise and track the response of the soil resistome following multiple applications of dairy cattle slurry in order to better understand the extent to which activities at a conventional (model) dairy farm might pose a risk to human health through the dissemination of antibiotic resistant bacteria. The source of slurry was also sampled to inform the distinction between autochthonous and introduced ARGs/bacteria. A site with no history of livestock waste application was also studied to provide an ecological baseline. Various meta-data were also collected to further assist the determination of key factors which might influence the soil resistome. Changes in bacterial communities were considered, as well as the relationship between specific taxa and antimicrobial resistance determinants. Critically, popular correlation analyses were complemented with assembly-based analyses to provide robust linkage between ARGs and their hosts.

In Chapter 2, the composition of soil and slurry resistomes were shown to be distinct from each other, irrespective of whether soil originated from sites with a history of exposure to dairy cattle slurry. Consequently, it became apparent that the long-term application of slurry had not dramatically altered the composition of dominant ARGs in soil. Nonetheless, a core member of the slurry resistome, *tetM* was significantly more abundant in slurry-impacted soil relative to the untreated site. Literature suggests this tetracycline resistance gene was most likely circulating on transposons such as Tn*916* and Tn*1545*. Subsequent examination of *tetM*-containing contigs confirmed Tn*916* carriage (Chapter 4, Figure 4.4). Interestingly, while an increased abundance of *tetM*/Tn*916* has previously been associated with long- and short-term exposure to livestock waste, soil samples collected within 24 hours of slurry application did not

exhibit a significant increase in *tetM*. It was therefore proposed that repeated exposure events had led to the saturation of soil carrying capacity for these genes at this site. On the other hand, a number of ARGs identified as slurry biomarkers were transiently enriched within soil after slurry application in May 2017 and were otherwise present at similar relative abundances in the undisturbed site (Chapter 2, Figure 2.11). This suggested that few slurry-associated ARGs possessed the capacity for naturalisation. In contrast, an earlier application (late winter) of slurry in 2018 appeared to increase the detection period of some slurry-borne ARGs by at least two months (Chapter, Figure 2.11).

Likewise, the response of the soil bacterial community to slurry-amendment described in Chapter 3 highlighted very few slurry indicator taxa were significantly enriched in soil immediately after application events. Furthermore, these enriched organisms are not classically considered indicators of faecal contamination. Meanwhile, pH was shown to correlate with phylum-level differences between the treated and untreated sites, which may correspond to alkalising effects of long-term slurry application.

Network analyses were used to consider ARG-ARG and taxon-taxon correlations based on unassembled reads in Chapters 2 and 3. Subsequent contig-based analysis with ABRicate demonstrated that correlations between select aminoglycoside ARGs (*ant6* and *ant9*, *aph3* and *aph6*) were indeed due to their co-occurrence on slurry contigs. Furthermore, certain taxon-ARG correlations were also corroborated by contig-based annotations (Mycobacteriaceae and *rbpA*, *cfxA* and *Bacteroides* sp./Bacteroidetes). In contrast, there were also many putative associations which could not be validated by contig analyses. Although it is far from ideal for so many un-validated inferences to remain, this is unfortunately indicative of studying bacterial assemblages (particularly in soil) where such much biological context is yet to be properly understood. Lastly, it is important to remember that a considerable quanitity of data is lost during assembly; network analyses offer some means by which to harness (if imperfectly) unassembled short-read data for association mining.

Analysis of both unassembled and assembled data (Chapter 1 and 4, respectively) showed that aspects of the core slurry resistome were consistent with the literature concerning dairy cattle, but also pigs. This is of interest as it implies a core-gut resistome may exist even across species with very different digestive physiology (i.e. ruminant vs. monogastric) and antibiotic therapy requirements; mastitis in cattle and respiratory disease in pigs (De Briyne et al., 2014). Indeed, penicillins and tetracyclines have found widespread use across livestock industries (Dibner and Richards, 2005). In contrast, extensive research has also demonstrated clear differences between these animals with respect to the prevalence of specific ARG-subgroups (swine manure typically has a more diverse resistome).

5.2 Informing Policy: Best Practice and One Health

The following discussion focuses on framing the results of the present study within the context of existing UK government guidelines and policy regarding dairy farming and the threat of emergent AMR.

It would be impossible to discuss the UK's approach to tackling AMR without making reference to the global recommendations outlined by O'Neill (2016), which have clearly influenced successive policy decisions regarding this subject. One core tenet of the aforementioned report concerns the reduction of antibiotic use in agriculture, with particular focus on the highest priority - critically important antimicrobials (HP-CIAs), as defined by the WHO (2017a). Accordingly, UK-VARSS (2020) showed UK veterinary sales of HP-CIAs declined by ~74% between 2015-2019. Furthermore, a survey covering 34% UK dairy cattle revealed HP-CIA use reduced by 87% between 2017-2019 (UK-VARSS, 2020). Accordingly, the model farm used in the current study reflects this trend in policy (Chapter 2, Table 2.1). Specifically, the frequency and range of third and fourth generation cephalosporin use on the farm diminished greatly between 2015-2017, with total doses of HP-CIAs falling by 75% over three years. In addition, the survey

presented in the UK-VARSS (2020) report included farms with a mean herd size of 215, which is similar to that of the study dairy farm (~250) at the time of study. On the other hand, the overall UK mean is only 148, although there is a trend towards fewer, larger dairy farms (AHDB, 2020a). It has been shown that the short-term effects of Brexit may lead to a contraction in the herd size of some dairy operations due to the loss of key workers from the European Union (NFU, 2017). In the long-term, this may precipitate the acceleration of automation, subsequent expansion and further loss of smaller dairy farms unable to diversify or afford automation.

Consequently, the semi-automated dairy farm studied here gives an insight into the environmental impacts of slurry fertilisation under stringent antibiotic stewardship at a scale which is likely to remain relevant into the future. Although there are shortcomings to focussing on a single farm (e.g., unusual local factors may skew data interpretation) it is proposed that in this case the depth of temporal analysis at a commercial farm with open access to clearly defined operating procedures offered an unparalleled opportunity to assess *in situ* management practice with a level of detail scarcely found in existing literature. Furthermore, the noise inherent to real-world environmental data has been placed in check by ensuring it is viewed with an awareness of studies which have opted for a small-scale, controlled (artificial) and highly replicated approach to tackling a similar question.

Less than three months after the release of the final report from the Review on Antimicrobial Resistance (O'Neill, 2016), the UK government issued new guidance on the handling of manure and slurry to reduce antibiotic resistance (GOV.UK, 2016).

One proposed mitigation strategy discourages slurry application on grazing pasture or land which is cropped for feed production (i.e. silage). However, where this takes place it is recommended that fields are withdrawn from grazing or cropping for at least eight weeks (GOV.UK, 2016). The primary objective of this measure is to minimise the likelihood of livestock consuming feed contaminated with antibiotic resistant bacteria,

pathogens and possibly unmetabolised antibiotics or bioactive degradation products. Indeed, in the current study, the first application of slurry in spring 2017 was accompanied by a transient increase in the detection of slurry-associated ARGs, which declined to pre-treatment levels within eight weeks. Contrarily, select slurry-associated ARGs were elevated above the field baseline 12 weeks after the resumption of organic fertilisation the following year in February. This indicates slurry-borne ARGs could contaminate cut-grass at least four weeks after the advised eight week wait-period. Moreover, the current study also showed *tetM*-like reads were consistently more prevalent in soil subject to long-term slurry exposure compared to soil with no history of amendment. As a result, contamination of cut grass with *tetM*-carrying organisms may represent a persistent risk. Although soil rather than cut grass was sampled, the contamination of ensiling material with soil organisms is an acknowledged hazard (Drouin and Lafrenière, 2012, Queiroz et al., 2018). It is thought that the contamination of cut grass with soil organisms, including those introduced by animal waste amendments is more likely to occur when a lower cut-height is employed (Drouin and Lafrenière, 2012).

Studies evaluating silage contamination typically involve organisms, which might compromise nutritional quality or represent a direct risk to livestock and human health; including *Clostridium* spp, *Listeria* spp and *Bacillus* spp (Driehuis and Elferink, 2000, Drouin and Lafrenière, 2012, Driehuis et al., 2018, Queiroz et al., 2018). Meanwhile, the potential for silage to act as a reservoir for ARGs on farms has largely escaped attention. Although Wu et al. (2020) found tetracycline and macrolide resistance genes dominated sweet corn kernel silage and described the effects of various silage additives, it appears no work to date clarifies the extent to which contamination with soil and animal waste amendments may influence the resistome of ensiled material. The likelihood of antibiotic residues contaminating silage in the same fashion is equally unclear. While crop plants destined for human consumption have been shown to absorb small quantities of

antibiotics within 6 weeks of exposure, including chlortetracycline (Kumar et al., 2005a) and sulfamethazine (Dolliver et al., 2007), little information is available for silage.

Assuming grass cut after eight weeks of slurry application could be contaminated with slurry-borne ARGs, it might be supposed that the fermentation stage of silage production could limit the proliferation of bacteria carrying slurry ARGs. Secondly, unlike slurry, which is neutral to alkaline, silage is often acidic and this may represent an additional obstacle to the survival of these bacteria. Alternatively, *Lactobacillus* sp. form a major component of silage and select species are often used as inoculants to enhance fermentation and silage quality (Ellis et al., 2016, Selwet, 2020). Indeed, based on unassembled data slurry-borne *Lactobacillus* sp. were among those bacteria, which persisted at elevated levels within soil 12 weeks after slurry application. Furthermore, mobile elements with a broad-host range like *tetM* may not require the prolonged survival of their original host to proliferate. Interestingly, Enterococci are also commonly used as LAB (lactic acid bacteria) for inoculating silage and while current evidence suggests pathogenic Enterococci are distinct from probiotic strains, debate exists about their safe use in view of transferable ARGs like *tetM* (Hanchi et al., 2018, Santos et al., 2020).

While it remains to be determined whether contaminated cut grass poses a credible route for ARG dispersal, it could be argued that even if slurry-associated ARGs were to be ingested by cattle, this would not dramatically change their resistome and therefore constitute negligible risk. Another perspective is that it could present an additional opportunity for cattle ARGs to make their way into humans through handling of silage. Furthermore, any antibiotic residues in the silage would effectively act as a subtherapeutic dose with the potential to exert selection pressure at herd level. The most likely means by which contaminated silage might promote AMR is if it contains pathogens that go on to cause illness in livestock and necessitate antibiotic therapy. In any case, it is interesting to note best practice stipulates material for feed should not be cropped within eight weeks of slurry application on the basis that antibiotic resistant

bacteria might be present (GOV.UK, 2016), while the likelihood of silage becoming contaminated with ARB is still largely untested.

Where the direct quantification of risk remains elusive, as in the case of AMR transfer from the environment to humans, Manaia (2017) highlighted the virtues of the precautionary principle and stresses measures should be implemented pre-emptively to contain potential hazards. Given the widely endorsed concept of "One Health" (Davis et al., 2017, Collignon and McEwen, 2019, Tiedje et al., 2019), it therefore seems appropriate to extend the precautionary principle to livestock feed production. Although evidence of this can be seen in the UK guidelines previously discussed, it is interesting to note that following the application of animal waste to land, there is a disparity of 40 weeks between the recommended wait-period for harvesting silage for animal feed and when ready-to-eat crops should be collected for human consumption. It is possible the wait-period guidelines for animal feed are not as strict simply because the produce is not destined for direct human consumption. Alternatively, the wait-period may take into account the additional processing animal feed undergoes during ensiling.

The present work also highlighted variability in the rate at which both slurry-borne ARGs and bacteria decline after slurry application. As a result, it may be difficult to firmly establish fixed recommendations for when material for livestock or human consumption may be safely harvested. In the present work, it was proposed that the variability relates to prevailing environmental conditions at the time of slurry application, and is likely to reflect season. Since seasonal effects are challenging to replicate and delineate at the field scale, few studies have explored the relationship between ARGs, taxa and season. Nonetheless, studies have shown surface transport of ARGs can be enhanced during rainfall events (Joy et al., 2013, Huang et al., 2019), while others indicated increased dispersal via run-off during the growing season (Neher et al., 2020). In light of the consistent relationship between ARGs and taxa in the current work, it is also pertinent to note that the survival of faecal bacteria in soil (especially FIOs), is modulated by

environmental conditions such as temperature (Van Elsas et al., 2011, Park et al., 2016) and rewetting events (Zaleski et al., 2005) which are liable to change with the seasons.

However, seasonal conditions in the UK are far from uniform, and there is a danger that aberrant climate events can confound the benefits of policies defined by the calendar. Indeed, a government report raised similar concerns from farmers regarding the inflexibility of UK NVZ closed-period policy (DEFRA, 2013).

Alternatively, current government guidelines also suggest shallow injection of slurry can reduce the spread of antibiotic resistant bacteria to leafy crops whether destined for animal feed or human consumption (GOV.UK, 2016). Although shallow injection may limit wildlife exposure, minimise surface transport of ARGs, and even curtail methane emissions when compared to broadcast methods, Hodgson et al. (2016) demonstrated sub-surface injection increased FIO survival by several weeks. Shallow injection was not practiced at the farm studied in the current work and slurry applied to the soil surface was subject to UV exposure and the consequent effects of that.

UK best practice, which was likely informed by the work of Marti et al. (2014), recommends a wait-period of 12 months before spreading untreated manure or slurry on land where ready-to-eat produce is cultivated (GOV.UK, 2016). Despite the absence of a 12-month cessation in slurry application, the 14 week (3.5 months) NVZ closed period was sufficient for almost all detected slurry-associated ARGs to decline to levels commensurate with samples collected over the same period from a site with no history of slurry application. One of the few exceptions to this was *ugd*-like ARGs, which remained enriched relative to the untreated site for the duration of the survey.

In summary, the present study illustrated the eight week withdrawal period suggested by the UK government may not always be sufficient to prevent silage from becoming contaminated with slurry-borne ARGs and bacteria. Meanwhile, the current work also showed many ARGs can dissipate rapidly over the 14 week closed period. Ultimately, exploration of UK policy and wider scientific literature shows those seeking to devise

evidence-based policy for tackling AMR in agriculture have the unenviable task of delivering safe and tractable industry recommendations that accommodate the high degree of uncertainty remaining in this research area.

5.3 Challenges for Evaluating Resistome Risk

As already identified in Chapter 4, establishing the relative risk posed by the land application of dairy cattle slurry can be problematic. More specifically, developing a standardised measure for quantifying risk which remains meaningful when applied across environmental compartments is especially challenging. The principal hindrances to doing so will now be discussed, with reference to the current work and more general research.

5.3.1 Defining Pathogens in the Context of Resistome Risk

While it is widely accepted that the emergence of multidrug resistant human pathogens should be considered the primary hazard, there are many nuances to consider when calculating how the environment contributes to the risk of this occurring. For instance, ARGs are abundant in commensals residing in healthy humans and other animals (Poeta et al., 2006, Card et al., 2014, Li et al., 2014a) as well as environmental bacteria which are only rarely pathogenic to humans (Martinez, 2009). Environmental bacteria may therefore act as vectors for ARGs, especially if they are associated with MGEs, which may facilitate the transfer of these genes to *bona fide* human pathogens. Alternatively, commensals can become pathogenic under certain circumstances. For example, Proença et al. (2017) demonstrated a single transposon insertion event enabled macrophage evasion by commensal *E. coli*; thus promoting pathogen-like traits. Finally, the 'silent colonisation' of healthy humans with ARG-carrying environmental bacteria or

commensals may manifest pathogenically in scenarios where these individuals become immune-compromised (Manaia, 2017).

In the context of the present study, *Aeromonas* sp. represent a useful example to illustrate and further explore some of the aforementioned challenges. Aeromonads are generally described as autochthonous environmental bacteria, which are found in a diverse array of terrestrial and aquatic settings (Janda and Abbott, 2010, Batra et al., 2016). It is therefore unsurprising that *Aeromonas* sp. were detected in metagenomic samples collected from both soil sites and slurry. However, select species of *Aeromonas* are increasingly implicated in human infections including septicaemia and gastroenteritis (Merino et al., 1995, Janda and Abbott, 2010, Fernández-Bravo and Figueras, 2020). Moreover, the incidences of Aeromonad-related infections are not limited to immune-compromised individuals (Janda and Abbott, 2010, Batra et al., 2016). Despite this, they are not included within the PATRIC database of pathogenic organisms used by MetaCompare to calculate risk scores for the current work.

Other aspects worthy of consideration include sanitation infrastructure and changing climatic context, the importance of which is also exemplified by *Aeromonas* sp. In particular, infections with *Aeromomas* sp. are known to be associated with untreated water resources (Carvalho et al., 2012). Similarly, infections have been linked to natural disasters which cause severe flooding and damage to sanitation systems, such as hurricanes and tsunamis (Hiransuthikul et al., 2005, Presley et al., 2006). With specific reference to agriculture and the storage of slurry, increasing flood events pose a genuine risk to uncovered slurry lagoons which may become overwhelmed by rainfall or subsumed by uncontrolled floodwaters. This also raises larger questions about how global climate change may force researchers to reconsider which antibiotic resistant organisms in the environment represent a cause for concern in the future. In this regard, researchers should guard against the tendency to focus exclusively on the 'usual suspects' and be wary of database bias.

Aeromonads may also be relevant from a One Health perspective, as antibiotic resistant isolates have been recovered from the faeces of healthy sheep, cattle and horses (Ceylan et al., 2009), in addition to wild animals (Dias et al., 2018). The authors of the aforementioned studies also make specific reference to risk to human health.

In the current work, a slurry sample contig had high BLAST homology with a plasmid-borne region (pAB5S9b) associated with *Aeromonas* sp. (sequence length = 3kb, identity 99.97%, e-value = 0.0). This contig harboured a combination of *strA\B* and *sul2*, a widespread ARG combination previously discussed (Chapter 4, Figure 4.7). In addition, culture-based analysis of soil samples collected from the slurry-impacted site (data not shown because the thesis focus is on the metagenomic aspects of the study), resulted in the isolation of an Aeromonad exhibiting phenotypic resistance (via disc diffusion assays) to ceflexin, ceftiofur (first and third generation cephalosporin, respectively), ampicillin (penicillin), nalidixic acid (flouroquinolone), florfenicol (phenicol) and nitrofurantoin (nitrofuran). Although resistance to ampicillin and narrow spectrum cephalosporins are thought to be common among Aeromonads, the other phenotypes described here are believed to be less common in clinical isolates (Janda and Abbott, 2010).

Defining what constitutes an organism of concern is therefore complex and will likely continue to frustrate attempts to quantify the risk AMR in the environment poses to human and animal health.

5.3.2 Defining ARGs and Other AMR Determinants in the Context of Resistome Risk

Although touched on when discussing Metacompare in Chapter 4, defining what constitutes an ARG and which are relevant for environmental risk assessment is a complex and often subjective undertaking. One of the most widely cited expositions on defining ARGs in the context of risk was authored by Martínez et al. (2015). In brief, the

opinion article contends researchers should be wary of considering ARG database entries, which relate to mutated gene variants, predicted ARGs based on partial sequence homology, intrinsic multidrug efflux pumps and host-antibiotic defence mechanisms. Similar recommendations are echoed elsewhere (Bengtsson-Palme et al., 2017, Wright, 2019).

With particular reference to mutant gene variants, Bengtsson-Palme et al. (2017) highlighted the ease of conflating rifampicin-susceptible *rpoB* genes with genetically similar rifampicin-resistant *rpoB2* mutants. Indeed, analyses of unassembled short reads are unlikely to resolve these differences, and so the wisdom of including *rpoB2* in DeepARG screening for the present work could be questioned.

Additionally, several groups have expressed concern about the number of predicted ARGs in databases lacking further functional validation (Bengtsson-Palme and Larsson, 2015, Martínez et al., 2015, Bengtsson-Palme et al., 2017). To this end, CARD is one database which consciously avoids including putative ARGs, although predicted variants of confirmed ARGs are curated separately (Alcock et al., 2019). However, while it may be justifiable to exclude predicted ARGs when evaluating well-characterised resistomes such as the human gut, perhaps such stringency is less appropriate for more microbially diverse environments like soil, which are thought to be a reservoir of many uncharacterised ARGs (D'Costa et al., 2007).

Interestingly, it has been proposed that intrinsic resistances could still be considered important contributors to environmental resistome risk when they become de-contextualised with the aid of mobilisable elements (Martínez et al., 2015, Bengtsson-Palme et al., 2017). However, Waglechner and Wright (2017) argue that such a view places an unnecessary restriction on the study of the wider resistome.

Ultimately, Martínez et al. (2015) suggested the presence of mobilisable ARGs in known pathogens poses the greatest risk to human health, whereas Bengtsson-Palme and Larsson (2015) argue this framework underestimates the risk posed by mobilisable ARGs

present in non-pathogenic organisms. The inherent short-comings already outlined of attempting to define pathogen lists, only serve to further reinforce the value of a more holistic approach. Nevertheless, the importance of establishing the genetic context of ARGs has been reiterated on numerous occasions (Bengtsson-Palme and Larsson, 2015, Martínez et al., 2015, Wright, 2019, Rice et al., 2020) and this information presently remains scarce with regard to soil resistomes in particular.

Other sources of confusion include the terminology used to distinguish ARGs from genes conferring resistance to other substances which also select bacterial resistance mechanisms (Singer et al., 2016). Certainly, genes which specifically confer metal resistance should not be called ARGs, although in instances where these genes are non-specific efflux pumps, they may straddle these definitions and lead to inconsistent annotations in databases. Furthermore, while the framework outlined by Martínez et al. (2015) refers to mobilisable elements, MRGs are not discussed, despite the fact one could argue they are equally part of the resistome *in toto*. Aside from possible cross-selection, MRGs are relevant since heavy metal-based materials are increasingly being trialled as antimicrobial surfaces in clinical healthcare settings (Page et al., 2009, Weber and Rutala, 2013, García and Parga-Landa, 2021).

