COMPARATIVE GENOMICS OF ENTEROCOCCI AND CARBAPENEM RESISTANT BACTERIA ACROSS A ONE-HEALTH CONTINUUM

SANI-E-ZEHRA ZAIDI
M.Phil. Molecular Medicine, University of Karachi, 2018

A thesis submitted
in partial fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

BIOMOLECULAR SCIENCE

Department of Chemistry and Biochemistry
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

Thesis Examination Committee

Date of defense: May 29, 2023

Dr. Athanasios Zovoilis Thesis Co-Supervisor	Associate Professor	PhD, M.D
Dr. Tim A. McAllister Thesis Co-Supervisor Lethbridge Research and Development Centre, Agriculture Agri-food Canada	Research Scientist	PhD
Dr. Rahat Zaheer Thesis Examination Committee Member Lethbridge Research and Development Centre, Agriculture Agri-food Canada	Biologist	PhD
Dr. Chad Liang Thesis