5.3.3 Biomarkers as a Tool for Risk Assessment

Resistance gene databases typically include thousands of genes and accordingly their manual curation is an arduous, if nigh on impossible task for researchers to undertake individually. Although some databases are actively curated by the wider scientific community (e.g. CARD), no single database is likely to cover all resistance determinants of interest such as MGEs and MRGs, necessitating database merging and the possible introduction of errors. An alternative approach could involve the identification of biomarkers or indicators. Biomarkers could act as routine indicators for a wider range of resistance determinants which signify an environmental risk of AMR dispersal (Li et al.,

2015a, Jiao et al., 2018, Cacace et al., 2019, Ishii, 2020), or allow metagenomic studies to focus on characterising the genetic context of specific gene targets associated with high clinical risk.

For example, elevated recovery of *cfxA* genes have been associated with antibiotic exposure in newborn infants (Zain et al., 2018, Loo et al., 2020) and adult humans (Duan et al., 2020). As previously mentioned in Chapter 2 and 4, *cfxA* was strongly associated with dairy slurry in the present study and previous publications. In this regard *cfxA* may represent an excellent biomarker for faecal contamination of the environment where antibiotic selection occurs. However, literature concerning *cfxA* selection in cattle is less consistent (discussed in Chapter 4). Of further note, *cfxA* was not detectable in soil within 24 hours of slurry application, which may impair its use as a long-term indicator of manure exposure.

Likewise, although *tetM* is a well-established biomarker of animal waste contamination, long-term application of slurry appeared to mask any immediate increases in soil following amendment. On the other hand, genes like *tet36*, *mefB* and *mefA*, showed initial increases directly after slurry application and degradation over time. Consequently, a combination of slurry biomarkers which allow the differentiation of short- and long-term exposure could be developed, although further validation would be required. In any case, the current work refines and consolidates the host of potential biomarkers which are likely to signify the dispersal of animal waste-associated ARGs in the environment.

The identification of biomarkers which might signify the presence of larger gene clusters has been practiced in the past using network analysis (Li et al., 2015a, Chen et al., 2019b). On the other hand, the sensitivity of methods must be taken into account. For instance, while ARG-ARG network analysis successfully demonstrated connections between *aph3* and *aph6*, which could be verified by the literature and was later further validated by assembled contigs; *sul2* was also shown to co-occur with these genes on

assembled contigs even though network analysis failed to detect this association. It therefore remains evident that such methods are neither infallible nor comprehensive. Consequently, an ensemble approach which uses assembled metagenomic data to support HT-qPCR or unassembled short-read data offers the greatest opportunity for selecting meaningful biomarkers. Finally, the development of concise biomarker arrays could enable the massive parallelisation of future environmental risk surveillance efforts by reducing the need for costly, high-depth metagenomics and extremely large q-PCR arrays.

5.4 Further Research and New Directions

The results of the current work highlight several areas of research that require further attention to clarify policy and make thorough risk assessment practical. Firstly, the presented work suggests an eight week wait-period is not always sufficient to reduce slurry-borne ARGs and bacteria to baseline levels in field soil receiving dairy cattle slurry. In line with the concept of One Health, further work should investigate the likelihood that cut-grass and other material for ensiling can act as a vector for the transfer of antimicrobial resistance determinants to livestock and whether in doing so, there is a potential to exacerbate AMR in the wider farming environment. Secondly, although the present work was able to provide genetic context for the most abundant ARGs in slurry, it was not possible to glean the same for ARGs in soil. As discussed, the acquisition of this information is vital if the behaviour and evolution of ARGs in the soil environment is to be properly understood and the associated risks evaluated. The importance of moving beyond quantification towards genetic context (with special attention to the mobilome) cannot be underestimated. An improved comprehension of how ARGs are configured within the environment can then be used to inform the assembly of quick, affordable, biomarker assays which can be incorporated into routine site-based risk assessment.

References

AAGAARD, K., RIEHLE, K., MA, J., SEGATA, N., MISTRETTA, T.-A., COARFA, C., RAZA, S., ROSENBAUM, S., VAN DEN VEYVER, I. & MILOSAVLJEVIC, A. 2012. A metagenomic approach to characterization of the vaginal microbiome signature in pregnancy. *PloS one,* 7.

AARESTRUP, F. M., AGERSO, Y., GERNER–SMIDT, P., MADSEN, M. & JENSEN, L. B. 2000. Comparison of antimicrobial resistance phenotypes and resistance genes in Enterococcus faecalis and Enterococcus faecium from humans in the community, broilers, and pigs in Denmark. *Diagnostic microbiology and infectious disease,* 37**,** 127-137.

AARESTRUP, F. M., WEGENER, H. C. & COLLIGNON, P. 2008. Resistance in bacteria of the food chain: epidemiology and control strategies. *Expert review of anti-infective therapy,* 6**,** 733-750.

ABUBAKER, J., CEDERLUND, H., ARTHURSON, V. & PELL, M. 2013. Bacterial community structure and microbial activity in different soils amended with biogas residues and cattle slurry. *Applied Soil Ecology,* 72**,** 171-180.

ACEA, M. J., MOORE, C. R. & ALEXANDER, M. 1988. Survival and growth of bacteria introduced into soil. *Soil Biology and Biochemistry,* 20**,** 509-515.

ACHARD, A., VILLERS, C., PICHEREAU, V. & LECLERCQ, R. 2005. New lnu (C) gene conferring resistance to lincomycin by nucleotidylation in Streptococcus agalactiae UCN36. *Antimicrobial agents and chemotherapy,* 49**,** 2716-2719.

AGERSØ, Y., PEDERSEN, A. G. & AARESTRUP, F. M. 2006a. Identification of Tn 5397-like and Tn 916-like transposons and diversity of the tetracycline resistance gene tet (M) in enterococci from humans, pigs and poultry. *Journal of Antimicrobial Chemotherapy,* 57**,** 832-839.

AGERSØ, Y., SENGELØV, G. & JENSEN, L. B. 2004. Development of a rapid method for direct detection of tet (M) genes in soil from Danish farmland. *Environment international,* 30**,** 117-122.

AGERSØ, Y., WULFF, G., VACLAVIK, E., HALLING-SØRENSEN, B. & JENSEN, L. B. 2006b. Effect of tetracycline residues in pig manure slurry on tetracycline-resistant bacteria and resistance gene tet (M) in soil microcosms. *Environment international,* 32**,** 876-882.

AGGA, G. E., ARTHUR, T. M., DURSO, L. M., HARHAY, D. M. & SCHMIDT, J. W. 2015. Antimicrobial-resistant bacterial populations and antimicrobial resistance genes obtained from environments impacted by livestock and municipal waste. *PLoS One,* 10**,** e0132586.

AHDB. 2020a. *UK and EU cow numbers* [Online]. Available: https://ahdb.org.uk/dairy/uk-and-eu-cow-numbers [Accessed 2020].

AHDB. 2020b. *UK producer numbers* [Online]. Available: https://ahdb.org.uk/dairy/uk-producer-numbers [Accessed 2021].

AISLABIE, J., DESLIPPE, J. R. & DYMOND, J. 2013. Soil microbes and their contribution to soil services. *Ecosystem services in New Zealand–conditions and trends. Manaaki Whenua Press, Lincoln, New Zealand***,** 143-161.

ALCALDE-RICO, M., HERNANDO-AMADO, S., BLANCO, P. & MARTÍNEZ, J. L. 2016. Multidrug efflux pumps at the crossroad between antibiotic resistance and bacterial virulence. *Frontiers in microbiology,* 7**,** 1483.

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. & MCARTHUR, A. G. 2019. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Research,* 48**,** D517-D525.

ALONSO, H. & GREADY, J. E. 2006. Integron-sequestered dihydrofolate reductase: a recently redeployed enzyme. *Trends in microbiology,* 14**,** 236-242.

ANANTHAM, S. & HALL, R. M. 2012. pCERC1, a small, globally disseminated plasmid carrying the dfrA14 cassette in the strA gene of the sul2-strA-strB gene cluster. *Microbial drug resistance,* 18**,** 364-371.

ANANTHARAMAN, K., HAUSMANN, B., JUNGBLUTH, S. P., KANTOR, R. S., LAVY, A., WARREN, L. A., RAPPÉ, M. S., PESTER, M., LOY, A. & THOMAS, B. C. 2018. Expanded diversity of microbial groups that shape the dissimilatory sulfur cycle. *The ISME journal,* 12**,** 1715-1728.

ANDERSON, M., GORLEY, R. & CLARKE, K. 2008. PERMANOVA+ for PRIMER: guide to software and statistical methods. PRIMER-E Ltd.

ANDERSSON, D. I. & HUGHES, D. 2011. Persistence of antibiotic resistance in bacterial populations. *FEMS microbiology reviews,* 35**,** 901-911.

ANDERSSON, D. I. & HUGHES, D. 2012. Evolution of antibiotic resistance at non-lethal drug concentrations. *Drug Resistance Updates,* 15**,** 162-172.

ANDREWS JR, R. E., JOHNSON, W. S., GUARD, A. R. & MARVIN, J. D. 2004. Survival of enterococci and Tn 916-like conjugative transposons in soil. *Canadian journal of microbiology,* 50**,** 957-966.

ANDREWS, S. 2015. *FastQC: A Quality Control Tool for High Throughput Sequence Data* [Online]. Available: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ [Accessed].

ANTUNES, N. T., LAMOUREAUX, T. L., TOTH, M., STEWART, N. K., FRASE, H. & VAKULENKO, S. B. 2014. Class D β-lactamases: are they all carbapenemases? *Antimicrobial agents and chemotherapy,* 58**,** 2119-2125.

ARANGO-ARGOTY, G., GARNER, E., PRUDEN, A., HEATH, L. S., VIKESLAND, P. & ZHANG, L. 2018. DeepARG: a deep learning approach for predicting antibiotic resistance genes from metagenomic data. *Microbiome,* 6**,** 1-15.

ARDANUY, C., TUBAU, F., LINARES, J., DOMÍNGUEZ, M. A., PALLARÉS, R. & MARTÍN, R. 2005. Distribution of subclasses mefA and mefE of the mefA gene among clinical isolates of macrolide-resistant (M-phenotype) Streptococcus pneumoniae, viridans group streptococci, and Streptococcus pyogenes. *Antimicrobial agents and chemotherapy,* 49**,** 827-829.

ARPIN, C., DAUBE, H., TESSIER, F. & QUENTIN, C. 1999. Presence of mefA and mefEGenes in Streptococcus agalactiae. *Antimicrobial agents and chemotherapy,* 43**,** 944-946.

ARREDONDO, A., BLANC, V., MOR, C., NART, J. & LEÓN, R. 2020. Resistance to β-lactams and distribution of β-lactam resistance genes in subgingival microbiota from Spanish patients with periodontitis. *Clinical Oral Investigations*.

ASAI, T., KOJIMA, A., HARADA, K., ISHIHARA, K., TAKAHASHI, T. & TAMURA, Y. 2005. Correlation between the usage volume of veterinary therapeutic antimicrobials and resistance in Escherichia coli isolated from the feces of food-producing animals in Japan. *Japanese journal of infectious diseases,* 58**,** 369.

ASHBOLT, N. J., AMÉZQUITA, A., BACKHAUS, T., BORRIELLO, P., BRANDT, K. K., COLLIGNON, P., COORS, A., FINLEY, R., GAZE, W. H. & HEBERER, T. 2013. Human health risk assessment (HHRA) for environmental development and transfer of antibiotic resistance. *Environmental health perspectives,* 121**,** 993-1001.

ASLAM, B., WANG, W., ARSHAD, M. I., KHURSHID, M., MUZAMMIL, S., RASOOL, M. H., NISAR, M. A., ALVI, R. F., ASLAM, M. A. & QAMAR, M. U. 2018. Antibiotic resistance: a rundown of a global crisis. *Infection and drug resistance,* 11**,** 1645.

ATEBA, C. N., TABI, N. M., FRI, J., BISSONG, M. E. A. & BEZUIDENHOUT, C. C. 2020. Occurrence of Antibiotic-Resistant Bacteria and Genes in Two Drinking Water Treatment and Distribution Systems in the North-West Province of South Africa. *Antibiotics,* 9**,** 745.

AYLING, M., CLARK, M. D. & LEGGETT, R. M. 2020. New approaches for metagenome assembly with short reads. *Briefings in bioinformatics,* 21**,** 584-594.

AZARGUN, R., GHOLIZADEH, P., SADEGHI, V., HOSAINZADEGAN, H., TARHRIZ, V., MEMAR, M. Y., PORMOHAMMAD, A. & EYVAZI, S. 2020. Molecular mechanisms associated with quinolone resistance in Enterobacteriaceae: review and update. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 114**,** 770-781.

BAKER-AUSTIN, C., WRIGHT, M. S., STEPANAUSKAS, R. & MCARTHUR, J. 2006. Co-selection of antibiotic and metal resistance. *Trends in microbiology,* 14**,** 176-182.
BAKER, M., HOBMAN, J. L., DODD, C. E., RAMSDEN, S. J. & STEKEL, D. J. 2016. Mathematical modelling of antimicrobial resistance in agricultural waste highlights importance of gene transfer rate. *FEMS Microbiology Ecology,* 92.
BALCAZAR, J. L. 2014. Bacteriophages as vehicles for antibiotic resistance genes in the environment. *PLoS Pathog,* 10**,** e1004219.
BARTLETT, P. C., MILLER, G. Y., LANCE, S. E. & HEIDER, L. E. 1992. Environmental and managerial determinants of somatic cell counts and clinical mastitis incidence in Ohio dairy herds. *Preventive Veterinary Medicine,* 14**,** 195-207.
BATRA, P., MATHUR, P. & MISRA, M. C. 2016. Aeromonas spp.: an emerging nosocomial pathogen. *Journal of laboratory physicians,* 8**,** 1.
BELL, S. C., GARLAND, S. & ALFORD, R. A. 2018. Increased numbers of culturable inhibitory bacterial taxa may mitigate the effects of Batrachochytrium dendrobatidis in australian wet tropics frogs. *Frontiers in microbiology,* 9**,** 1604.
BELLANGER, X., PAYOT, S., LEBLOND-BOURGET, N. & GUÉDON, G. 2014. Conjugative and mobilizable genomic islands in bacteria: evolution and diversity. *FEMS Microbiology Reviews,* 38**,** 720-760.
BEN, Y., FU, C., HU, M., LIU, L., WONG, M. H. & ZHENG, C. 2019. Human health risk assessment of antibiotic resistance associated with antibiotic residues in the environment: a review. *Environmental research,* 169**,** 483-493.
BENGTSSON-PALME, J., KRISTIANSSON, E. & LARSSON, D. J. 2018. Environmental factors influencing the development and spread of antibiotic resistance. *FEMS microbiology reviews,* 42**,** fux053.
BENGTSSON-PALME, J. & LARSSON, D. J. 2015. Antibiotic resistance genes in the environment: prioritizing risks. *Nature Reviews Microbiology,* 13**,** 396-396.
BENGTSSON-PALME, J., LARSSON, D. J. & KRISTIANSSON, E. 2017. Using metagenomics to investigate human and environmental resistomes. *Journal of Antimicrobial Chemotherapy,* 72**,** 2690-2703.
BERGLUND, B. 2015. Environmental dissemination of antibiotic resistance genes and correlation to anthropogenic contamination with antibiotics. *Infection ecology & epidemiology,* 5**,** 28564.
BEYE, M., BAKOUR, S., TRAORE, S. I., RATHORED, J., LABAS, N., RAOULT, D. & FOURNIER, P.-E. 2018. Draft genome sequence of Fermentimonas caenicola strain SIT8, isolated from the human gut. *Standards in genomic sciences,* 13**,** 8.
BIGGS, A., BARRETT, D., BRADLEY, A., GREEN, M., REYHER, K. & ZADOKS, R. 2016. Antibiotic dry cow therapy: where next? *Veterinary Record,* 178**,** 93-94.
BINTA, B. & PATEL, M. 2016. Detection of cfxA2, cfxA3, and cfxA6 genes in beta-lactamase producing oral anaerobes. *Journal of Applied Oral Science,* 24**,** 142-147.
BLACK, C. A. 1965. *Methods of Soil Analysis: Part 1 Physical and Mineralogical Properties, Including Statistics of Measurement and Sampling*, American Society of Agronomy.
BLAU, K., CASADEVALL, L., WOLTERS, B., VAN DEN MEERSCHE, T., KREUZIG, R., SMALLA, K. & JECHALKE, S. 2018. Soil texture-depending effects of doxycycline and streptomycin applied with manure on the bacterial community composition and resistome. *FEMS microbiology ecology,* 94**,** fix145.
BOLÓN-CANEDO, V., SÁNCHEZ-MAROÑO, N. & ALONSO-BETANZOS, A. 2015. Recent advances and emerging challenges of feature selection in the context of big data. *Knowledge-Based Systems,* 86**,** 33-45.
BOTELHO, J. & SCHULENBURG, H. 2020. The role of integrative and conjugative elements in antibiotic resistance evolution. *Trends in Microbiology*.
BOUCHER, H. W., TALBOT, G. H., BRADLEY, J. S., EDWARDS, J. E., GILBERT, D., RICE, L. B., SCHELD, M., SPELLBERG, B. & BARTLETT, J. 2009. Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society of America. *Clinical infectious diseases,* 48**,** 1-12.

BRADLEY, A., LEACH, K., BREEN, J., GREEN, L. & GREEN, M. 2007. Survey of the incidence and aetiology of mastitis on dairy farms in England and Wales. *Veterinary Record,* 160**,** 253-258.
BRADLEY, A. J. 2002. Bovine mastitis: an evolving disease. *The veterinary journal,* 164**,** 116-128.
BREEN, J., BRADLEY, A., MAY, B., ANSCOMBE, J., BREARLEY, B. & ELKINS, P. 2014. Clinical forum: the role of dry cow therapy in the modern dairy herd. *Livestock,* 19**,** 262-270.
BREITWIESER, F. P., LU, J. & SALZBERG, S. L. 2017. A review of methods and databases for metagenomic classification and assembly. *Briefings in bioinformatics,* 20**,** 1125-1136.
BREITWIESER, F. P., LU, J. & SALZBERG, S. L. 2019. A review of methods and databases for metagenomic classification and assembly. *Briefings in bioinformatics,* 20**,** 1125-1136.
BROOKS-POLLOCK, E. & KEELING, M. 2009. Herd size and bovine tuberculosis persistence in cattle farms in Great Britain. *Preventive veterinary medicine,* 92**,** 360-365.
BROWN, C. T., HUG, L. A., THOMAS, B. C., SHARON, I., CASTELLE, C. J., SINGH, A., WILKINS, M. J., WRIGHTON, K. C., WILLIAMS, K. H. & BANFIELD, J. F. 2015. Unusual biology across a group comprising more than 15% of domain Bacteria. *Nature,* 523**,** 208-211.
BRUNTON, L., DUNCAN, D., COLDHAM, N., SNOW, L. & JONES, J. 2012. A survey of antimicrobial usage on dairy farms and waste milk feeding practices in England and Wales. *Veterinary Record,* 171**,** 296-296.
BRYAN, A., SHAPIR, N. & SADOWSKY, M. J. 2004. Frequency and distribution of tetracycline resistance genes in genetically diverse, nonselected, and nonclinical Escherichia coli strains isolated from diverse human and animal sources. *Appl. Environ. Microbiol.,* 70**,** 2503-2507.
BUKIN, Y. S., GALACHYANTS, Y. P., MOROZOV, I., BUKIN, S., ZAKHARENKO, A. & ZEMSKAYA, T. 2019. The effect of 16S rRNA region choice on bacterial community metabarcoding results. *Scientific data,* 6**,** 190007.
BURGOS, J., ELLINGTON, B. & VARELA, M. 2005. Presence of multidrug-resistant enteric bacteria in dairy farm topsoil. *Journal of dairy science,* 88**,** 1391-1398.
BUTLER, M. S. & PATERSON, D. L. 2020. Antibiotics in the clinical pipeline in October 2019. *The Journal of antibiotics,* 73**,** 329-364.
BYAPPANAHALLI, M. N., NEVERS, M. B., KORAJKIC, A., STALEY, Z. R. & HARWOOD, V. J. 2012. Enterococci in the environment. *Microbiol. Mol. Biol. Rev.,* 76**,** 685-706.
BYRNE-BAILEY, K., GAZE, W., KAY, P., BOXALL, A., HAWKEY, P. & WELLINGTON, E. 2009. Prevalence of sulfonamide resistance genes in bacterial isolates from manured agricultural soils and pig slurry in the United Kingdom. *Antimicrobial Agents and Chemotherapy,* 53**,** 696-702.
CACACE, D., FATTA-KASSINOS, D., MANAIA, C. M., CYTRYN, E., KREUZINGER, N., RIZZO, L., KARAOLIA, P., SCHWARTZ, T., ALEXANDER, J. & MERLIN, C. 2019. Antibiotic resistance genes in treated wastewater and in the receiving water bodies: A pan-European survey of urban settings. *Water research,* 162**,** 320-330.
CANTÓN, R., GONZÁLEZ-ALBA, J. M. & GALÁN, J. C. 2012. CTX-M enzymes: origin and diffusion. *Frontiers in microbiology,* 3**,** 110.
CARD, R. M., WARBURTON, P. J., MACLAREN, N., MULLANY, P., ALLAN, E. & ANJUM, M. F. 2014. Application of microarray and functional-based screening methods for the detection of antimicrobial resistance genes in the microbiomes of healthy humans. *PloS one,* 9**,** e86428.
CARDING, S., VERBEKE, K., VIPOND, D. T., CORFE, B. M. & OWEN, L. J. 2015. Dysbiosis of the gut microbiota in disease. *Microbial ecology in health and disease,* 26**,** 26191.
CARINI, P., MARSDEN, P. J., LEFF, J. W., MORGAN, E. E., STRICKLAND, M. S. & FIERER, N. 2016. Relic DNA is abundant in soil and obscures estimates of soil microbial diversity. *Nature microbiology,* 2**,** 1-6.
CARO-QUINTERO, A., RITALAHTI, K., CUSICK, K., LÖFFLER, F. & KONSTANTINIDIS, K. 2012. The chimeric genome of Sphaerochaeta: nonspiral spirochetes that break with the prevalent dogma in spirochete biology. *MBio,* 3.
CARVALHO, M. J., MARTÍNEZ-MURCIA, A., ESTEVES, A. C., CORREIA, A. & SAAVEDRA, M. J. 2012. Phylogenetic diversity, antibiotic resistance and virulence traits of Aeromonas spp. from

untreated waters for human consumption. *International journal of food microbiology,* 159**,** 230-239.
CASSINI, A., HÖGBERG, L. D., PLACHOURAS, D., QUATTROCCHI, A., HOXHA, A., SIMONSEN, G. S., COLOMB-COTINAT, M., KRETZSCHMAR, M. E., DEVLEESSCHAUWER, B. & CECCHINI, M. 2019. Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the EU and the European Economic Area in 2015: a population-level modelling analysis. *The Lancet infectious diseases,* 19**,** 56-66.
CERVANTES, J., YOKOBORI, N. & HONG, B.-Y. 2020. Genetic Identification and Drug-Resistance Characterization of Mycobacterium tuberculosis Using a Portable Sequencing Device. A Pilot Study. *Antibiotics,* 9**,** 548.
CEYLAN, E., BERKTAS, M. & AĞAOĞLU, Z. 2009. The occurrence and antibiotic resistance of motile Aeromonas in livestock. *Tropical animal health and production,* 41**,** 199-204.
CHAMBERS, L., YANG, Y., LITTIER, H., RAY, P., ZHANG, T., PRUDEN, A., STRICKLAND, M. & KNOWLTON, K. 2015. Metagenomic analysis of antibiotic resistance genes in dairy cow feces following therapeutic administration of third generation cephalosporin. *PloS one,* 10.
CHANDRA, N. & KUMAR, S. 2017. Antibiotics producing soil microorganisms. *Antibiotics and Antibiotics Resistance Genes in Soils.* Springer.
CHANG, Q., WANG, W., REGEV-YOCHAY, G., LIPSITCH, M. & HANAGE, W. P. 2015. Antibiotics in agriculture and the risk to human health: how worried should we be? *Evolutionary applications,* 8**,** 240-247.
CHANTZIARAS, I., BOYEN, F., CALLENS, B. & DEWULF, J. 2014. Correlation between veterinary antimicrobial use and antimicrobial resistance in food-producing animals: a report on seven countries. *Journal of Antimicrobial Chemotherapy,* 69**,** 827-834.
CHAO, A., GOTELLI, N. J., HSIEH, T., SANDER, E. L., MA, K., COLWELL, R. K. & ELLISON, A. M. 2014. Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. *Ecological monographs,* 84**,** 45-67.
CHATTERJEE, A., SHARMA, A. K., MAHATHA, A. C., BANERJEE, S. K., KUMAR, M., SAHA, S., BASU, J. & KUNDU, M. 2018. Global mapping of MtrA-binding sites links MtrA to regulation of its targets in Mycobacterium tuberculosis. *Microbiology,* 164**,** 99-110.
CHAUDHRY, V., REHMAN, A., MISHRA, A., CHAUHAN, P. S. & NAUTIYAL, C. S. 2012. Changes in bacterial community structure of agricultural land due to long-term organic and chemical amendments. *Microbial ecology,* 64**,** 450-460.
CHAUMEIL, P.-A., MUSSIG, A. J., HUGENHOLTZ, P. & PARKS, D. H. 2020. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Oxford University Press.
CHAWG 2020. GB Cattle Health & Welfare Group, 5th Report.
CHE, Y., XIA, Y., LIU, L., LI, A.-D., YANG, Y. & ZHANG, T. 2019. Mobile antibiotic resistome in wastewater treatment plants revealed by Nanopore metagenomic sequencing. *Microbiome,* 7**,** 1-13.
CHECCUCCI, A., TREVISI, P., LUISE, D., MODESTO, M., BLASIOLI, S., BRASCHI, I. & MATTARELLI, P. 2020. Exploring the animal waste resistome: the spread of antimicrobial resistance genes through the use of livestock manure. *Frontiers in Microbiology,* 11**,** 1416.
CHEE-SANFORD, J. C., AMINOV, R., KRAPAC, I. J., GARRIGUES-JEANJEAN, N. & MACKIE, R. I. 2001. Occurrence and diversity of tetracycline resistance genes in lagoons and groundwater underlying two swine production facilities. *Applied and environmental microbiology,* 67**,** 1494-1502.
CHEE-SANFORD, J. C., MACKIE, R. I., KOIKE, S., KRAPAC, I. G., LIN, Y. F., YANNARELL, A. C., MAXWELL, S. & AMINOV, R. I. 2009. Fate and transport of antibiotic residues and antibiotic resistance genes following land application of manure waste. *Journal of environmental quality,* 38**,** 1086-1108.
CHEN, C., PANKOW, C. A., OH, M., HEATH, L. S., ZHANG, L., DU, P., XIA, K. & PRUDEN, A. 2019a. Effect of antibiotic use and composting on antibiotic resistance gene abundance and resistome

risks of soils receiving manure-derived amendments. *Environment international,* 128**,** 233-243.
CHEN, J. & CHEN, M. J. 2018. Package 'GUniFrac'. *The Comprehensive R Archive Network (CRAN)*.
CHEN, L.-X., ANANTHARAMAN, K., SHAIBER, A., EREN, A. M. & BANFIELD, J. F. 2020. Accurate and complete genomes from metagenomes. *Genome research,* 30**,** 315-333.
CHEN, Q.-L., AN, X.-L., LI, H., ZHU, Y.-G., SU, J.-Q. & CUI, L. 2017. Do manure-borne or indigenous soil microorganisms influence the spread of antibiotic resistance genes in manured soil? *Soil Biology and Biochemistry,* 114**,** 229-237.
CHEN, Q., AN, X., LI, H., SU, J., MA, Y. & ZHU, Y.-G. 2016a. Long-term field application of sewage sludge increases the abundance of antibiotic resistance genes in soil. *Environment International,* 92**,** 1-10.
CHEN, Y., JI, F., GUO, J., SHI, D., FANG, D. & LI, L. 2016b. Dysbiosis of small intestinal microbiota in liver cirrhosis and its association with etiology. *Scientific reports,* 6**,** 1-9.
CHEN, Z., ZHANG, W., YANG, L., STEDTFELD, R. D., PENG, A., GU, C., BOYD, S. A. & LI, H. 2019b. Antibiotic resistance genes and bacterial communities in cornfield and pasture soils receiving swine and dairy manures. *Environmental Pollution,* 248**,** 947-957.
CHENG, J.-H., TANG, X.-Y. & CUI, J.-F. 2019. Effect of long-term manure slurry application on the occurrence of antibiotic resistance genes in arable purple soil (entisol). *Science of The Total Environment,* 647**,** 853-861.
CIRIC, L., JASNI, A., DE VRIES, E., AGERSO, Y., MULLANY, P. & ROBERTS, A. P. 2013. The Tn916/Tn1545 Family of Conjugative Transposon. Landes Bioscience.
CLANCY, J., PETITPAS, J., DIB-HAJJ, F., YUAN, W., CRONAN, M., KAMATH, A. V., BERGERON, J. & RETSEMA, J. A. 1996. Molecular cloning and functional analysis of a novel macrolide-resistance determinant, mefA, from Streptococcus pyogenes. *Molecular microbiology,* 22**,** 867-879.
CLARKE, K. & GORLEY, R. 2006. PRIMER v6: User Manual/Tutorial. PRIMER-E

6ed. Plymouth

CLEARY, D. W., BISHOP, A. H., ZHANG, L., TOPP, E., WELLINGTON, E. M. & GAZE, W. H. 2016. Long-term antibiotic exposure in soil is associated with changes in microbial community structure and prevalence of class 1 integrons. *FEMS microbiology ecology,* 92.
CLEMENTE, J. C., PEHRSSON, E. C., BLASER, M. J., SANDHU, K., GAO, Z., WANG, B., MAGRIS, M., HIDALGO, G., CONTRERAS, M. & NOYA-ALARCÓN, Ó. 2015. The microbiome of uncontacted Amerindians. *Science advances,* 1**,** e1500183.
COELHO, J. J., HENNESSY, A., CASEY, I., BRAGANÇA, C. R. S., WOODCOCK, T. & KENNEDY, N. 2020. Biofertilisation with anaerobic digestates: A field study of effects on soil microbial abundance and diversity. *Applied Soil Ecology,* 147**,** 103403.
COLACO, S., KUMAR, S., TAMANG, A. & BIJU, V. G. 2019. A Review on Feature Selection Algorithms. *Emerging Research in Computing, Information, Communication and Applications.* Springer.
COLLIGNON, P. & VOSS, A. 2015. China, what antibiotics and what volumes are used in food production animals? : BioMed Central.
COLLIGNON, P. C., CONLY, J. M., ANDREMONT, A., MCEWEN, S. A., AIDARA-KANE, A., WORLD HEALTH ORGANIZATION ADVISORY GROUP, B. M. O. I. S. O. A. R., AGERSO, Y., ANDREMONT, A., COLLIGNON, P. & CONLY, J. 2016. World Health Organization ranking of antimicrobials according to their importance in human medicine: a critical step for developing risk management strategies to control antimicrobial resistance from food animal production. *Clinical Infectious Diseases,* 63**,** 1087-1093.
COLLIGNON, P. J. & MCEWEN, S. A. 2019. One health—its importance in helping to better control antimicrobial resistance. *Tropical Medicine and Infectious Disease,* 4**,** 22.

COOLS, D., MERCKX, R., VLASSAK, K. & VERHAEGEN, J. 2001. Survival of E. coli and Enterococcus spp. derived from pig slurry in soils of different texture. *Applied Soil Ecology,* 17**,** 53-62.
COX, G. & WRIGHT, G. D. 2013. Intrinsic antibiotic resistance: mechanisms, origins, challenges and solutions. *International Journal of Medical Microbiology,* 303**,** 287-292.
CYCOŃ, M., MROZIK, A. & PIOTROWSKA-SEGET, Z. 2019. Antibiotics in the soil environment—degradation and their impact on microbial activity and diversity. *Frontiers in microbiology,* 10.
D'COSTA, V. M., GRIFFITHS, E. & WRIGHT, G. D. 2007. Expanding the soil antibiotic resistome: exploring environmental diversity. *Current opinion in microbiology,* 10**,** 481-489.
D'COSTA, V. M., KING, C. E., KALAN, L., MORAR, M., SUNG, W. W., SCHWARZ, C., FROESE, D., ZAZULA, G., CALMELS, F. & DEBRUYNE, R. 2011. Antibiotic resistance is ancient. *Nature,* 477**,** 457-461.
DAI, Z., SU, W., CHEN, H., BARBERÁN, A., ZHAO, H., YU, M., YU, L., BROOKES, P. C., SCHADT, C. W. & CHANG, S. X. 2018. Long-term nitrogen fertilization decreases bacterial diversity and favors the growth of Actinobacteria and Proteobacteria in agro-ecosystems across the globe. *Global change biology,* 24**,** 3452-3461.
DANG, B., ZHANG, H., LI, Z., MA, S. & XU, Z. 2020. Coexistence of the bla NDM-1-carrying plasmid pWLK-NDM and the bla KPC-2-carrying plasmid pWLK-KPC in a Raoultella ornithinolytica isolate. *Scientific reports,* 10**,** 1-9.
DAVIES, J. & DAVIES, D. 2010. Origins and evolution of antibiotic resistance. *Microbiology and molecular biology reviews,* 74**,** 417-433.
DAVIES, R. & WALES, A. 2019. Antimicrobial resistance on farms: a review including biosecurity and the potential role of disinfectants in resistance selection. *Comprehensive reviews in food science and food safety,* 18**,** 753-774.
DAVIS, M. A., BESSER, T. E., ORFE, L. H., BAKER, K. N., LANIER, A. S., BROSCHAT, S. L., NEW, D. & CALL, D. R. 2011. Genotypic-phenotypic discrepancies between antibiotic resistance characteristics of Escherichia coli isolates from calves in management settings with high and low antibiotic use. *Applied and environmental microbiology,* 77**,** 3293-3299.
DAVIS, M. F., RANKIN, S. C., SCHURER, J. M., COLE, S., CONTI, L., RABINOWITZ, P., GRAY, G., KAHN, L., MACHALABA, C. & MAZET, J. 2017. Checklist for one health epidemiological reporting of evidence (COHERE). *One Health,* 4**,** 14-21.
DE BRIYNE, N., ATKINSON, J., BORRIELLO, S. & POKLUDOVÁ, L. 2014. Antibiotics used most commonly to treat animals in Europe. *Veterinary Record,* 175**,** 325-325.
DE KRAKER, M. E., STEWARDSON, A. J. & HARBARTH, S. 2016. Will 10 million people die a year due to antimicrobial resistance by 2050? *PLoS medicine,* 13**,** e1002184.
DE MENEZES, A. B., LEWIS, E., O'DONOVAN, M., O'NEILL, B. F., CLIPSON, N. & DOYLE, E. M. 2011. Microbiome analysis of dairy cows fed pasture or total mixed ration diets. *FEMS microbiology ecology,* 78**,** 256-265.
DE OLIVEIRA, D. M., FORDE, B. M., KIDD, T. J., HARRIS, P. N., SCHEMBRI, M. A., BEATSON, S. A., PATERSON, D. L. & WALKER, M. J. 2020. Antimicrobial resistance in ESKAPE pathogens. *Clinical microbiology reviews,* 33**,** e00181-19.
DEFRA 2013. Report of the Joint Government and Industry Slurry Management and Storage Project.
DEFRA 2016. The national inventory and map of livestock manure loadings to agricultural land (Manures-GIS) - WQ0103.
DEGENHARDT, F., SEIFERT, S. & SZYMCZAK, S. 2019. Evaluation of variable selection methods for random forests and omics data sets. *Briefings in bioinformatics,* 20**,** 492-503.
DEGRUTTOLA, A. K., LOW, D., MIZOGUCHI, A. & MIZOGUCHI, E. 2016. Current understanding of dysbiosis in disease in human and animal models. *Inflammatory bowel diseases,* 22**,** 1137-1150.
DEY, A., ADITHI, V. & CHATTERJI, D. 2012. Co-evolution of RNA polymerase with RbpA in the phylum Actinobacteria. *Applied & translational genomics,* 1**,** 9-20.

DIAS, C., BORGES, A., SAAVEDRA, M. J. & SIMÕES, M. 2018. Biofilm formation and multidrug-resistant Aeromonas spp. from wild animals. *Journal of global antimicrobial resistance,* 12**,** 227-234.

DIBNER, J. & RICHARDS, J. 2005. Antibiotic growth promoters in agriculture: history and mode of action. *Poultry science,* 84**,** 634-643.

DING, Y., TEO, J. W., DRAUTZ-MOSES, D. I., SCHUSTER, S. C., GIVSKOV, M. & YANG, L. 2018. Acquisition of resistance to carbapenem and macrolide-mediated quorum sensing inhibition by Pseudomonas aeruginosa via ICE Tn4371 6385. *Communications biology,* 1**,** 1-10.

DITZLER, G., MORRISON, J. C., LAN, Y. & ROSEN, G. L. 2015. Fizzy: feature subset selection for metagenomics. *BMC bioinformatics,* 16**,** 358.

DOLLIVER, H., KUMAR, K. & GUPTA, S. 2007. Sulfamethazine uptake by plants from manure-amended soil. *Journal of environmental quality,* 36**,** 1224-1230.

DOMINGUES, S., DA SILVA, G. J. & NIELSEN, K. M. 2012. Integrons: vehicles and pathways for horizontal dissemination in bacteria. *Mobile genetic elements,* 2**,** 211-223.

DOSTER, E., LAKIN, S. M., DEAN, C. J., WOLFE, C., YOUNG, J. G., BOUCHER, C., BELK, K. E., NOYES, N. R. & MORLEY, P. S. 2020. MEGARes 2.0: a database for classification of antimicrobial drug, biocide and metal resistance determinants in metagenomic sequence data. *Nucleic acids research,* 48**,** D561-D569.

DOUGHARI, H. J., NDAKIDEMI, P. A., HUMAN, I. S. & BENADE, S. 2009. The ecology, biology and pathogenesis of Acinetobacter spp.: an overview. *Microbes and environments***,** 1103150282-1103150282.

DRIEHUIS, F. & ELFERINK, S. O. 2000. The impact of the quality of silage on animal health and food safety: a review. *Veterinary Quarterly,* 22**,** 212-216.

DRIEHUIS, F., WILKINSON, J., JIANG, Y., OGUNADE, I. & ADESOGAN, A. 2018. Silage review: animal and human health risks from silage. *Journal of Dairy Science,* 101**,** 4093-4110.

DROUIN, P. & LAFRENIÈRE, C. 2012. Clostridial spores in animal feeds and milk. *Milk Production-An up-to-date overview of animal nutrition, management and health.* IntechOpen.

DUAN, Y., CHEN, Z., TAN, L., WANG, X., XUE, Y., WANG, S., WANG, Q., DAS, R., LIN, H. & HOU, J. 2020. Gut resistomes, microbiota and antibiotic residues in Chinese patients undergoing antibiotic administration and healthy individuals. *Science of the Total Environment,* 705**,** 135674.

DUNGAN, R. S., STRAUSBAUGH, C. A. & LEYTEM, A. B. 2019. Survey of selected antibiotic resistance genes in agricultural and non-agricultural soils in south-central Idaho. *FEMS microbiology ecology,* 95**,** fiz071.

DURSO, L. M., HARHAY, G. P., BONO, J. L. & SMITH, T. P. 2011. Virulence-associated and antibiotic resistance genes of microbial populations in cattle feces analyzed using a metagenomic approach. *Journal of microbiological methods,* 84**,** 278-282.

DURSO, L. M., WEDIN, D. A., GILLEY, J. E., MILLER, D. N. & MARX, D. B. 2016. Assessment of selected antibiotic resistances in ungrazed native Nebraska prairie soils. *Journal of Environmental Quality,* 45**,** 454-462.

EC 2003. Council Directive 2003/1831/EC of 22nd September 2003 on Additives For Use In Animal Nutrition. *In:* COUNCIL, E. (ed.).

ECDC 2020. Antimicrobial resistance in the EU/EEA (EARS-Net), Annual Epidemiological Report for 2019.

ELIOPOULOS, G. M. & HUOVINEN, P. 2001. Resistance to trimethoprim-sulfamethoxazole. *Clinical infectious diseases,* 32**,** 1608-1614.

ELLIS, J., HINDRICHSEN, I., KLOP, G., KINLEY, R., MILORA, N., BANNINK, A. & DIJKSTRA, J. 2016. Effects of lactic acid bacteria silage inoculation on methane emission and productivity of Holstein Friesian dairy cattle. *Journal of dairy science,* 99**,** 7159-7174.

EVANS, N. & BORDER, P. 2018. Reducing UK Antibiotic Use in Animals. Houses of Parliament, Parliamentary Office of Science and Technology.

FAHRENFELD, N., KNOWLTON, K., KROMETIS, L. A., HESSION, W. C., XIA, K., LIPSCOMB, E., LIBUIT, K., GREEN, B. L. & PRUDEN, A. 2014. Effect of manure application on abundance of antibiotic resistance genes and their attenuation rates in soil: field-scale mass balance approach. *Environmental science & technology,* 48**,** 2643-2650.

FANG, P., PENG, F., GAO, X., XIAO, P. & YANG, J. 2019. Decoupling the dynamics of bacterial taxonomy and antibiotic resistance function in a subtropical urban reservoir as revealed by high-frequency sampling. *Frontiers in microbiology,* 10**,** 1448.

FAO 2016. The FAO Action Plan on Antimicrobial Resistance 2016-2020.

FAUST, K. & RAES, J. 2016. CoNet app: inference of biological association networks using Cytoscape. *F1000Research,* 5.

FDA 2013. New Animal Drugs and New Animal Drug Combination Products Administered in or on Medicated Feed or Drinking Water of Food-Producing Animals: Recommendations for Drug Sponsors for Voluntarily Aligning Product Use Conditions with GFI #209.

FDA 2019. Summary Report On Antimicrobials Sold or Distributed for Use in Food-Producing Animals.

FENG, X., CHAMBERS, L. R. & KNOWLTON, K. F. 2020. Antibiotic resistance genes in the faeces of dairy cows following short-term therapeutic and prophylactic antibiotic administration. *Journal of Applied Animal Research,* 48**,** 34-37.

FENG, X., LITTIER, H. M., KNOWLTON, K. F., GARNER, E. & PRUDEN, A. 2019. The impacts of feeding milk with antibiotics on the fecal microbiome and antibiotic resistance genes in dairy calves. *Canadian Journal of Animal Science,* 100**,** 69-76.

FENLON, D., OGDEN, I., VINTEN, A. & SVOBODA, I. 2000. The fate of Escherichia coli and E. coli O157 in cattle slurry after application to land. *Journal of Applied Microbiology,* 88**,** 149S-156S.

FERNÁNDEZ-BRAVO, A. & FIGUERAS, M. J. 2020. An update on the genus Aeromonas: taxonomy, epidemiology, and pathogenicity. *Microorganisms,* 8**,** 129.

FEWTRELL, L. & BARTRAM, J. 2001. *Water quality: guidelines, standards & health*, IWA publishing.

FIERER, N., BRADFORD, M. A. & JACKSON, R. B. 2007. Toward an ecological classification of soil bacteria. *Ecology,* 88**,** 1354-1364.

FISCHER, A., LILJANDER, A., KASPAR, H., MURIUKI, C., FUXELIUS, H.-H., BONGCAM-RUDLOFF, E., DE VILLIERS, E. P., HUBER, C. A., FREY, J. & DAUBENBERGER, C. 2013. Camel Streptococcus agalactiae populations are associated with specific disease complexes and acquired the tetracycline resistance gene tetM via a Tn916-like element. *Veterinary research,* 44**,** 86.

FISHER, J. C., EREN, A. M., GREEN, H. C., SHANKS, O. C., MORRISON, H. G., VINEIS, J. H., SOGIN, M. L. & MCLELLAN, S. L. 2015. Comparison of sewage and animal fecal microbiomes by using oligotyping reveals potential human fecal indicators in multiple taxonomic groups. *Applied and environmental microbiology,* 81**,** 7023-7033.

FJALSTAD, J. W., ESAIASSEN, E., JUVET, L. K., VAN DEN ANKER, J. N. & KLINGENBERG, C. 2018. Antibiotic therapy in neonates and impact on gut microbiota and antibiotic resistance development: a systematic review. *Journal of Antimicrobial Chemotherapy,* 73**,** 569-580.

FORSBERG, K. J., PATEL, S., GIBSON, M. K., LAUBER, C. L., KNIGHT, R., FIERER, N. & DANTAS, G. 2014. Bacterial phylogeny structures soil resistomes across habitats. *Nature,* 509**,** 612-616.

FRANCO-URÍA, A., LÓPEZ-MATEO, C., ROCA, E. & FERNÁNDEZ-MARCOS, M. L. 2009. Source identification of heavy metals in pastureland by multivariate analysis in NW Spain. *Journal of hazardous materials,* 165**,** 1008-1015.

FREIRE-MORAN, L., ARONSSON, B., MANZ, C., GYSSENS, I. C., SO, A. D., MONNET, D. L., CARS, O. & GROUP, E.-E. W. 2011. Critical shortage of new antibiotics in development against multidrug-resistant bacteria—Time to react is now. *Drug resistance updates,* 14**,** 118-124.

FREITAG, C., MICHAEL, G. B., KADLEC, K., HASSEL, M. & SCHWARZ, S. 2017. Detection of plasmid-borne extended-spectrum β-lactamase (ESBL) genes in Escherichia coli isolates from bovine mastitis. *Veterinary microbiology,* 200**,** 151-156.

FROST, L. S., LEPLAE, R., SUMMERS, A. O. & TOUSSAINT, A. 2005. Mobile genetic elements: the agents of open source evolution. *Nature Reviews Microbiology,* 3**,** 722-732.
FTP. 2020. *Age of animals slaughtered* [Online]. Farm Transparency Project. Available: http://www.aussieabattoirs.com/facts/age-slaughtered [Accessed 2020].
FUJIMURA, S., KAWAMURA, T., KATO, S., TATENO, H. & WATANABE, A. 2002. Detection of Helicobacter pylori in cow's milk. *Letters in Applied Microbiology,* 35**,** 504-507.
FURET, J.-P., FIRMESSE, O., GOURMELON, M., BRIDONNEAU, C., TAP, J., MONDOT, S., DORÉ, J. & CORTHIER, G. 2009. Comparative assessment of human and farm animal faecal microbiota using real-time quantitative PCR. *FEMS microbiology ecology,* 68**,** 351-362.
GARCÍA, N., GUTIÉRREZ, G., LORENZO, M., GARCÍA, J. E., PÍRIZ, S. & QUESADA, A. 2008. Genetic determinants for cfxA expression in Bacteroides strains isolated from human infections. *Journal of antimicrobial chemotherapy,* 62**,** 942-947.
GARCÍA, P. A. & PARGA-LANDA, B. 2021. An improved proposal for using laminar copper as a biocidal material in touch surfaces in a hospital Intensive Care Unit (ICU). *Environmental Science and Pollution Research***,** 1-9.
GARDNER, P. P., WATSON, R. J., MORGAN, X. C., DRAPER, J. L., FINN, R. D., MORALES, S. E. & STOTT, M. B. 2019. Identifying accurate metagenome and amplicon software via a meta-analysis of sequence to taxonomy benchmarking studies. *PeerJ,* 7**,** e6160.
GHOSH, S. & LAPARA, T. M. 2007. The effects of subtherapeutic antibiotic use in farm animals on the proliferation and persistence of antibiotic resistance among soil bacteria. *The ISME journal,* 1**,** 191-203.
GIACOMETTI, F., LUCCHI, A., DI FRANCESCO, A., DELOGU, M., GRILLI, E., GUARNIERO, I., STANCAMPIANO, L., MANFREDA, G., MERIALDI, G. & SERRAINO, A. 2015. Arcobacter butzleri, Arcobacter cryaerophilus, and Arcobacter skirrowii circulation in a dairy farm and sources of milk contamination. *Applied and environmental microbiology,* 81**,** 5055-5063.
GILLER, K. E., WITTER, E. & MCGRATH, S. P. 2009. Heavy metals and soil microbes. *Soil Biology and Biochemistry,* 41**,** 2031-2037.
GILLINGS, M. R. 2014. Integrons: past, present, and future. *Microbiology and Molecular Biology Reviews,* 78**,** 257-277.
GLOOR, G. B., MACKLAIM, J. M., PAWLOWSKY-GLAHN, V. & EGOZCUE, J. J. 2017. Microbiome datasets are compositional: and this is not optional. *Frontiers in microbiology,* 8**,** 2224.
GOLDSTEIN, B. P. 2014. Resistance to rifampicin: a review. *The Journal of antibiotics,* 67**,** 625-630.
GÖPFERT, C., PFANNSCHMIDT, L., GÖPFERT, J. P. & HAMMER, B. 2018. Interpretation of linear classifiers by means of feature relevance bounds. *Neurocomputing,* 298**,** 69-79.
GÓRECKI, K. & MCEVOY, M. M. 2020. Phylogenetic analysis reveals an ancient gene duplication as the origin of the MdtABC efflux pump. *PloS one,* 15**,** e0228877.
GOU, M., HU, H.-W., ZHANG, Y.-J., WANG, J.-T., HAYDEN, H., TANG, Y.-Q. & HE, J.-Z. 2018. Aerobic composting reduces antibiotic resistance genes in cattle manure and the resistome dissemination in agricultural soils. *Science of the Total Environment,* 612**,** 1300-1310.
GOV.UK. 2016. *Handling of manure and slurry to reduce antibiotic resistance* [Online]. Available: https://www.gov.uk/guidance/handling-of-manure-and-slurry-to-reduce-antibiotic-resistance [Accessed].
GOV.UK 2019. Contained and controlled The UK's 20-year vision for antimicrobial resistance.
GOV.UK. 2020a. *Nitrate vulnerable zones* [Online]. Available: https://www.gov.uk/government/collections/nitrate-vulnerable-zones [Accessed].
GOV.UK. 2020b. *Using nitrogen fertilisers in nitrate vulnerable zones* [Online]. Available: https://www.gov.uk/guidance/using-nitrogen-fertilisers-in-nitrate-vulnerable-zones#how-much-organic-manure-you-can-use-farm-and-field-limits [Accessed 2020].
GOW, S. P., WALDNER, C. L., HAREL, J. & BOERLIN, P. 2008. Associations between antimicrobial resistance genes in fecal generic Escherichia coli isolates from cow-calf herds in western Canada. *Appl. Environ. Microbiol.,* 74**,** 3658-3666.

GRAHAM, D. W., KNAPP, C. W., CHRISTENSEN, B. T., MCCLUSKEY, S. & DOLFING, J. 2016. Appearance of β-lactam Resistance Genes in Agricultural Soils and Clinical Isolates over the 20 th Century. *Scientific reports,* 6**,** 21550.

GRAVEY, F., GALOPIN, S., GRALL, N., AUZOU, M., ANDREMONT, A., LECLERCQ, R. & CATTOIR, V. 2013. Lincosamide resistance mediated by lnu (C)(L phenotype) in a Streptococcus anginosus clinical isolate. *Journal of Antimicrobial Chemotherapy,* 68**,** 2464-2467.

GRIFFITHS, B. E., DAI WHITE, G. & OIKONOMOU, G. 2018. A cross-sectional study into the prevalence of dairy cattle lameness and associated herd-level risk factors in England and Wales. *Frontiers in veterinary science,* 5**,** 65.

GROS, M., MARTI, E., BALCÁZAR, J. L., BOY-ROURA, M., BUSQUETS, A., COLON, J., SANCHEZ-MELSIO, A., LEKUNBERRI, I., BORREGO, C. M. & PONSÁ, S. 2019. Fate of pharmaceuticals and antibiotic resistance genes in a full-scale on-farm livestock waste treatment plant. *Journal of hazardous materials,* 378**,** 120716.

GUDETA, D. D., MOODLEY, A., BORTOLAIA, V. & GUARDABASSI, L. 2014. vanO, a new glycopeptide resistance operon in environmental Rhodococcus equi isolates. *Antimicrobial agents and chemotherapy,* 58**,** 1768-1770.

GULLBERG, E., CAO, S., BERG, O. G., ILBÄCK, C., SANDEGREN, L., HUGHES, D. & ANDERSSON, D. I. 2011. Selection of resistant bacteria at very low antibiotic concentrations. *PLoS Pathog,* 7**,** e1002158.

GURON, G. K., ARANGO-ARGOTY, G., ZHANG, L., PRUDEN, A. & PONDER, M. A. 2019. Effects of Dairy Manure-Based Amendments and Soil Texture on Lettuce-and Radish-Associated Microbiota and Resistomes. *Msphere,* 4.

HABTEWOLD, J., GORDON, R., SOKOLOV, V., VANDERZAAG, A., WAGNER-RIDDLE, C. & DUNFIELD, K. 2018. Targeting bacteria and methanogens to understand the role of residual slurry as an inoculant in stored liquid dairy manure. *Applied and environmental microbiology,* 84.

HAENNI, M., SARAS, E., BERTIN, S., LEBLOND, P., MADEC, J.-Y. & PAYOT, S. 2010. Diversity and mobility of integrative and conjugative elements in bovine isolates of Streptococcus agalactiae, S. dysgalactiae subsp. dysgalactiae, and S. uberis. *Applied and environmental microbiology,* 76**,** 7957-7965.

HAHNKE, S., LANGER, T., KOECK, D. E. & KLOCKE, M. 2016. Description of Proteiniphilum saccharofermentans sp. nov., Petrimonas mucosa sp. nov. and Fermentimonas caenicola gen. nov., sp. nov., isolated from mesophilic laboratory-scale biogas reactors, and emended description of the genus Proteiniphilum. *International journal of systematic and evolutionary microbiology,* 66**,** 1466-1475.

HAMMESFAHR, U., BIERL, R. & THIELE-BRUHN, S. 2011. Combined effects of the antibiotic sulfadiazine and liquid manure on the soil microbial-community structure and functions. *Journal of Plant Nutrition and Soil Science,* 174**,** 614-623.

HAN, X.-M., HU, H.-W., CHEN, Q.-L., YANG, L.-Y., LI, H.-L., ZHU, Y.-G., LI, X.-Z. & MA, Y.-B. 2018. Antibiotic resistance genes and associated bacterial communities in agricultural soils amended with different sources of animal manures. *Soil Biology and Biochemistry,* 126**,** 91-102.

HANCHI, H., MOTTAWEA, W., SEBEI, K. & HAMMAMI, R. 2018. The genus Enterococcus: between probiotic potential and safety concerns—an update. *Frontiers in microbiology,* 9**,** 1791.

HANDAL, T., GIRAUD-MORIN, C., CAUGANT, D. A., MADINIER, I., OLSEN, I. & FOSSE, T. 2005. Chromosome-and plasmid-encoded β-lactamases in Capnocytophaga spp. *Antimicrobial agents and chemotherapy,* 49**,** 3940-3943.

HANKS, J. & KOSSAIBATI, M. 2019. Key Performance Indicators for the UK national dairy herd.

HAO, X. & CHANG, C. 2002. Effect of 25 annual cattle manure applications on soluble and exchangeable cations in soil. *Soil science,* 167**,** 126-134.

HASKELL, M., LANGFORD, F., JACK, M., SHERWOOD, L., LAWRENCE, A. & RUTHERFORD, K. 2009. The effect of organic status and management practices on somatic cell counts on UK dairy farms. *Journal of dairy science,* 92**,** 3775-3780.

HE, Y., YUAN, Q., MATHIEU, J., STADLER, L., SENEHI, N., SUN, R. & ALVAREZ, P. J. 2020. Antibiotic resistance genes from livestock waste: Occurrence, dissemination, and treatment. *npj Clean Water,* 3**,** 1-11.

HEGSTAD, K., MIKALSEN, T., COQUE, T., WERNER, G. & 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**,** 541-554.

HERNANDO-AMADO, S., COQUE, T. M., BAQUERO, F. & MARTÍNEZ, J. L. 2019. Defining and combating antibiotic resistance from One Health and Global Health perspectives. *Nature Microbiology,* 4**,** 1432-1442.

HEUER, H., SOLEHATI, Q., ZIMMERLING, U., KLEINEIDAM, K., SCHLOTER, M., MÜLLER, T., FOCKS, A., THIELE-BRUHN, S. & SMALLA, K. 2011. Accumulation of sulfonamide resistance genes in arable soils due to repeated application of manure containing sulfadiazine. *Applied and Environmental Microbiology,* 77**,** 2527-2530.

HIGGINS, H., GOLDING, S., MOUNCEY, J., NANJIANI, I. & COOK, A. 2017. Understanding veterinarians' prescribing decisions on antibiotic dry cow therapy. *Journal of dairy science,* 100**,** 2909-2916.

HILL, A. E., GREEN, A. L., WAGNER, B. A. & DARGATZ, D. A. 2009. Relationship between herd size and annual prevalence of and primary antimicrobial treatments for common diseases on dairy operations in the United States. *Preventive veterinary medicine,* 88**,** 264-277.

HILLERTON, J. & BERRY, E. 2005. Treating mastitis in the cow–a tradition or an archaism. *Journal of applied microbiology,* 98**,** 1250-1255.

HIRANO, H. & TAKEMOTO, K. 2019. Difficulty in inferring microbial community structure based on co-occurrence network approaches. *BMC bioinformatics,* 20**,** 329.

HIRANSUTHIKUL, N., TANTISIRIWAT, W., LERTUTSAHAKUL, K., VIBHAGOOL, A. & BOONMA, P. 2005. Skin and soft-tissue infections among tsunami survivors in southern Thailand. *Clinical Infectious Diseases,* 41**,** e93-e96.

HO, A., DI LONARDO, D. P. & BODELIER, P. L. 2017. Revisiting life strategy concepts in environmental microbial ecology. *FEMS microbiology ecology,* 93**,** fix006.

HODGSON, C. J., OLIVER, D. M., FISH, R. D., BULMER, N. M., HEATHWAITE, A. L., WINTER, M. & CHADWICK, D. R. 2016. Seasonal persistence of faecal indicator organisms in soil following dairy slurry application to land by surface broadcasting and shallow injection. *Journal of environmental management,* 183**,** 325-332.

HOE, F. & RUEGG, P. L. 2006. Opinions and practices of Wisconsin dairy producers about biosecurity and animal well-being. *Journal of dairy science,* 89**,** 2297-2308.

HOELZER, K., WONG, N., THOMAS, J., TALKINGTON, K., JUNGMAN, E. & COUKELL, A. 2017. Antimicrobial drug use in food-producing animals and associated human health risks: what, and how strong, is the evidence? *BMC veterinary research,* 13**,** 1-38.

HOFMEYR, S., EGAN, R., GEORGANAS, E., COPELAND, A. C., RILEY, R., CLUM, A., ELOE-FADROSH, E., ROUX, S., GOLTSMAN, E. & BULUÇ, A. 2020. Terabase-scale metagenome coassembly with MetaHipMer. *Scientific reports,* 10**,** 1-11.

HOLMES, A. J., GILLINGS, M. R., NIELD, B. S., MABBUTT, B. C., NEVALAINEN, K. H. & STOKES, H. 2003. The gene cassette metagenome is a basic resource for bacterial genome evolution. *Environmental microbiology,* 5**,** 383-394.

HOLMES, S. & HUBER, W. 2018. *Modern statistics for modern biology*, Cambridge University Press.

HORMEÑO, L., UGARTE-RUIZ, M., PALOMO, G., BORGE, C., FLOREZ-CUADRADO, D., VADILLO, S., PÍRIZ, S., DOMÍNGUEZ, L., CAMPOS, M. J. & QUESADA, A. 2018. Ant (6)-I genes encoding aminoglycoside O-Nucleotidyltransferases are widely spread among streptomycin resistant strains of Campylobacter jejuni and Campylobacter coli. *Frontiers in microbiology,* 9**,** 2515.

HSIEH, T., MA, K. & CHAO, A. 2016. iNEXT: an R package for rarefaction and extrapolation of species diversity (H ill numbers). *Methods in Ecology and Evolution,* 7**,** 1451-1456.
HU, H.-W., HAN, X.-M., SHI, X.-Z., WANG, J.-T., HAN, L.-L., CHEN, D. & HE, J.-Z. 2016. Temporal changes of antibiotic-resistance genes and bacterial communities in two contrasting soils treated with cattle manure. *FEMS Microbiology Ecology,* 92.
HU, Y., YANG, X., QIN, J., LU, N., CHENG, G., WU, N., PAN, Y., LI, J., ZHU, L. & WANG, X. 2013. Metagenome-wide analysis of antibiotic resistance genes in a large cohort of human gut microbiota. *Nature communications,* 4**,** 1-7.
HU, Y. J. & COWLING, B. J. 2020. Reducing antibiotic use in livestock, China. *Bulletin of the World Health Organization,* 98**,** 360.
HUANG, L., XU, Y., XU, J., LING, J., ZHENG, L., ZHOU, X. & XIE, G. 2019. Dissemination of antibiotic resistance genes (ARGs) by rainfall on a cyclic economic breeding livestock farm. *International Biodeterioration & Biodegradation,* 138**,** 114-121.
HURST, J. J., OLIVER, J. P., SCHUELER, J., GOOCH, C., LANSING, S., CROSSETTE, E., WIGGINTON, K., RASKIN, L., AGA, D. S. & SASSOUBRE, L. M. 2019. Trends in antimicrobial resistance genes in manure blend pits and long-term storage across dairy farms with comparisons to antimicrobial usage and residual concentrations. *Environmental science & technology,* 53**,** 2405-2415.
HUTCHISON, M., WALTERS, L., MOORE, A., CROOKES, K. & AVERY, S. 2004. Effect of length of time before incorporation on survival of pathogenic bacteria present in livestock wastes applied to agricultural soil. *Applied and environmental microbiology,* 70**,** 5111-5118.
HYDE, R. M., REMNANT, J. G., BRADLEY, A. J., BREEN, J. E., HUDSON, C. D., DAVIES, P. L., CLARKE, T., CRITCHELL, Y., HYLANDS, M. & LINTON, E. 2017. Quantitative analysis of antimicrobial use on British dairy farms. *Veterinary Record***,** vetrec-2017-104614.
IBRAHIM, D. R., DODD, C. E., STEKEL, D. J., RAMSDEN, S. J. & HOBMAN, J. L. 2016. Multidrug resistant, extended spectrum β-lactamase (ESBL)-producing Escherichia coli isolated from a dairy farm. *FEMS microbiology ecology,* 92.
IM, S., PETERSEN, S. O., LEE, D. & KIM, D.-H. 2020. Effects of storage temperature on CH4 emissions from cattle manure and subsequent biogas production potential. *Waste Management,* 101**,** 35-43.
INGLIS, G. D., MCALLISTER, T. A., LARNEY, F. J. & TOPP, E. 2010. Prolonged survival of Campylobacter species in bovine manure compost. *Applied and Environmental Microbiology,* 76**,** 1110-1119.
ISHII, S. 2020. Quantification of antibiotic resistance genes for environmental monitoring: Current methods and future directions. *Current Opinion in Environmental Science & Health,* 16**,** 47-53.
ISIDRO, J., MENEZES, J., SERRANO, M., BORGES, V., PAIXÃO, P., MIMOSO, M., MARTINS, F., TOSCANO, C., SANTOS, A. & HENRIQUES, A. O. 2018. Genomic study of a Clostridium difficile multidrug resistant outbreak-related clone reveals novel determinants of resistance. *Frontiers in microbiology,* 9**,** 2994.
IWAHARA, K., KURIYAMA, T., SHIMURA, S., WILLIAMS, D. W., YANAGISAWA, M., NAKAGAWA, K. & KARASAWA, T. 2006. Detection of cfxA and cfxA2, the β-lactamase genes of Prevotella spp., in clinical samples from dentoalveolar infection by real-time PCR. *Journal of clinical microbiology,* 44**,** 172-176.
JAMIESON, R., GORDON, R., SHARPLES, K., STRATTON, G. & MADANI, A. 2002. Movement and persistence of fecal bacteria in agricultural soils and subsurface drainage water: A review. *Canadian biosystems engineering,* 44**,** 1-9.
JANDA, J. M. & ABBOTT, S. L. 2010. The genus Aeromonas: taxonomy, pathogenicity, and infection. *Clinical microbiology reviews,* 23**,** 35-73.
JANG, J., HUR, H. G., SADOWSKY, M. J., BYAPPANAHALLI, M., YAN, T. & ISHII, S. 2017. Environmental Escherichia coli: ecology and public health implications—a review. *Journal of applied microbiology,* 123**,** 570-581.

JANSSEN, P. H. 2006. Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. *Applied and environmental microbiology,* 72**,** 1719-1728.
JAYARAMAN, R. 2009. Antibiotic resistance: an overview of mechanisms and a paradigm shift. *Current science***,** 1475-1484.
JIANG, X., HALL, A. B., XAVIER, R. J. & ALM, E. J. 2019. Comprehensive analysis of chromosomal mobile genetic elements in the gut microbiome reveals phylum-level niche-adaptive gene pools. *PloS one,* 14**,** e0223680.
JIAO, Y.-N., ZHOU, Z.-C., CHEN, T., WEI, Y.-Y., ZHENG, J., GAO, R.-X. & CHEN, H. 2018. Biomarkers of antibiotic resistance genes during seasonal changes in wastewater treatment systems. *Environmental Pollution,* 234**,** 79-87.
JOHANSEN, A., CARTER, M. S., JENSEN, E. S., HAUGGARD-NIELSEN, H. & AMBUS, P. 2013. Effects of digestate from anaerobically digested cattle slurry and plant materials on soil microbial community and emission of CO2 and N2O. *Applied Soil Ecology,* 63**,** 36-44.
JOHNSON, C. M. & GROSSMAN, A. D. 2015. Integrative and conjugative elements (ICEs): what they do and how they work. *Annual review of genetics,* 49**,** 577-601.
JONES, P., MARIER, E., TRANTER, R., WU, G., WATSON, E. & TEALE, C. 2015. Factors affecting dairy farmers' attitudes towards antimicrobial medicine usage in cattle in England and Wales. *Preventive veterinary medicine,* 121**,** 30-40.
JOY, S. R., BARTELT-HUNT, S. L., SNOW, D. D., GILLEY, J. E., WOODBURY, B. L., PARKER, D. B., MARX, D. B. & LI, X. 2013. Fate and transport of antimicrobials and antimicrobial resistance genes in soil and runoff following land application of swine manure slurry. *Environmental science & technology,* 47**,** 12081-12088.
JOY, S. R., LI, X., SNOW, D. D., GILLEY, J. E., WOODBURY, B. & BARTELT-HUNT, S. L. 2014. Fate of antimicrobials and antimicrobial resistance genes in simulated swine manure storage. *Science of the Total Environment,* 481**,** 69-74.
JU, F., BECK, K., YIN, X., MACCAGNAN, A., MCARDELL, C. S., SINGER, H. P., JOHNSON, D. R., ZHANG, T. & BÜRGMANN, H. 2019. Wastewater treatment plant resistomes are shaped by bacterial composition, genetic exchange, and upregulated expression in the effluent microbiomes. *The ISME journal,* 13**,** 346-360.
JUHAS, M., VAN DER MEER, J. R., GAILLARD, M., HARDING, R. M., HOOD, D. W. & CROOK, D. W. 2009. Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS microbiology reviews,* 33**,** 376-393.
JURADO-RABADÁN, S., DE LA FUENTE, R., RUIZ-SANTA-QUITERIA, J. A., ORDEN, J. A., DE VRIES, L. E. & AGERSØ, Y. 2014. Detection and linkage to mobile genetic elements of tetracycline resistance gene tet (M) in Escherichia coli isolates from pigs. *BMC veterinary research,* 10**,** 155.
KAAKOUSH, N. O. 2015. Insights into the role of Erysipelotrichaceae in the human host. *Frontiers in cellular and infection microbiology,* 5**,** 84.
KANG, D. D., LI, F., KIRTON, E., THOMAS, A., EGAN, R., AN, H. & WANG, Z. 2019. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ,* 7**,** e7359.
KANTOR, R. S., MILLER, S. E. & NELSON, K. L. 2019. The Water microbiome through a pilot scale advanced treatment facility for direct potable reuse. *Frontiers in microbiology,* 10**,** 993.
KÄPPELI, N., MORACH, M., ZURFLUH, K., CORTI, S., NÜESCH-INDERBINEN, M. & STEPHAN, R. 2019. Sequence types and antimicrobial resistance profiles of Streptococcus uberis isolated from bovine mastitis. *Frontiers in veterinary science,* 6**,** 234.
KARAM, G., CHASTRE, J., WILCOX, M. H. & VINCENT, J.-L. 2016. Antibiotic strategies in the era of multidrug resistance. *Critical Care,* 20**,** 1-9.
KAYE, K. S., POGUE, J. M., TRAN, T. B., NATION, R. L. & LI, J. 2016. Agents of last resort: polymyxin resistance. *Infectious disease clinics of North America,* 30**,** 391-414.

KELLER, R., PEDROSO, M. Z., RITCHMANN, R. & SILVA, R. M. 1998. Occurrence of virulence-associated properties in Enterobacter cloacae. *Infection and immunity,* 66**,** 645-649.
KHAN, A. & MATHELIER, A. 2017. Intervene: a tool for intersection and visualization of multiple gene or genomic region sets. *BMC bioinformatics,* 18**,** 1-8.
KIM, K.-R., OWENS, G., KWON, S.-I., SO, K.-H., LEE, D.-B. & OK, Y. S. 2011. Occurrence and environmental fate of veterinary antibiotics in the terrestrial environment. *Water, Air, & Soil Pollution,* 214**,** 163-174.
KLÜMPER, U., RECKER, M., ZHANG, L., YIN, X., ZHANG, T., BUCKLING, A. & GAZE, W. H. 2019. Selection for antimicrobial resistance is reduced when embedded in a natural microbial community. *The ISME journal,* 13**,** 2927-2937.
KÖHLER, C.-D. & DOBRINDT, U. 2011. What defines extraintestinal pathogenic Escherichia coli? *International Journal of Medical Microbiology,* 301**,** 642-647.
KORAIMANN, G. 2018. Spread and Persistence of Virulence and Antibiotic Resistance Genes: A Ride on the F Plasmid Conjugation Module. *EcoSal Plus,* 8.
KRISHNASAMY, V., OTTE, J. & SILBERGELD, E. 2015. Antimicrobial use in Chinese swine and broiler poultry production. *Antimicrobial resistance and infection control,* 4**,** 1-9.
KRUEGER, F. 2012. *Trim Galore: a wrapper around Cutadapt and FastQC to consistently apply adapter and quality trimming to FastQ files, with extra functionality for RRBS data* [Online]. Available: https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/ [Accessed].
KUMAR, K., GUPTA, S. C., BAIDOO, S., CHANDER, Y. & ROSEN, C. J. 2005a. Antibiotic uptake by plants from soil fertilized with animal manure. *Journal of environmental quality,* 34**,** 2082-2085.
KUMAR, K., GUPTA, S. C., CHANDER, Y. & SINGH, A. K. 2005b. Antibiotic use in agriculture and its impact on the terrestrial environment. *Advances in agronomy,* 87**,** 1-54.
KURSA, M. B. & RUDNICKI, W. R. 2010. Feature selection with the Boruta package. *J Stat Softw,* 36**,** 1-13.
KURTZ, Z. D., MÜLLER, C. L., MIRALDI, E. R., LITTMAN, D. R., BLASER, M. J. & BONNEAU, R. A. 2015. Sparse and compositionally robust inference of microbial ecological networks. *PLoS computational biology,* 11.
KYSELKOVÁ, M., JIROUT, J., CHROŇÁKOVÁ, A., VRCHOTOVÁ, N., BRADLEY, R., SCHMITT, H. & ELHOTTOVÁ, D. 2013. Cow excrements enhance the occurrence of tetracycline resistance genes in soil regardless of their oxytetracycline content. *Chemosphere,* 93**,** 2413-2418.
KYSELKOVÁ, M., JIROUT, J., VRCHOTOVÁ, N., SCHMITT, H. & ELHOTTOVÁ, D. 2015a. Spread of tetracycline resistance genes at a conventional dairy farm. *Frontiers in microbiology,* 6**,** 536.
KYSELKOVÁ, M., KOTRBOVÁ, L., BHUMIBHAMON, G., CHROŇÁKOVÁ, A., JIROUT, J., VRCHOTOVÁ, N., SCHMITT, H. & ELHOTTOVÁ, D. 2015b. Tetracycline resistance genes persist in soil amended with cattle feces independently from chlortetracycline selection pressure. *Soil Biology and Biochemistry,* 81**,** 259-265.
LARSSON, D. J., ANDREMONT, A., BENGTSSON-PALME, J., BRANDT, K. K., DE RODA HUSMAN, A. M., FAGERSTEDT, P., FICK, J., FLACH, C.-F., GAZE, W. H. & KURODA, M. 2018. Critical knowledge gaps and research needs related to the environmental dimensions of antibiotic resistance. *Environment international,* 117**,** 132-138.
LAUBER, C. L., HAMADY, M., KNIGHT, R. & FIERER, N. 2009. Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Applied and environmental microbiology,* 75**,** 5111-5120.
LAXMINARAYAN, R., DUSE, A., WATTAL, C., ZAIDI, A. K., WERTHEIM, H. F., SUMPRADIT, N., VLIEGHE, E., HARA, G. L., GOULD, I. M. & GOOSSENS, H. 2013. Antibiotic resistance—the need for global solutions. *The Lancet infectious diseases,* 13**,** 1057-1098.
LEAR, G., DICKIE, I., BANKS, J., BOYER, S., BUCKLEY, H. L., BUCKLEY, T. R., CRUICKSHANK, R., DOPHEIDE, A., HANDLEY, K. M. & HERMANS, S. 2018. Methods for the extraction, storage, amplification and sequencing of DNA from environmental samples. *New Zealand Journal of Ecology,* 42**,** 10-50A.

LECLERCQ, R. 2002. Mechanisms of resistance to macrolides and lincosamides: nature of the resistance elements and their clinical implications. *Clinical Infectious Diseases,* 34**,** 482-492.
LECLERCQ, S. O., WANG, C., SUI, Z., WU, H., ZHU, B., DENG, Y. & FENG, J. 2016. A multiplayer game: species of Clostridium, Acinetobacter, and Pseudomonas are responsible for the persistence of antibiotic resistance genes in manure-treated soils. *Environmental microbiology,* 18**,** 3494-3508.
LEE, J.-Y., HAN, G. G., LEE, H.-B., LEE, S.-M., KANG, S.-K., JIN, G.-D., PARK, J., CHAE, B. J., CHOI, Y. H. & KIM, E. B. 2017. Prohibition of antibiotic growth promoters has affected the genomic profiles of Lactobacillus salivarius inhabiting the swine intestine. *Plos one,* 12**,** e0186671.
LEHTINEN, S., CHEWAPREECHA, C., LEES, J., HANAGE, W. P., LIPSITCH, M., CROUCHER, N. J., BENTLEY, S. D., TURNER, P., FRASER, C. & MOSTOWY, R. J. 2020. Horizontal gene transfer rate is not the primary determinant of observed antibiotic resistance frequencies in Streptococcus pneumoniae. *Science Advances,* 6**,** eaaz6137.
LEIGH, J. 1999. Streptococcus uberis: a permanent barrier to the control of bovine mastitis? *The Veterinary Journal,* 157**,** 225-238.
LEPLAE, R., LIMA-MENDEZ, G. & TOUSSAINT, A. 2010. ACLAME: a CLAssification of Mobile genetic Elements, update 2010. *Nucleic acids research,* 38**,** D57-D61.
LEVINGS, R. S., LIGHTFOOT, D., ELBOURNE, L. D., DJORDJEVIC, S. P. & HALL, R. M. 2006. New integron-associated gene cassette encoding a trimethoprim-resistant DfrB-type dihydrofolate reductase. *Antimicrobial agents and chemotherapy,* 50**,** 2863-2865.
LEVY, S. B. & MARSHALL, B. 2004. Antibacterial resistance worldwide: causes, challenges and responses. *Nature medicine,* 10**,** S122-S129.
LI, B., CHEN, Z., ZHANG, F., LIU, Y. & YAN, T. 2020. Abundance, diversity and mobility potential of antibiotic resistance genes in pristine Tibetan Plateau soil as revealed by soil metagenomics. *FEMS Microbiology Ecology,* 96**,** fiaa172.
LI, B., YANG, Y., MA, L., JU, F., GUO, F., TIEDJE, J. M. & ZHANG, T. 2015a. Metagenomic and network analysis reveal wide distribution and co-occurrence of environmental antibiotic resistance genes. *The ISME journal,* 9**,** 2490-2502.
LI, B., ZHAO, Z.-C., WANG, M.-H., HUANG, X.-H., PAN, Y.-H. & CAO, Y.-P. 2014a. Antimicrobial resistance and integrons of commensal Escherichia coli strains from healthy humans in China. *Journal of Chemotherapy,* 26**,** 190-192.
LI, D., LIU, C.-M., LUO, R., SADAKANE, K. & LAM, T.-W. 2015b. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics,* 31**,** 1674-1676.
LI, H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997*.
LI, H., HANDSAKER, B., WYSOKER, A., FENNELL, T., RUAN, J., HOMER, N., MARTH, G., ABECASIS, G. & DURBIN, R. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics,* 25**,** 2078-2079.
LI, J., RETTEDAL, E. A., VAN DER HELM, E., ELLABAAN, M., PANAGIOTOU, G. & SOMMER, M. O. 2019. Antibiotic treatment drives the diversification of the human gut resistome. *Genomics, proteomics & bioinformatics,* 17**,** 39-51.
LI, L.-Y., SHOEMAKER, N. B. & SALYERS, A. A. 1993. Characterization of the mobilization region of a Bacteroides insertion element (NBU1) that is excised and transferred by Bacteroides conjugative transposons. *Journal of bacteriology,* 175**,** 6588-6598.
LI, Y. F., CHEN, P. H. & YU, Z. 2014b. Spatial and temporal variations of microbial community in a mixed plug-flow loop reactor fed with dairy manure. *Microbial biotechnology,* 7**,** 332-346.
LIM, S.-K., KIM, D., MOON, D.-C., CHO, Y. & RHO, M. 2020. Antibiotic resistomes discovered in the gut microbiomes of Korean swine and cattle. *GigaScience,* 9**,** giaa043.
LIMA, T., DOMINGUES, S. & DA SILVA, G. J. 2020. Manure as a Potential Hotspot for Antibiotic Resistance Dissemination by Horizontal Gene Transfer Events. *Veterinary Sciences,* 7**,** 110.

LIN, H., CHAPMAN, S. J., FREITAG, T. E., KYLE, C., MA, J., YANG, Y. & ZHANG, Z. 2019. Fate of tetracycline and sulfonamide resistance genes in a grassland soil amended with different organic fertilizers. *Ecotoxicology and Environmental Safety,* 170**,** 39-46.

LIU, C., CHEN, Y., LI, X., ZHANG, Y., YE, J., HUANG, H. & ZHU, C. 2020. Temporal effects of repeated application of biogas slurry on soil antibiotic resistance genes and their potential bacterial hosts. *Environmental Pollution,* 258**,** 113652.

LIU, G., BOGAJ, K., BORTOLAIA, V., OLSEN, J. E. & THOMSEN, L. E. 2019a. Antibiotic-Induced, increased conjugative transfer is common to diverse naturally occurring ESBL plasmids in Escherichia coli. *Frontiers in microbiology,* 10**,** 2119.

LIU, J., LIU, M., WU, M., JIANG, C., CHEN, X., CAI, Z., WANG, B., ZHANG, J., ZHANG, T. & LI, Z. 2018. Soil pH rather than nutrients drive changes in microbial community following long-term fertilization in acidic Ultisols of southern China. *Journal of Soils and Sediments,* 18**,** 1853-1864.

LIU, J., TAFT, D. H., MALDONADO-GOMEZ, M. X., JOHNSON, D., TREIBER, M. L., LEMAY, D. G., DEPETERS, E. J. & MILLS, D. A. 2019b. The fecal resistome of dairy cattle is associated with diet during nursing. *Nature communications,* 10**,** 1-15.

LIU, P., JIA, S., HE, X., ZHANG, X. & YE, L. 2017. Different impacts of manure and chemical fertilizers on bacterial community structure and antibiotic resistance genes in arable soils. *Chemosphere,* 188**,** 455-464.

LIU, Y.-Y., WANG, Y., WALSH, T. R., YI, L.-X., ZHANG, R., SPENCER, J., DOI, Y., TIAN, G., DONG, B. & HUANG, X. 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**,** 161-168.

LOO, E. X. L., ZAIN, A., YAP, G. C., PURBOJATI, R. W., DRAUTZ-MOSES, D. I., KOH, Y. Q., CHONG, Y. S., TAN, K. H., GLUCKMAN, P. D. & YAP, F. 2020. Longitudinal assessment of antibiotic resistance gene profiles in gut microbiomes of infants at risk of eczema. *BMC infectious diseases,* 20**,** 1-12.

LOOFT, T., ALLEN, H. K., CASEY, T. A., ALT, D. P. & STANTON, T. B. 2014. Carbadox has both temporary and lasting effects on the swine gut microbiota. *Frontiers in microbiology,* 5**,** 276.

LOVELL, D., PAWLOWSKY-GLAHN, V., EGOZCUE, J. J., MARGUERAT, S. & BÄHLER, J. 2015. Proportionality: a valid alternative to correlation for relative data. *PLoS computational biology,* 11.

LOZUPONE, C. & KNIGHT, R. 2005. UniFrac: a new phylogenetic method for comparing microbial communities. *Applied and environmental microbiology,* 71**,** 8228-8235.

LOZUPONE, C. A., HAMADY, M., KELLEY, S. T. & KNIGHT, R. 2007. Quantitative and qualitative β diversity measures lead to different insights into factors that structure microbial communities. *Applied and environmental microbiology,* 73**,** 1576-1585.

LU, J., BREITWIESER, F. P., THIELEN, P. & SALZBERG, S. L. 2017. Bracken: estimating species abundance in metagenomics data. *PeerJ Computer Science,* 3**,** e104.

LU, Y., LI, J., MENG, J., ZHANG, J., ZHUANG, H., ZHENG, G., XIE, W., PING, L. & SHAN, S. 2020. Long-term biogas slurry application increased antibiotics accumulation and antibiotic resistance genes (ARGs) spread in agricultural soils with different properties. *Science of The Total Environment***,** 143473.

LUDVIGSEN, J., AMDAM, G. V., RUDI, K. & L'ABÉE-LUND, T. M. 2018. Detection and characterization of streptomycin resistance (strA-strB) in a honeybee gut symbiont (Snodgrassella alvi) and the associated risk of antibiotic resistance transfer. *Microbial ecology,* 76**,** 588-591.

MA, L., XIA, Y., LI, B., YANG, Y., LI, L.-G., TIEDJE, J. M. & ZHANG, T. 2016. Metagenomic assembly reveals hosts of antibiotic resistance genes and the shared resistome in pig, chicken, and human feces. *Environmental science & technology,* 50**,** 420-427.

MACEDO, G., HERNANDEZ-LEAL, L., VAN DER MAAS, P., HEEDERIK, D., MEVIUS, D. & SCHMITT, H. 2020. The impact of manure and soil texture on antimicrobial resistance gene levels in farmlands and adjacent ditches. *Science of The Total Environment*, 139563.

MADIGAN, M. T., MARTINKO, J. M., STAHL, D. A. & CLARK, D. P. 2006. *Brock biology of microorganisms, global edition*, Pearson Education Inc.

MADIGAN, M. T., MARTINKO, J. M., STAHL, D. A. & CLARK, D. P. 2012. *Brock biology of microorganisms*, Pearson Education, Inc.

MAEUSLI, M., LEE, B., MILLER, S., REYNA, Z., LU, P., YAN, J., ULHAQ, A., SKANDALIS, N., SPELLBERG, B. & LUNA, B. 2020. Horizontal Gene Transfer of Antibiotic Resistance from Acinetobacter baylyi to Escherichia coli on Lettuce and Subsequent Antibiotic Resistance Transmission to the Gut Microbiome. *Msphere,* 5.

MAGUIRE, F., JIA, B., GRAY, K., LAU, W. Y. V., BEIKO, R. G. & BRINKMAN, F. S. 2020. Metagenome-Assembled Genome Binning Methods with Short Reads Disproportionately Fail for Plasmids and Genomic Islands. *bioRxiv*.

MAHAJAN, G. B. & BALACHANDRAN, L. 2012. Antibacterial agents from actinomycetes-a review. *Frontiers in bioscience (Elite edition),* 4**,** 240-253.

MAHMOUDI, N., SLATER, G. F. & FULTHORPE, R. R. 2011. Comparison of commercial DNA extraction kits for isolation and purification of bacterial and eukaryotic DNA from PAH-contaminated soils. *Canadian Journal of Microbiology,* 57**,** 623-628.

MANAIA, C. M. 2017. Assessing the risk of antibiotic resistance transmission from the environment to humans: non-direct proportionality between abundance and risk. *Trends in microbiology,* 25**,** 173-181.

MARSHALL, B. M. & LEVY, S. B. 2011. Food animals and antimicrobials: impacts on human health. *Clinical microbiology reviews,* 24**,** 718-733.

MARTI, R., TIEN, Y.-C., MURRAY, R., SCOTT, A., SABOURIN, L. & TOPP, E. 2014. Safely coupling livestock and crop production systems: how rapidly do antibiotic resistance genes dissipate in soil following a commercial application of swine or dairy manure? *Applied and environmental microbiology,* 80**,** 3258-3265.

MARTIN, B. D., WITTEN, D. & WILLIS, A. D. 2020. Modeling microbial abundances and dysbiosis with beta-binomial regression. *Annals of Applied Statistics,* 14**,** 94-115.

MARTINEZ, J. L. 2009. The role of natural environments in the evolution of resistance traits in pathogenic bacteria. *Proceedings of the Royal Society B: Biological Sciences,* 276**,** 2521-2530.

MARTÍNEZ, J. L., COQUE, T. M. & BAQUERO, F. 2015. What is a resistance gene? Ranking risk in resistomes. *Nature Reviews Microbiology,* 13**,** 116-123.

MARTÍNEZ, N., MENDOZA, M. C., RODRÍGUEZ, I., SOTO, S., BANCES, M. & RODICIO, M. R. 2007. Detailed structure of integrons and transposons carried by large conjugative plasmids responsible for multidrug resistance in diverse genomic types of Salmonella enterica serovar Brandenburg. *Journal of antimicrobial chemotherapy,* 60**,** 1227-1234.

MCARTHUR, A. G. & TSANG, K. K. 2017. Antimicrobial resistance surveillance in the genomic age. *Annals of the New York Academy of Sciences,* 1388**,** 78-91.

MCEWEN, S. A. & COLLIGNON, P. J. 2018. Antimicrobial resistance: a one health perspective. *Antimicrobial resistance in bacteria from livestock and companion animals*, 521-547.

MCINTYRE, A. B., OUNIT, R., AFSHINNEKOO, E., PRILL, R. J., HÉNAFF, E., ALEXANDER, N., MINOT, S. S., DANKO, D., FOOX, J. & AHSANUDDIN, S. 2017. Comprehensive benchmarking and ensemble approaches for metagenomic classifiers. *Genome biology,* 18**,** 182.

MCKINNEY, C. W., DUNGAN, R. S., MOORE, A. & LEYTEM, A. B. 2018. Occurrence and abundance of antibiotic resistance genes in agricultural soil receiving dairy manure. *FEMS microbiology ecology,* 94**,** fiy010.

MEEK, R. W., VYAS, H. & PIDDOCK, L. J. V. 2015. Nonmedical uses of antibiotics: time to restrict their use? *PLoS Biol,* 13**,** e1002266.

MENCÍA-ARES, O., CABRERA-RUBIO, R., COBO-DÍAZ, J. F., ÁLVAREZ-ORDÓÑEZ, A., GÓMEZ-GARCÍA, M., PUENTE, H., COTTER, P. D., CRISPIE, F., CARVAJAL, A. & RUBIO, P. 2020. Antimicrobial use and production system shape the fecal, environmental, and slurry resistomes of pig farms. *Microbiome,* 8**,** 1-17.

MENZEL, P., NG, K. L. & KROGH, A. 2016. Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nature communications,* 7**,** 1-9.

MERINO, S., RUBIRES, X., KNOCHEL, S. & TOMÁS, J. M. 1995. Emerging pathogens: Aeromonas spp. *International journal of food microbiology,* 28**,** 157.

MIMA, T., KOHIRA, N., LI, Y., SEKIYA, H., OGAWA, W., KURODA, T. & TSUCHIYA, T. 2009. Gene cloning and characteristics of the RND-type multidrug efflux pump MuxABC-OpmB possessing two RND components in Pseudomonas aeruginosa. *Microbiology,* 155**,** 3509-3517.

MORRISON, K. E., JAŠAREVIĆ, E., HOWARD, C. D. & BALE, T. L. 2020. It's the fiber, not the fat: significant effects of dietary challenge on the gut microbiome. *Microbiome,* 8**,** 15.

MOYNIHAN, E. L., RICHARDS, K. G., BRENNAN, F. P., TYRREL, S. F. & RITZ, K. 2015. Enteropathogen survival in soil from different land-uses is predominantly regulated by microbial community composition. *Applied Soil Ecology,* 89**,** 76-84.

MUSOVIC, S., KLÜMPER, U., DECHESNE, A., MAGID, J. & SMETS, B. F. 2014. Long-term manure exposure increases soil bacterial community potential for plasmid uptake. *Environmental microbiology reports,* 6**,** 125-130.

MUURINEN, J., STEDTFELD, R., KARKMAN, A., PARNANEN, K., TIEDJE, J. & VIRTA, M. 2017. Influence of manure application on the environmental resistome under Finnish agricultural practice with restricted antibiotic use. *Environmental science & technology,* 51**,** 5989-5999.

MYROLD, D. D., ZEGLIN, L. H. & JANSSON, J. K. 2014. The potential of metagenomic approaches for understanding soil microbial processes. *Soil Science Society of America Journal,* 78**,** 3-10.

NAKAGAWA, S. & TAKAKI, Y. 2009. Nonpathogenic epsilonproteobacteria. *eLS*.

NAKATSU, C. H., HRISTOVA, K., HANADA, S., MENG, X.-Y., HANSON, J. R., SCOW, K. M. & KAMAGATA, Y. 2006. Methylibium petroleiphilum gen. nov., sp. nov., a novel methyl tert-butyl ether-degrading methylotroph of the Betaproteobacteria. *International journal of systematic and evolutionary microbiology,* 56**,** 983-989.

NATARAJAN, M. & ORIEL, P. 1992. Transfer of transposon Tn916 from Bacillus subtilis into a natural soil population. *Appl. Environ. Microbiol.,* 58**,** 2701-2703.

NAVON-VENEZIA, S., CHMELNITSKY, I., LEAVITT, A., SCHWABER, M. J., SCHWARTZ, D. & CARMELI, Y. 2006. Plasmid-mediated imipenem-hydrolyzing enzyme KPC-2 among multiple carbapenem-resistant Escherichia coli clones in Israel. *Antimicrobial agents and chemotherapy,* 50**,** 3098-3101.

NAYFACH, S., SHI, Z. J., SESHADRI, R., POLLARD, K. S. & KYRPIDES, N. C. 2019. New insights from uncultivated genomes of the global human gut microbiome. *Nature,* 568**,** 505-510.

NAYLOR, N. R., ATUN, R., ZHU, N., KULASABANATHAN, K., SILVA, S., CHATTERJEE, A., KNIGHT, G. M. & ROBOTHAM, J. V. 2018. Estimating the burden of antimicrobial resistance: a systematic literature review. *Antimicrobial Resistance & Infection Control,* 7**,** 1-17.

NEHER, T. P., MA, L., MOORMAN, T. B., HOWE, A. & SOUPIR, M. L. 2020. Seasonal variations in export of antibiotic resistance genes and bacteria in runoff from an agricultural watershed in Iowa. *Science of The Total Environment,* 738**,** 140224.

NEWELL, K. V., THOMAS, D. P., BREKASIS, D. & PAGET, M. S. 2006. The RNA polymerase-binding protein RbpA confers basal levels of rifampicin resistance on Streptomyces coelicolor. *Molecular microbiology,* 60**,** 687-696.

NFU 2017. NFU Consultation Response to Migration Advisory Committee Call For Evidence.

NGUYEN, H. T., WOLFF, K. A., CARTABUKE, R. H., OGWANG, S. & NGUYEN, L. 2010. A lipoprotein modulates activity of the MtrAB two-component system to provide intrinsic multidrug resistance, cytokinetic control and cell wall homeostasis in Mycobacterium. *Molecular microbiology,* 76**,** 348-364.

NGUYEN, L. H. & HOLMES, S. 2019. Ten quick tips for effective dimensionality reduction. *PLoS computational biology,* 15**,** e1006907.
NICHOLSON, F. A., SMITH, S. R., ALLOWAY, B., CARLTON-SMITH, C. & CHAMBERS, B. 2003. An inventory of heavy metals inputs to agricultural soils in England and Wales. *Science of the total environment,* 311**,** 205-219.
NIE, L., LV, Y., YUAN, M., HU, X., NIE, T., YANG, X., LI, G., PANG, J., ZHANG, J. & LI, C. 2014. Genetic basis of high level aminoglycoside resistance in Acinetobacter baumannii from Beijing, China. *Acta Pharmaceutica Sinica B,* 4**,** 295-300.
NÕLVAK, H., TRUU, M., KANGER, K., TAMPERE, M., ESPENBERG, M., LOIT, E., RAAVE, H. & TRUU, J. 2016. Inorganic and organic fertilizers impact the abundance and proportion of antibiotic resistance and integron-integrase genes in agricultural grassland soil. *Science of the Total Environment,* 562**,** 678-689.
NOYES, N. R., YANG, X., LINKE, L. M., MAGNUSON, R. J., COOK, S. R., ZAHEER, R., YANG, H., WOERNER, D. R., GEORNARAS, I. & MCART, J. A. 2016. Characterization of the resistome in manure, soil and wastewater from dairy and beef production systems. *Scientific reports,* 6**,** 1-12.
NURK, S., MELESHKO, D., KOROBEYNIKOV, A. & PEVZNER, P. A. 2017. metaSPAdes: a new versatile metagenomic assembler. *Genome research,* 27**,** 824-834.
O'NEILL, J. 2014. Antimicrobial Resistance: Tackling a crisis for the health and wealth of nations.
O'NEILL, J. 2016. Tackling drug-resistant infections globally: final report and recommendations.
OH, M., PRUDEN, A., CHEN, C., HEATH, L. S., XIA, K. & ZHANG, L. 2018. MetaCompare: a computational pipeline for prioritizing environmental resistome risk. *FEMS microbiology ecology,* 94**,** fiy079.
OKABE, S., OKAYAMA, N., SAVICHTCHEVA, O. & ITO, T. 2007. Quantification of host-specific Bacteroides–Prevotella 16S rRNA genetic markers for assessment of fecal pollution in freshwater. *Applied Microbiology and Biotechnology,* 74**,** 890-901.
OKADA, B. K. & SEYEDSAYAMDOST, M. R. 2017. Antibiotic dialogues: induction of silent biosynthetic gene clusters by exogenous small molecules. *FEMS microbiology reviews,* 41**,** 19-33.
OKUBO, T., AE, R., NODA, J., IIZUKA, Y., USUI, M. & TAMURA, Y. 2019. Detection of the sul2–strA–strB gene cluster in an ice core from Dome Fuji Station, East Antarctica. *Journal of global antimicrobial resistance,* 17**,** 72-78.
OLIVER, J. P., GOOCH, C. A., LANSING, S., SCHUELER, J., HURST, J. J., SASSOUBRE, L., CROSSETTE, E. M. & AGA, D. S. 2020a. Invited review: Fate of antibiotic residues, antibiotic-resistant bacteria, and antibiotic resistance genes in US dairy manure management systems. *Journal of Dairy Science,* 103**,** 1051-1071.
OLIVER, J. P., HURST, J. J., GOOCH, C. A., STAPPENBECK, A., SASSOUBRE, L. & AGA, D. S. 2020b. On-farm screw-press/rotary drum treatment of dairy manure associated antibiotic residues and resistance. Wiley Online Library.
OLIVER, S. P. & MURINDA, S. E. 2012. Antimicrobial resistance of mastitis pathogens. *Veterinary Clinics: Food Animal Practice,* 28**,** 165-185.
OLIVER, S. P., MURINDA, S. E. & JAYARAO, B. M. 2011. Impact of antibiotic use in adult dairy cows on antimicrobial resistance of veterinary and human pathogens: a comprehensive review. *Foodborne pathogens and disease,* 8**,** 337-355.
OLM, M. R., BROWN, C. T., BROOKS, B. & BANFIELD, J. F. 2017. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *The ISME journal,* 11**,** 2864-2868.
OSTRER, L., KHODURSKY, R. F., JOHNSON, J. R., HIASA, H. & KHODURSKY, A. 2019. Analysis of mutational patterns in quinolone resistance-determining regions of GyrA and ParC of clinical isolates. *International journal of antimicrobial agents,* 53**,** 318-324.

PAGE, K., WILSON, M. & PARKIN, I. P. 2009. Antimicrobial surfaces and their potential in reducing the role of the inanimate environment in the incidence of hospital-acquired infections. *Journal of materials chemistry,* 19**,** 3819-3831.
PAGET, M. S., MOLLE, V., COHEN, G., AHARONOWITZ, Y. & BUTTNER, M. J. 2001. Defining the disulphide stress response in Streptomyces coelicolor A3 (2): identification of the σR regulon. *Molecular microbiology,* 42**,** 1007-1020.
PAL, C., ASIANI, K., ARYA, S., RENSING, C., STEKEL, D. J., LARSSON, D. J. & HOBMAN, J. L. 2017. Metal resistance and its association with antibiotic resistance. *Advances in microbial physiology.* Elsevier.
PAL, C., BENGTSSON-PALME, J., KRISTIANSSON, E. & LARSSON, D. J. 2015. Co-occurrence of resistance genes to antibiotics, biocides and metals reveals novel insights into their co-selection potential. *BMC genomics,* 16**,** 964.
PANDEY, P., CHIU, C., MIAO, M., WANG, Y., SETTLES, M., DEL RIO, N. S., CASTILLO, A., SOUZA, A., PEREIRA, R. & JEANNOTTE, R. 2018. 16S rRNA analysis of diversity of manure microbial community in dairy farm environment. *PloS one,* 13.
PARHAM, J., DENG, S., DA, H., SUN, H. & RAUN, W. 2003. Long-term cattle manure application in soil. II. Effect on soil microbial populations and community structure. *Biology and Fertility of Soils,* 38**,** 209-215.
PARK, Y., PACHEPSKY, Y., SHELTON, D., JEONG, J. & WHELAN, G. 2016. Survival of manure-borne Escherichia coli and fecal coliforms in soil: temperature dependence as affected by site-specific factors. *Journal of environmental quality,* 45**,** 949-957.
PARK, Y. K., FOX, L. K., HANCOCK, D. D., MCMAHAN, W. & PARK, Y. H. 2012. Prevalence and antibiotic resistance of mastitis pathogens isolated from dairy herds transitioning to organic management. *Journal of Veterinary Science,* 13**,** 103.
PARKS, D. H., IMELFORT, M., SKENNERTON, C. T., HUGENHOLTZ, P. & TYSON, G. W. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome research,* 25**,** 1043-1055.
PARKS, D. H., RINKE, C., CHUVOCHINA, M., CHAUMEIL, P.-A., WOODCROFT, B. J., EVANS, P. N., HUGENHOLTZ, P. & TYSON, G. W. 2017. Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nature microbiology,* 2**,** 1533-1542.
PARTRIDGE, S. R., KWONG, S. M., FIRTH, N. & JENSEN, S. O. 2018. Mobile genetic elements associated with antimicrobial resistance. *Clinical microbiology reviews,* 31.
PARTRIDGE, S. R., TSAFNAT, G., COIERA, E. & IREDELL, J. R. 2009. Gene cassettes and cassette arrays in mobile resistance integrons. *FEMS microbiology reviews,* 33**,** 757-784.
PAUN, V. I., LAVIN, P., CHIFIRIUC, M. C. & PURCAREA, C. 2021. First report on antibiotic resistance and antimicrobial activity of bacterial isolates from 13,000-year old cave ice core. *Scientific reports,* 11**,** 1-15.
PENDLETON, J. N., GORMAN, S. P. & GILMORE, B. F. 2013. Clinical relevance of the ESKAPE pathogens. *Expert review of anti-infective therapy,* 11**,** 297-308.
PENG, S., FENG, Y., WANG, Y., GUO, X., CHU, H. & LIN, X. 2017. Prevalence of antibiotic resistance genes in soils after continually applied with different manure for 30 years. *Journal of hazardous materials,* 340**,** 16-25.
PENG, S., ZHOU, B., WANG, Y., LIN, X., WANG, H. & QIU, C. 2016. Bacteria play a more important role than nutrients in the accumulation of tetracycline resistance in manure-treated soil. *Biology and Fertility of Soils,* 52**,** 655-663.
PENG, Y., LEUNG, H. C., YIU, S.-M. & CHIN, F. Y. 2012. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics,* 28**,** 1420-1428.
PEOPLE'S REPUBLIC OF CHINA, M. O. A. A. R. A. 2017. National action plan to combat animal resources antimicrobial resistance (2017-2020).

PÉREZ-COBAS, A. E., GOMEZ-VALERO, L. & BUCHRIESER, C. 2020. Metagenomic approaches in microbial ecology: an update on whole-genome and marker gene sequencing analyses. *Microbial Genomics,* 6**,** e000409.

PÉREZ-VALERA, E., KYSELKOVÁ, M., AHMED, E., SLADECEK, F. X. J., GOBERNA, M. & ELHOTTOVÁ, D. 2019. Native soil microorganisms hinder the soil enrichment with antibiotic resistance genes following manure applications. *Scientific reports,* 9**,** 1-10.

PÉREZ-VARELA, M., CORRAL, J., ARANDA, J. & BARBÉ, J. 2018. Functional characterization of AbaQ, a novel efflux pump mediating quinolone resistance in Acinetobacter baumannii. *Antimicrobial agents and chemotherapy,* 62.

PERICHON, B. & COURVALIN, P. 2009. Antibiotic resistance. *In:* SCHAECHTER, M. (ed.) *The Desk Encyclopedia of Microbiology.* Elsevier.

PERRY, J., WAGLECHNER, N. & WRIGHT, G. 2016. The prehistory of antibiotic resistance. *Cold Spring Harbor perspectives in medicine,* 6**,** a025197.

PETERSON, E. & 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.

PETINAKI, E., GUÉRIN-FAUBLÉE, V., PICHEREAU, V., VILLERS, C., ACHARD, A., MALBRUNY, B. & LECLERCQ, R. 2008. Lincomycin resistance gene lnu (D) in Streptococcus uberis. *Antimicrobial agents and chemotherapy,* 52**,** 626-630.

PETROVICH, M. L., ZILBERMAN, A., KAPLAN, A., ELIRAZ, G. R., WANG, Y., LANGENFELD, K., DUHAIME, M., WIGGINTON, K., PORETSKY, R. & AVISAR, D. 2020. Microbial and Viral Communities and Their Antibiotic Resistance Genes Throughout a Hospital Wastewater Treatment System. *Frontiers in microbiology,* 11**,** 153.

PHILLIPS, I., CASEWELL, M., COX, T., DE GROOT, B., FRIIS, C., JONES, R., NIGHTINGALE, C., PRESTON, R. & WADDELL, J. 2004. Does the use of antibiotics in food animals pose a risk to human health? A critical review of published data. *Journal of antimicrobial Chemotherapy,* 53**,** 28-52.

PIDDOCK, L. J. 2006. Multidrug-resistance efflux pumps? not just for resistance. *Nature Reviews Microbiology,* 4**,** 629-636.

PIDDOCK, L. J., RICCI, V., STANLEY, K. & JONES, K. 2000. Activity of antibiotics used in human medicine for Campylobacter jejuni isolated from farm animals and their environment in Lancashire, UK. *Journal of antimicrobial chemotherapy,* 46**,** 303-306.

PINTO, T. C. A., COSTA, N. S., CORRÊA, A. B. D. A., OLIVEIRA, I. C. M. D., MATTOS, M. C. D., ROSADO, A. S. & BENCHETRIT, L. C. 2014. Conjugative transfer of resistance determinants among human and bovine Streptococcus agalactiae. *Brazilian Journal of Microbiology,* 45**,** 785-789.

PODMIRSEG, S. M., WALDHUBER, S., KNAPP, B. A., INSAM, H. & GOBERNA, M. 2019. Robustness of the autochthonous microbial soil community after amendment of cattle manure or its digestate. *Biology and Fertility of Soils,* 55**,** 565-576.

POETA, P., COSTA, D., RODRIGUES, J. & TORRES, C. 2006. Antimicrobial resistance and the mechanisms implicated in faecal enterococci from healthy humans, poultry and pets in Portugal. *International journal of antimicrobial agents,* 27**,** 131-137.

POIDEVIN, M., SATO, M., ALTINOGLU, I., DELAPLACE, M., SATO, C. & YAMAICHI, Y. 2018. Mutation in ESBL plasmid from Escherichia coli O104: H4 leads autoagglutination and enhanced plasmid dissemination. *Frontiers in microbiology,* 9**,** 130.

POIREL, L., NAAS, T. & NORDMANN, P. 2010. Diversity, epidemiology, and genetics of class D β-lactamases. *Antimicrobial agents and chemotherapy,* 54**,** 24-38.

POURCHER, A.-M., FRANÇOISE, P.-B., VIRGINIE, F., AGNIESZKA, G., VASILICA, S. & GÉRARD, M. 2007. Survival of faecal indicators and enteroviruses in soil after land-spreading of municipal sewage sludge. *Applied Soil Ecology,* 35**,** 473-479.

PREM, E. M., STRES, B., ILLMER, P. & WAGNER, A. O. 2020. Microbial community dynamics in mesophilic and thermophilic batch reactors under methanogenic, phenyl acid-forming conditions. *Biotechnology for Biofuels,* 13**,** 1-17.

PRESLEY, S. M., RAINWATER, T. R., AUSTIN, G. P., PLATT, S. G., ZAK, J. C., COBB, G. P., MARSLAND, E. J., TIAN, K., ZHANG, B. & ANDERSON, T. A. 2006. Assessment of pathogens and toxicants in New Orleans, LA following Hurricane Katrina. *Environmental Science & Technology,* 40**,** 468-474.

PROENÇA, J. T., BARRAL, D. C. & GORDO, I. 2017. Commensal-to-pathogen transition: One-single transposon insertion results in two pathoadaptive traits in Escherichia coli-macrophage interaction. *Scientific reports,* 7**,** 1-12.

QIAN, X., GU, J., SUN, W., WANG, X.-J., SU, J.-Q. & STEDFELD, R. 2018. Diversity, abundance, and persistence of antibiotic resistance genes in various types of animal manure following industrial composting. *Journal of Hazardous materials,* 344**,** 716-722.

QIAO, M., YING, G.-G., SINGER, A. C. & ZHU, Y.-G. 2018. Review of antibiotic resistance in China and its environment. *Environment international,* 110**,** 160-172.

QUEIROZ, O., OGUNADE, I., WEINBERG, Z. & ADESOGAN, A. 2018. Silage review: Foodborne pathogens in silage and their mitigation by silage additives. *Journal of dairy science,* 101**,** 4132-4142.

QUINCE, C., WALKER, A. W., SIMPSON, J. T., LOMAN, N. J. & SEGATA, N. 2017. Shotgun metagenomics, from sampling to analysis. *Nature biotechnology,* 35**,** 833.

R CORE TEAM 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

RAJEEV, L., SALYERS, A. A. & GARDNER, J. F. 2006. Characterization of the integrase of NBU1, a Bacteroides mobilizable transposon. *Molecular microbiology,* 61**,** 978-990.

RANKIN, D. J., ROCHA, E. P. & BROWN, S. P. 2011. What traits are carried on mobile genetic elements, and why? *Heredity,* 106**,** 1-10.

RASTOGI, G. & SANI, R. K. 2011. Molecular techniques to assess microbial community structure, function, and dynamics in the environment. *Microbes and microbial technology.* Springer.

RECORBET, G., STEINBERG, C. & FAURIE, G. 1992. Survival in soil of genetically engineered Escherichia coli as related to inoculum density, predation and competition. *FEMS Microbiology Letters,* 101**,** 251-260.

REESE, A. T. & DUNN, R. R. 2018. Drivers of microbiome biodiversity: a review of general rules, feces, and ignorance. *MBio,* 9.

RENDUELES, O., DE SOUSA, J. A. M., BERNHEIM, A., TOUCHON, M. & ROCHA, E. P. 2018. Genetic exchanges are more frequent in bacteria encoding capsules. *PLoS genetics,* 14**,** e1007862.

REYGAERT, W. C. 2018. An overview of the antimicrobial resistance mechanisms of bacteria. *AIMS microbiology,* 4**,** 482.

RIBER, L., POULSEN, P. H., AL-SOUD, W. A., SKOV HANSEN, L. B., BERGMARK, L., BREJNROD, A., NORMAN, A., HANSEN, L. H., MAGID, J. & SØRENSEN, S. J. 2014. Exploring the immediate and long-term impact on bacterial communities in soil amended with animal and urban organic waste fertilizers using pyrosequencing and screening for horizontal transfer of antibiotic resistance. *FEMS Microbiology Ecology,* 90**,** 206-224.

RICE, E. W., WANG, P., SMITH, A. L. & STADLER, L. B. 2020. Determining Hosts of Antibiotic Resistance Genes: A Review of Methodological Advances. *Environmental Science & Technology Letters,* 7**,** 282-291.

RICE, L. B. 2008. Federal funding for the study of antimicrobial resistance in nosocomial pathogens: no ESKAPE. The University of Chicago Press.

RIEKE, E. L., SOUPIR, M. L., MOORMAN, T. B., YANG, F. & HOWE, A. C. 2018. Temporal dynamics of bacterial communities in soil and leachate water after swine manure application. *Frontiers in microbiology,* 9**,** 3197.

ROBERTS, A. P. & MULLANY, P. 2011. Tn 916-like genetic elements: a diverse group of modular mobile elements conferring antibiotic resistance. *FEMS microbiology reviews,* 35**,** 856-871.
ROBERTS, M. C. 2005. Update on acquired tetracycline resistance genes. *FEMS microbiology letters,* 245**,** 195-203.
ROBERTS, M. C., SCHWARZ, S. & AARTS, H. J. 2012. Erratum: Acquired antibiotic resistance genes: an overview. *Frontiers in microbiology,* 3**,** 384.
ROESCH, L. F., FULTHORPE, R. R., RIVA, A., CASELLA, G., HADWIN, A. K., KENT, A. D., DAROUB, S. H., CAMARGO, F. A., FARMERIE, W. G. & TRIPLETT, E. W. 2007. Pyrosequencing enumerates and contrasts soil microbial diversity. *The ISME journal,* 1**,** 283-290.
ROTHROCK JR, M. J., KEEN, P. L., COOK, K. L., DURSO, L. M., FRANKLIN, A. M. & DUNGAN, R. S. 2016. How should we be determining background and baseline antibiotic resistance levels in agroecosystem research? *Journal of environmental quality,* 45**,** 420-431.
ROUSK, J., BÅÅTH, E., BROOKES, P. C., LAUBER, C. L., LOZUPONE, C., CAPORASO, J. G., KNIGHT, R. & FIERER, N. 2010. Soil bacterial and fungal communities across a pH gradient in an arable soil. *The ISME journal,* 4**,** 1340-1351.
ROVIRA, P., MCALLISTER, T., LAKIN, S. M., COOK, S. R., DOSTER, E., NOYES, N. R., WEINROTH, M. D., YANG, X., PARKER, J. K. & BOUCHER, C. 2019. Characterization of the microbial resistome in conventional and "raised without antibiotics" beef and dairy production systems. *Frontiers in microbiology,* 10**,** 1980.
RUDNICKI, W. R., WRZESIEŃ, M. & PAJA, W. 2015. All relevant feature selection methods and applications. *Feature Selection for Data and Pattern Recognition.* Springer.
RUFETE, B., PEREZ-MURCIA, M., PEREZ-ESPINOSA, A., MORAL, R., MORENO-CASELLES, J. & PAREDES, C. 2006. Total and faecal coliform bacteria persistence in a pig slurry amended soil. *Livestock Science,* 102**,** 211-215.
RUUSKANEN, M., MUURINEN, J., MEIERJOHAN, A., PÄRNÄNEN, K., TAMMINEN, M., LYRA, C., KRONBERG, L. & VIRTA, M. 2016. Fertilizing with animal manure disseminates antibiotic resistance genes to the farm environment. *Journal of environmental quality,* 45**,** 488-493.
SALDANHA, G. Z., PIRES, R. N., RAUBER, A. P., DE LIMA-MORALES, D., FALCI, D. R., CAIERÃO, J., PASQUALOTTO, A. C. & MARTINS, A. F. 2020. Genetic relatedness, Virulence factors and Antimicrobial Resistance of C. difficile strains from hospitalized patients in a multicentric study in Brazil. *Journal of Global Antimicrobial Resistance*.
SÁNCHEZ-OSUNA, M., CORTÉS, P., LLAGOSTERA, M., BARBÉ, J. & ERILL, I. 2020. Exploration into the origins and mobilization of di-hydrofolate reductase genes and the emergence of clinical resistance to trimethoprim. *Microbial Genomics***,** mgen000440.
SANDBERG, K. D. & LAPARA, T. M. 2016. The fate of antibiotic resistance genes and class 1 integrons following the application of swine and dairy manure to soils. *FEMS microbiology ecology,* 92.
SANDEGREN, L. 2014. Selection of antibiotic resistance at very low antibiotic concentrations. *Upsala journal of medical sciences,* 119**,** 103-107.
SANTOS, D. D. S., CALAÇA, P. R. D. A., PORTO, A. L. F., DE SOUZA, P. R. E., DE FREITAS, N. S. A. & CAVALCANTI VIEIRA SOARES, M. T. 2020. What Differentiates Probiotic from Pathogenic Bacteria? The Genetic Mobility of Enterococcus faecium Offers New Molecular Insights. *OMICS: A Journal of Integrative Biology,* 24**,** 706-713.
SAXENA, A., KUMARI, R., MUKHERJEE, U., SINGH, P. & LAL, R. 2014. Draft genome sequence of the rifamycin producer Amycolatopsis rifamycinica DSM 46095. *Genome Announcements,* 2.
SCHMITT, H., STOOB, K., HAMSCHER, G., SMIT, E. & SEINEN, W. 2006. Tetracyclines and tetracycline resistance in agricultural soils: microcosm and field studies. *Microbial ecology,* 51**,** 267-276.
SCOTT, L., MCGEE, P., WALSH, C., FANNING, S., SWEENEY, T., BLANCO, J., KARCZMARCZYK, M., EARLEY, B., LEONARD, N. & SHERIDAN, J. 2009. Detection of numerous verotoxigenic E. coli serotypes, with multiple antibiotic resistance from cattle faeces and soil. *Veterinary microbiology,* 134**,** 288-293.

SCZYRBA, A., HOFMANN, P., BELMANN, P., KOSLICKI, D., JANSSEN, S., DRÖGE, J., GREGOR, I., MAJDA, S., FIEDLER, J. & DAHMS, E. 2017. Critical assessment of metagenome interpretation—a benchmark of metagenomics software. *Nature methods,* 14**,** 1063-1071.

SEEMANN, T. 2020. Abricate. Github.

SEGATA, N., IZARD, J., WALDRON, L., GEVERS, D., MIROPOLSKY, L., GARRETT, W. S. & HUTTENHOWER, C. 2011. Metagenomic biomarker discovery and explanation. *Genome biology,* 12**,** R60.

SELWET, M. 2020. Influence of inoculation with Lactobacillus on fermentation, production of 1, 2-propanediol and 1-propanol as well as Maize silage aerobic stability. *Open Life Sciences,* 15**,** 373-378.

SENGELØV, G., AGERSØ, Y., HALLING-SØRENSEN, B., BALODA, S. B., ANDERSEN, J. S. & JENSEN, L. B. 2003. Bacterial antibiotic resistance levels in Danish farmland as a result of treatment with pig manure slurry. *Environment international,* 28**,** 587-595.

SHANKS, O. C., KELTY, C. A., ARCHIBEQUE, S., JENKINS, M., NEWTON, R. J., MCLELLAN, S. L., HUSE, S. M. & SOGIN, M. L. 2011. Community structures of fecal bacteria in cattle from different animal feeding operations. *Applied and environmental microbiology,* 77**,** 2992-3001.

SHARMA, M. & REYNNELLS, R. 2016. Importance of Soil Amendments: Survival of Bacterial Pathogens in Manure and Compost Used as Organic Fertilizers. *Microbiology Spectrum,* 4.

SHAW, J. & CLEWELL, D. 1985. Complete nucleotide sequence of macrolide-lincosamide-streptogramin B-resistance transposon Tn917 in Streptococcus faecalis. *Journal of Bacteriology,* 164**,** 782-796.

SHEN, Z., WANG, Y., SHEN, Y., SHEN, J. & WU, C. 2016. Early emergence of mcr-1 in Escherichia coli from food-producing animals. *The Lancet infectious diseases,* 16**,** 293.

SHENAGARI, M., BAKHTIARI, M., MOJTAHEDI, A. & ROUSHAN, Z. A. 2018. High frequency of mutations in gyrA gene associated with quinolones resistance in uropathogenic Escherichia coli isolates from the north of Iran. *Iranian journal of basic medical sciences,* 21**,** 1226.

SHOEMAKER, N. B., WANG, G.-R. & SALYERS, A. A. 1996. The Bacteroides mobilizable insertion element, NBU1, integrates into the 3'end of a Leu-tRNA gene and has an integrase that is a member of the lambda integrase family. *Journal of Bacteriology,* 178**,** 3594-3600.

SIEGWALD, L., CABOCHE, S., EVEN, G., VISCOGLIOSI, E., AUDEBERT, C. & CHABÉ, M. 2019. The impact of bioinformatics pipelines on microbiota studies: does the analytical “microscope” affect the biological interpretation? *Microorganisms,* 7**,** 393.

SIMON, H. Y., SIDDLE, K. J., PARK, D. J. & SABETI, P. C. 2019. Benchmarking metagenomics tools for taxonomic classification. *Cell,* 178**,** 779-794.

SINGER, A. C., SHAW, H., RHODES, V. & HART, A. 2016. Review of antimicrobial resistance in the environment and its relevance to environmental regulators. *Frontiers in microbiology,* 7**,** 1728.

SINGER, R. S. & WILLIAMS-NGUYEN, J. 2014. Human health impacts of antibiotic use in agriculture: a push for improved causal inference. *Current Opinion in Microbiology,* 19**,** 1-8.

SINTON, L. W., BRAITHWAITE, R. R., HALL, C. H. & MACKENZIE, M. L. 2007. Survival of indicator and pathogenic bacteria in bovine feces on pasture. *Applied and environmental microbiology,* 73**,** 7917-7925.

SLIZOVSKIY, I. B., MUKHERJEE, K., DEAN, C. J., BOUCHER, C. & NOYES, N. R. 2020. Mobilization of antibiotic resistance: Are current approaches for colocalizing resistomes and mobilomes useful? *Frontiers in Microbiology,* 11**,** 1376.

SMAILL, F. M. & GRIVELL, R. M. 2014. Antibiotic prophylaxis versus no prophylaxis for preventing infection after cesarean section. *Cochrane Database of Systematic Reviews*.

SMITH, K. & WILLIAMS, A. 2016. Production and management of cattle manure in the UK and implications for land application practice. *Soil Use and Management,* 32**,** 73-82.

SMITH, R. & COAST, J. 2013. The true cost of antimicrobial resistance. *Bmj,* 346.

SMOOT, M. E., ONO, K., RUSCHEINSKI, J., WANG, P.-L. & IDEKER, T. 2011. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics,* 27**,** 431-432.
SOM, N. F., HEINE, D., HOLMES, N., KNOWLES, F., CHANDRA, G., SEIPKE, R. F., HOSKISSON, P. A., WILKINSON, B. & HUTCHINGS, M. I. 2017. The MtrAB two-component system controls antibiotic production in Streptomyces coelicolor A3 (2). *Microbiology,* 163**,** 1415.
SONG, J., RENSING, C., HOLM, P. E., VIRTA, M. & BRANDT, K. K. 2017. Comparison of metals and tetracycline as selective agents for development of tetracycline resistant bacterial communities in agricultural soil. *Environmental science & technology,* 51**,** 3040-3047.
SOULSBY, L. 2007. Antimicrobials and animal health: a fascinating nexus. *Journal of Antimicrobial Chemotherapy,* 60**,** i77-i78.
SPAIN, A. M., KRUMHOLZ, L. R. & ELSHAHED, M. S. 2009. Abundance, composition, diversity and novelty of soil Proteobacteria. *The ISME journal,* 3**,** 992-1000.
SPÍŽEK, J. & ŘEZANKA, T. 2017. Lincosamides: Chemical structure, biosynthesis, mechanism of action, resistance, and applications. *Biochemical pharmacology,* 133**,** 20-28.
SRINIVASAN, P. & SARMAH, A. K. 2014. Dissipation of sulfamethoxazole in pasture soils as affected by soil and environmental factors. *Science of the Total Environment,* 479**,** 284-291.
STAMPS, B. W. & SPEAR, J. R. 2020. Identification of Metagenome-Assembled Genomes Containing Antimicrobial Resistance Genes, Isolated from an Advanced Water Treatment Facility. *Microbiology Resource Announcements,* 9.
STOGIOS, P. J. & SAVCHENKO, A. 2020. Molecular mechanisms of vancomycin resistance. *Protein Science,* 29**,** 654-669.
SUI, Q., ZHANG, J., CHEN, M., TONG, J., WANG, R. & WEI, Y. 2016. Distribution of antibiotic resistance genes (ARGs) in anaerobic digestion and land application of swine wastewater. *Environmental Pollution,* 213**,** 751-759.
SUN, R., ZHANG, X.-X., GUO, X., WANG, D. & CHU, H. 2015. Bacterial diversity in soils subjected to long-term chemical fertilization can be more stably maintained with the addition of livestock manure than wheat straw. *Soil Biology and Biochemistry,* 88**,** 9-18.
SUNDIN, G. & BENDER, C. 1996. Dissemination of the strA-strB streptomycin-resistance genes among commensal and pathogenic bacteria from humans, animals, and plants. *Molecular ecology,* 5**,** 133-143.
TACCONELLI, E. & PEZZANI, M. D. 2019. Public health burden of antimicrobial resistance in Europe. *The Lancet Infectious Diseases,* 19**,** 4-6.
TAN, G., HU, M., LI, X., PAN, Z., LI, M., LI, L., ZHENG, Z. & YANG, M. 2021. Metagenomics reveals the diversity and taxonomy of antibiotic resistance genes in sufu bacterial communities. *Food Control,* 121**,** 107641.
TANGO, C. N., SEO, S.-S., KWON, M., LEE, D.-O., CHANG, H. K. & KIM, M. K. 2020. Taxonomic and Functional Differences in Cervical Microbiome Associated with Cervical Cancer Development. *Scientific reports,* 10**,** 1-10.
TASHO, R. P. & CHO, J. Y. 2016. Veterinary antibiotics in animal waste, its distribution in soil and uptake by plants: a review. *Science of the Total Environment,* 563**,** 366-376.
TEDIJANTO, C., OLESEN, S. W., GRAD, Y. H. & LIPSITCH, M. 2018. Estimating the proportion of bystander selection for antibiotic resistance among potentially pathogenic bacterial flora. *Proceedings of the National Academy of Sciences,* 115**,** E11988-E11995.
TEILLANT, A., GANDRA, S., BARTER, D., MORGAN, D. J. & LAXMINARAYAN, R. 2015. Potential burden of antibiotic resistance on surgery and cancer chemotherapy antibiotic prophylaxis in the USA: a literature review and modelling study. *The Lancet infectious diseases,* 15**,** 1429-1437.
TENSON, T., LOVMAR, M. & EHRENBERG, M. 2003. The mechanism of action of macrolides, lincosamides and streptogramin B reveals the nascent peptide exit path in the ribosome. *Journal of molecular biology,* 330**,** 1005-1014.

THIELE-BRUHN, S., SEIBICKE, T., SCHULTEN, H. R. & LEINWEBER, P. 2004. Sorption of sulfonamide pharmaceutical antibiotics on whole soils and particle-size fractions. *Journal of Environmental Quality,* 33**,** 1331-1342.

THOMAS, M., WEBB, M., GHIMIRE, S., BLAIR, A., OLSON, K., FENSKE, G. J., FONDER, A. T., CHRISTOPHER-HENNINGS, J., BRAKE, D. & SCARIA, J. 2017. Metagenomic characterization of the effect of feed additives on the gut microbiome and antibiotic resistome of feedlot cattle. *Scientific reports,* 7**,** 1-13.

TIEDJE, J. M., FANG, W., MANAIA, C. M., VIRTA, M., SHENG, H., LIPING, M., ZHANG, T. & EDWARD, T. 2019. Antibiotic resistance genes in the human-impacted environment: A One Health perspective. *Pedosphere,* 29**,** 273-282.

TOMAZETTO, G., HAHNKE, S., WIBBERG, D., PÜHLER, A., KLOCKE, M. & SCHLÜTER, A. 2018. Proteiniphilum saccharofermentans str. M3/6T isolated from a laboratory biogas reactor is versatile in polysaccharide and oligopeptide utilization as deduced from genome-based metabolic reconstructions. *Biotechnology Reports,* 18**,** e00254.

TONG, J., FANG, P., ZHANG, J., WEI, Y., SU, Y. & ZHANG, Y. 2019. Microbial community evolution and fate of antibiotic resistance genes during sludge treatment in two full-scale anaerobic digestion plants with thermal hydrolysis pretreatment. *Bioresource technology,* 288**,** 121575.

TOULOUSE, J. L., EDENS, T. J., ALEJALDRE, L., MANGES, A. R. & PELLETIER, J. N. 2017. Integron-associated DfrB4, a previously uncharacterized member of the trimethoprim-resistant dihydrofolate reductase B family, is a clinically identified emergent source of antibiotic resistance. *Antimicrobial agents and chemotherapy,* 61.

TOULOUSE, J. L., SHI, G., LEMAY-ST-DENIS, C., EBERT, M. C., DEON, D., GAGNON, M., RUEDIGER, E., SAINT-JACQUES, K., FORGE, D. & VANDEN EYNDE, J. J. 2020. Dual-target inhibitors of the folate pathway inhibit intrinsically trimethoprim-resistant DfrB dihydrofolate reductases. *ACS Medicinal Chemistry Letters,* 11**,** 2261-2267.

TOVO, A., MENZEL, P., KROGH, A., LAGOMARSINO, M. C. & SUWEIS, S. 2020. Taxonomic classification method for metagenomics based on core protein families with Core-Kaiju. *bioRxiv*.

TRIVEDI, P., ANDERSON, I. C. & SINGH, B. K. 2013. Microbial modulators of soil carbon storage: integrating genomic and metabolic knowledge for global prediction. *Trends in Microbiology,* 21**,** 641-651.

TYRRELL, C., BURGESS, C. M., BRENNAN, F. P. & WALSH, F. 2019. Antibiotic resistance in grass and soil. *Biochemical Society Transactions,* 47**,** 477-486.

UDIKOVIC-KOLIC, N., WICHMANN, F., BRODERICK, N. A. & HANDELSMAN, J. 2014. Bloom of resident antibiotic-resistant bacteria in soil following manure fertilization. *Proceedings of the National Academy of Sciences,* 111**,** 15202-15207.

UK-VARSS 2020. Veterinary Antibiotic Resistance and Sales Surveillance

Report (UK-VARSS 2019).

UNC, A. & GOSS, M. J. 2003. Movement of faecal bacteria through the vadose zone. *Water, Air, and Soil Pollution,* 149**,** 327-337.

UNC, A. & GOSS, M. J. 2004. Transport of bacteria from manure and protection of water resources. *Applied Soil Ecology,* 25**,** 1-18.

UNNO, T., JANG, J., HAN, D., KIM, J. H., SADOWSKY, M. J., KIM, O.-S., CHUN, J. & HUR, H.-G. 2010. Use of barcoded pyrosequencing and shared OTUs to determine sources of fecal bacteria in watersheds. *Environmental science & technology,* 44**,** 7777-7782.

VAN DER KOLK, J. H., ENDIMIANI, A., GRAUBNER, C., GERBER, V. & PERRETEN, V. 2019. Acinetobacter in veterinary medicine, with an emphasis on Acinetobacter baumannii. *Journal of global antimicrobial resistance,* 16**,** 59-71.

VAN DER WALT, A. J., VAN GOETHEM, M. W., RAMOND, J.-B., MAKHALANYANE, T. P., REVA, O. & COWAN, D. A. 2017. Assembling metagenomes, one community at a time. *BMC genomics,* 18**,** 1-13.

VAN ELSAS, J. D., SEMENOV, A. V., COSTA, R. & TREVORS, J. T. 2011. Survival of Escherichia coli in the environment: fundamental and public health aspects. *The ISME journal,* 5**,** 173-183.
VAN GOETHEM, M. W., PIERNEEF, R., BEZUIDT, O. K., VAN DE PEER, Y., COWAN, D. A. & MAKHALANYANE, T. P. 2018. A reservoir of 'historical'antibiotic resistance genes in remote pristine Antarctic soils. *Microbiome,* 6**,** 1-12.
VAN HOEK, A. H., MEVIUS, D., GUERRA, B., MULLANY, P., ROBERTS, A. P. & AARTS, H. J. 2011. Acquired antibiotic resistance genes: an overview. *Frontiers in microbiology,* 2**,** 203.
VAN LOOVEREN, M., GOOSSENS, H. & GROUP, A. S. 2004. Antimicrobial resistance of Acinetobacter spp. in Europe. *Clinical microbiology and infection,* 10**,** 684-704.
VAN OVERBEEK, L. S., WELLINGTON, E. M., EGAN, S., SMALLA, K., HEUER, H., COLLARD, J.-M., GUILLAUME, G., KARAGOUNI, A. D., NIKOLAKOPOULOU, T. L. & VAN ELSAS, J. D. 2002. Prevalence of streptomycin-resistance genes in bacterial populations in European habitats. *FEMS microbiology ecology,* 42**,** 277-288.
VAN VEEN, J. A., VAN OVERBEEK, L. S. & VAN ELSAS, J. D. 1997. Fate and activity of microorganisms introduced into soil. *Microbiology and Molecular Biology Reviews,* 61**,** 121-135.
VENTOLA, C. L. 2015. The antibiotic resistance crisis: part 1: causes and threats. *Pharmacy and therapeutics,* 40**,** 277.
VERMA, A. K. & CHATTERJI, D. 2014. Dual role of MsRbpA: transcription activation and rescue of transcription from the inhibitory effect of rifampicin. *Microbiology,* 160**,** 2018-2029.
VIA, L., CURCIC, R., MUDD, M., DHANDAYUTHAPANI, S., ULMER, R. & DERETIC, V. 1996. Elements of signal transduction in Mycobacterium tuberculosis: in vitro phosphorylation and in vivo expression of the response regulator MtrA. *Journal of Bacteriology,* 178**,** 3314-3321.
VISHNIVETSKAYA, T. A., LAYTON, A. C., LAU, M. C., CHAUHAN, A., CHENG, K. R., MEYERS, A. J., MURPHY, J. R., ROGERS, A. W., SAARUNYA, G. S. & WILLIAMS, D. E. 2014. Commercial DNA extraction kits impact observed microbial community composition in permafrost samples. *FEMS microbiology ecology,* 87**,** 217-230.
VON MEIJENFELDT, F. B., ARKHIPOVA, K., CAMBUY, D. D., COUTINHO, F. H. & DUTILH, B. E. 2019. Robust taxonomic classification of uncharted microbial sequences and bins with CAT and BAT. *Genome biology,* 20**,** 217.
VON WINTERSDORFF, C. J., PENDERS, J., VAN NIEKERK, J. M., MILLS, N. D., MAJUMDER, S., VAN ALPHEN, L. B., SAVELKOUL, P. H. & WOLFFS, P. F. 2016. Dissemination of antimicrobial resistance in microbial ecosystems through horizontal gene transfer. *Frontiers in microbiology,* 7**,** 173.
WAGLECHNER, N. & WRIGHT, G. D. 2017. Antibiotic resistance: it's bad, but why isn't it worse? *BMC biology,* 15**,** 1-8.
WALES, A. D. & DAVIES, R. H. 2015. Co-selection of resistance to antibiotics, biocides and heavy metals, and its relevance to foodborne pathogens. *Antibiotics,* 4**,** 567-604.
WALLACE, J. S., GARNER, E., PRUDEN, A. & AGA, D. S. 2018. Occurrence and transformation of veterinary antibiotics and antibiotic resistance genes in dairy manure treated by advanced anaerobic digestion and conventional treatment methods. *Environmental pollution,* 236**,** 764-772.
WALTHER-RASMUSSEN, J. & HØIBY, N. 2006. OXA-type carbapenemases. *Journal of Antimicrobial Chemotherapy,* 57**,** 373-383.
WANG, C.-Y., ZHOU, X., GUO, D., ZHAO, J.-H., YAN, L., FENG, G.-Z., GAO, Q., YU, H. & ZHAO, L.-P. 2019. Soil pH is the primary factor driving the distribution and function of microorganisms in farmland soils in northeastern China. *Annals of Microbiology***,** 1-13.
WANG, F., HAN, W., CHEN, S., DONG, W., QIAO, M., HU, C. & LIU, B. 2020a. Fifteen-Year Application of Manure and Chemical Fertilizers Differently Impacts Soil ARGs and Microbial Community Structure. *Frontiers in microbiology,* 11**,** 62.
WANG, L., MANKIN, K. & MARCHIN, G. 2004. Survival of fecal bacteria in dairy cow manure. *Transactions of the ASAE,* 47**,** 1239.

WANG, N., GUO, X., YAN, Z., WANG, W., CHEN, B., GE, F. & YE, B. 2016a. A comprehensive analysis on spread and distribution characteristic of antibiotic resistance genes in livestock farms of Southeastern China. *PloS one,* 11.

WANG, Y., JI, H. & GAO, C. 2016b. Differential responses of soil bacterial taxa to long-term P, N, and organic manure application. *Journal of soils and sediments,* 16**,** 1046-1058.

WANG, Z., WANG, Y., FUHRMAN, J. A., SUN, F. & ZHU, S. 2020b. Assessment of metagenomic assemblers based on hybrid reads of real and simulated metagenomic sequences. *Briefings in bioinformatics,* 21**,** 777-790.

WASEEM, H., JAMEEL, S., ALI, J., SALEEM UR REHMAN, H., TAUSEEF, I., FAROOQ, U., JAMAL, A. & ALI, M. I. 2019. Contributions and challenges of high throughput qPCR for determining antimicrobial resistance in the environment: a critical review. *Molecules,* 24**,** 163.

WATERS, J. L. & SALYERS, A. A. 2013. Regulation of CTnDOT conjugative transfer is a complex and highly coordinated series of events. *MBio,* 4.

WATTENBERG, M., VIÉGAS, F. & JOHNSON, I. 2016. *How to use t-SNE effectively* [Online]. Distill. Available: https://distill.pub/2016/misread-tsne/ [Accessed].

WATTS, J. L. 1988. Etiological agents of bovine mastitis. *Veterinary microbiology,* 16**,** 41-66.

WEBER, D. J. & RUTALA, W. A. 2013. Self-disinfecting surfaces: review of current methodologies and future prospects. *American journal of infection control,* 41**,** S31-S35.

WEI, Z.-Q., DU, X.-X., YU, Y.-S., SHEN, P., CHEN, Y.-G. & LI, L.-J. 2007. Plasmid-mediated KPC-2 in a Klebsiella pneumoniae isolate from China. *Antimicrobial agents and chemotherapy,* 51**,** 763-765.

WERNER, G., NEUMANN, B., WEBER, R., KRESKEN, M., WENDT, C., BENDER, J., BECKER, K., BORGMANN, S., DIEFENBACH, A. & HAMPRECHT, A. 2020. Thirty years of VRE in Germany–"expect the unexpected": The view from the National Reference Centre for Staphylococci and Enterococci. *Drug Resistance Updates***,** 100732.

WHITTLE, G., WHITEHEAD, T. R., HAMBURGER, N., SHOEMAKER, N. B., COTTA, M. A. & SALYERS, A. A. 2003. Identification of a new ribosomal protection type of tetracycline resistance gene, tet (36), from swine manure pits. *Appl. Environ. Microbiol.,* 69**,** 4151-4158.

WHO 2014. Antimicrobial Resistance Global Report on Surveillance, 2014.

WHO 2015. Global Action Plan on Antimicrobial Resistance.

WHO 2017a. Critically important antimicrobials for human medicine: ranking of antimicrobial agents for risk management of antimicrobial resistance due to non-human use.

WHO 2017b. Global priority list of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics.

WHO 2019. Antibacterial agents in clinical development an analysis of the antibacterial clinical development pipeline, 2019.

WHO. 2020. *Disability-adjusted life years (DALYs)* [Online]. Available: https://www.who.int/data/gho/indicator-metadata-registry/imr-details/158 [Accessed].

WICHMANN, F., UDIKOVIC-KOLIC, N., ANDREW, S. & HANDELSMAN, J. 2014. Diverse antibiotic resistance genes in dairy cow manure. *MBio,* 5**,** e01017-13.

WILLIAMS-NGUYEN, J., SALLACH, J. B., BARTELT-HUNT, S., BOXALL, A. B., DURSO, L. M., MCLAIN, J. E., SINGER, R. S., SNOW, D. D. & ZILLES, J. L. 2016. Antibiotics and antibiotic resistance in agroecosystems: state of the science. *Journal of environmental quality,* 45**,** 394-406.

WIND, L., KROMETIS, L.-A., HESSION, W. C. & PRUDEN, A. 2020. Cross-comparison of methods for quantifying antibiotic resistance in agricultural soils amended with dairy manure and compost. *Science of The Total Environment***,** 144321.

WOLTERS, B., DING, G.-C., KREUZIG, R. & SMALLA, K. 2016. Full-scale mesophilic biogas plants using manure as C-source: bacterial community shifts along the process cause changes in the abundance of resistance genes and mobile genetic elements. *FEMS Microbiology Ecology,* 92.

WOLTERS, B., JACQUIOD, S., SØRENSEN, S. J., WIDYASARI-MEHTA, A., BECH, T. B., KREUZIG, R. & SMALLA, K. 2018. Bulk soil and maize rhizosphere resistance genes, mobile genetic elements and microbial communities are differently impacted by organic and inorganic fertilization. *FEMS microbiology ecology,* 94**,** fiy027.

WRIGHT, G. D. 2019. Environmental and clinical antibiotic resistomes, same only different. *Current opinion in microbiology,* 51**,** 57-63.

WU, Z., LUO, Y., BAO, J., LUO, Y. & YU, Z. 2020. Additives affect the distribution of metabolic profile, microbial communities and antibiotic resistance genes in high-moisture sweet corn kernel silage. *Bioresource Technology,* 315**,** 123821.

XIE, W.-Y., YUAN, S.-T., XU, M.-G., YANG, X.-P., SHEN, Q.-R., ZHANG, W.-W., SU, J.-Q. & ZHAO, F.-J. 2018a. Long-term effects of manure and chemical fertilizers on soil antibiotic resistome. *Soil Biology and Biochemistry,* 122**,** 111-119.

XIE, W. Y., SHEN, Q. & ZHAO, F. 2018b. Antibiotics and antibiotic resistance from animal manures to soil: a review. *European journal of soil science,* 69**,** 181-195.

XING, J., SUN, S., WANG, H., BROOKES, P. C. & XU, J. 2020. Response of soil native microbial community to Escherichia coli O157: H7 invasion. *Environmental Pollution,* 261**,** 114225.

XIONG, W., WANG, M., DAI, J., SUN, Y. & ZENG, Z. 2018. Application of manure containing tetracyclines slowed down the dissipation of tet resistance genes and caused changes in the composition of soil bacteria. *Ecotoxicology and environmental safety,* 147**,** 455-460.

XU, Y., SESHADRI, B., BOLAN, N., SARKAR, B., OK, Y. S., ZHANG, W., RUMPEL, C., SPARKS, D., FARRELL, M. & HALL, T. 2019. Microbial functional diversity and carbon use feedback in soils as affected by heavy metals. *Environment international,* 125**,** 478-488.

YANG, F., HAN, B., GU, Y. & ZHANG, K. 2020. Swine liquid manure: a hotspot of mobile genetic elements and antibiotic resistance genes. *Scientific reports,* 10**,** 1-10.

YAO, L., LI, Y., LI, Z., SHEN, D., FENG, H., ZHOU, H. & WANG, M. 2020. Prevalence of fluoroquinolone, macrolide and sulfonamide-related resistance genes in landfills from East China, mainly driven by MGEs. *Ecotoxicology and environmental safety,* 190**,** 110131.

YIN, X., DENG, Y., MA, L., WANG, Y., CHAN, L. Y. & ZHANG, T. 2019. Exploration of the antibiotic resistome in a wastewater treatment plant by a nine-year longitudinal metagenomic study. *Environment international,* 133**,** 105270.

ZAHEER, R., LAKIN, S. M., POLO, R. O., COOK, S. R., LARNEY, F. J., MORLEY, P. S., BOOKER, C. W., HANNON, S. J., VAN DOMSELAAR, G. & READ, R. R. 2019. Comparative diversity of microbiomes and Resistomes in beef feedlots, downstream environments and urban sewage influent. *BMC microbiology,* 19**,** 197.

ZAHRT, T. C. & DERETIC, V. 2000. An essential two-component signal transduction system in Mycobacterium tuberculosis. *Journal of bacteriology,* 182**,** 3832-3838.

ZAIN, A., YAP, G., PURBOJATI, R., MOSES, D., SHEK, L., GOH, A., VAN BEVER, H., TEOH, O., SOH, J. & THOMAS, B. Dynamic Nature of the Gut Resistome Among Infants in Singapore: 623. Open Forum Infectious Diseases, 2018.

ZALESKI, K. J., JOSEPHSON, K. L., GERBA, C. P. & PEPPER, I. L. 2005. Potential regrowth and recolonization of salmonellae and indicators in biosolids and biosolid-amended soil. *Applied and Environmental Microbiology,* 71**,** 3701-3708.

ZHANG, G., LECLERCQ, S. O., TIAN, J., WANG, C., YAHARA, K., AI, G., LIU, S. & FENG, J. 2017a. A new subclass of intrinsic aminoglycoside nucleotidyltransferases, ANT (3")-II, is horizontally transferred among Acinetobacter spp. by homologous recombination. *PLoS genetics,* 13**,** e1006602.

ZHANG, J., CHEN, M., SUI, Q., WANG, R., TONG, J. & WEI, Y. 2016. Fate of antibiotic resistance genes and its drivers during anaerobic co-digestion of food waste and sewage sludge based on microwave pretreatment. *Bioresource technology,* 217**,** 28-36.

ZHANG, J., SUI, Q., TONG, J., ZHONG, H., WANG, Y., CHEN, M. & WEI, Y. 2018a. Soil types influence the fate of antibiotic-resistant bacteria and antibiotic resistance genes following the land application of sludge composts. *Environment international,* 118**,** 34-43.

ZHANG, K., XIN, R., ZHAO, Z., LI, W., WANG, Y., WANG, Q., NIU, Z. & ZHANG, Y. 2020. Mobile genetic elements are the Major driver of High antibiotic resistance genes abundance in the Upper reaches of huaihe River Basin. *Journal of Hazardous Materials,* 401**,** 123271.

ZHANG, R., EGGLESTON, K., ROTIMI, V. & ZECKHAUSER, R. J. 2006. Antibiotic resistance as a global threat: evidence from China, Kuwait and the United States. *Globalization and Health,* 2**,** 6.

ZHANG, X.-X., ZHANG, T. & FANG, H. H. 2009. Antibiotic resistance genes in water environment. *Applied microbiology and biotechnology,* 82**,** 397-414.

ZHANG, Y.-J., HU, H.-W., GOU, M., WANG, J.-T., CHEN, D. & HE, J.-Z. 2017b. Temporal succession of soil antibiotic resistance genes following application of swine, cattle and poultry manures spiked with or without antibiotics. *Environmental Pollution,* 231**,** 1621-1632.

ZHANG, Y., HAO, X., ALEXANDER, T. W., THOMAS, B. W., SHI, X. & LUPWAYI, N. Z. 2018b. Long-term and legacy effects of manure application on soil microbial community composition. *Biology and Fertility of Soils,* 54**,** 269-283.

ZHANG, Y., SHEN, H., HE, X., THOMAS, B. W., LUPWAYI, N. Z., HAO, X., THOMAS, M. C. & SHI, X. 2017c. Fertilization shapes bacterial community structure by alteration of soil pH. *Frontiers in Microbiology,* 8**,** 1325.

ZHAO, Q., WENDLANDT, S., LI, H., LI, J., WU, C., SHEN, J., SCHWARZ, S. & WANG, Y. 2014. Identification of the novel lincosamide resistance gene lnu (E) truncated by ISEnfa5-cfr-ISEnfa5 insertion in Streptococcus suis: de novo synthesis and confirmation of functional activity in Staphylococcus aureus. *Antimicrobial agents and chemotherapy,* 58**,** 1785-1788.

ZHAO, R., YU, K., ZHANG, J., ZHANG, G., HUANG, J., MA, L., DENG, C., LI, X. & LI, B. 2020. Deciphering the mobility and bacterial hosts of antibiotic resistance genes under antibiotic selection pressure by metagenomic assembly and binning approaches. *Water Research,* 186**,** 116318.

ZHAO, X., WANG, J., ZHU, L. & WANG, J. 2019. Field-based evidence for enrichment of antibiotic resistance genes and mobile genetic elements in manure-amended vegetable soils. *Science of the Total Environment,* 654**,** 906-913.

ZHOU, B., WANG, C., ZHAO, Q., WANG, Y., HUO, M., WANG, J. & WANG, S. 2016. Prevalence and dissemination of antibiotic resistance genes and coselection of heavy metals in Chinese dairy farms. *Journal of hazardous materials,* 320**,** 10-17.

ZHOU, Z., RASKIN, L. & ZILLES, J. L. 2010. Effects of swine manure on macrolide, lincosamide, and streptogramin B antimicrobial resistance in soils. *Appl. Environ. Microbiol.,* 76**,** 2218-2224.

ZHU, Y., ZHANG, P., ZHANG, J., WANG, J., LU, Y. & PANG, X. 2020. Impact on Multiple Antibiotic Pathways Reveals MtrA as a Master Regulator of Antibiotic Production in Streptomyces spp. and Potentially in Other Actinobacteria. *Applied and environmental microbiology,* 86.

ZIELIŃSKA, S., RADKOWSKI, P., BLENDOWSKA, A., LUDWIG-GAŁĘZOWSKA, A., ŁOŚ, J. M. & ŁOŚ, M. 2017. The choice of the DNA extraction method may influence the outcome of the soil microbial community structure analysis. *MicrobiologyOpen,* 6**,** e00453.

Appendices

Appendix 1 Summary of soil physiochemical data, SE = standard error.

Site		pH	Moisture content %	TOC mg kg^{-1}	TC mg kg^{-1}	IC mg kg^{-1}	TN mg kg^{-1}	C/N_ratio	B mg kg^{-1}	Na mg kg^{-1}	Mg mg kg^{-1}	P mg kg^{-1}	S mg kg^{-1}	K mg kg^{-1}	Ca mg kg^{-1}
Slurry-impacted	Average	7.2	20.5	117.145	179.530	62.388	97.096	2.471	1.106	66.839	564.621	0.933	24.703	329.496	4176.775
	SE	0.04	0.80	9.632	9.796	2.640	9.597	0.243	0.277	2.278	18.243	0.246	1.863	25.725	85.183
Untreated	Average	6.5	27.3	94.072	133.880	19.747	54.073	2.798	2.729	22.094	264.003	2.199	12.575	367.033	2261.297
	SE	0.18	1.44	20.765	28.457	11.738	7.338	0.573	0.835	5.625	36.248	1.251	1.662	48.550	290.835

Site			Al mg kg^{-1}	Mn mg kg^{-1}	Fe mg kg^{-1}	Co mg kg^{-1}	Ni mg kg^{-1}	Cu mg kg^{-1}	Zn mg kg^{-1}	As mg kg^{-1}	Se mg kg^{-1}	Mo mg kg^{-1}	Cd mg kg^{-1}	Cs mg kg^{-1}	Ba mg kg^{-1}
Slurry-impacted	Average	-	27.824	0.144	0.246	0.002	0.020	0.083	0.238	0.007	0.003	4.349	0.004	0.004	101.441
	SE	-	8.746	0.030	0.058	<0.001	0.004	0.008	0.040	0.002	<0.001	1.837	<0.001	0.001	3.006
Untreated	Average	-	58.906	0.708	0.641	0.002	0.071	0.084	1.112	0.014	0.003	0.006	0.012	0.011	48.917
	SE	-	21.867	0.213	0.209	0.001	0.031	0.016	0.225	0.004	<0.001	0.002	0.002	0.001	6.450

Appendix 2 List and description of key software used for bioinformatic processing and analyses. Versions given where available.

General Description	Purpose	Platform (environment type)	Package(s)/Program(s)	Key Settings
DeepARG (short-read alogrithm)	ARG annotation (reads)	command-line	DeepARG v1.0.1	minimum ARG identity= 80%
				minimum ARG probability = 0.8
				mimimum e-value for ARG alignments = 1e-10
				mimimum gene coverage = 1
				minimum 16S identity = 0.8
PCA	data exploration	R Studio	ggbiplot v0.55	data scaled and centered
NMDS	data exploration	R Studio	vegan v2.5-6	distance measure = Bray-Curtis
				k= 2
tSNE	data exploration	R Studio	Rtsne	perplexity = 15
				theta = 0.0
				iterations = 1500
Corncob	differential abundance testing	R Studio	Corncob v0.1.0.	controlled for dispesrion in tested factors
iNEXT	diversity estimation (Hill numbers for q0, q1 and q2)	R Studio	v2.0.20	calculate q0, q1 and q2
UniFrac distances	calculating weighted UniFrac	R Studio	v1.3	used GUniFrac() function
PMCMRplus	calculate pair-wise multiple comparisons	R Studio	v1.9.3	used gamesHowellTest() function
FSA	calculate pair-wise multiple comparisons	R Studio	V0.9.1	used dunnTest() funtion
PRIMER	beta-diversity testing	PRIMER v6	PERMANOVA+ addon	NA
Boruta	feature selection	R Studio	Boruta	iterations = 99
LEfSe	feature selection	Online portal [https://huttenhower.sph.harvard.edu/galaxy]	LEfSe	alpha value Kruskal-Wallis= 0.05
				LDA threshold = 2.0
Heatmap	clustered heatmap	R Studio	Custom heatmap.2 script, dendextend	distance measure = euclidean
				cluster method = Ward.D2
				data scaled for heatmap visualisation
SPIECEASI	network analysis (co-occurance)	R Studio	SpiecEasi v.1.0.7	method =mb
				lambda min ratio = 1e-1
				n lambda = 50
				iterations = 200
correlation matrix	co-occurance	R Studio	corrplot v.0.84	type = Spearman's Rank
				fdr < 0.05
Procrustes	taxon-ARG correspondence	R Studio	vegan v2.5-6	p < 0.05
Kaiju	Taxonomic classification	command-line	Kaiju v1.7.1	default settings
Megahit	Metagenome assembly	command-line	Megahit v1.1.3	presets = meta-large
				min-count = 2
				k-step = variable (described in text)
TrimGalore	Sequence quality control	command-line	TrimGalore v0.0.4	quality = 20

FastQC	Sequence quality control	command-line	FastQC v0.11.8	NA
ABRicate	Contig ARG annotation	command-line	ABRicate v1.0.1	identity = 80%
CoNet	network analysis (co-occurance)	standalone GUI	CoNet v1.1.1.beta	gene coverage cut-off = 60% row minimum occurrence = 10 shuffle rows (edgeScores) renormalise (edgeScores) fdr < 0.05 p-value merge = Simes Spearman's Rank R > 0.8 bootstraps = 100 (filter unstable edges)
Cytoscape	CoNet network visualisation	standalone GUI	Cytoscape v3.7.2	yFiles organic layout
EasyFig	Contig-gene visualisation	standalone GUI	EasyFig v2.2.5	NA
MetaCompare	Resistome risk scores	command-line	MetaCompare v1.0	default settings
CAT/BAT	Taxonomic classification (contigs)	command-line	CAT v5.0.4	top = 10 range = 9
BWA-mem	Sequence alignment	command-line	BWA-mem v0.7.17 (r1188)	default settings
MetaBAT2	Binning (contigs)	command-line	MetaBAT2 v2.12.1	deafult settings
CheckM	MAG quality control	command-line	CheckM v1.1.2	default settings
GTDB-Tk	Taxonomic classification (contigs)	command-line	GTDB-Tk v0.3.2	deault settings
Prodigal	Gene (ORF) annotation	command-line	prodigal v2.6.3.	default settings
dRep	MAG clustering/dereplication	command-line	dRep v2.5.4	default settings
SamTools	sequence processing (helper program)	command-line	SamTools v1.9	NA
Cluster	PAM clustering of contigs	R Studio	cluster v2.1.0	k= 80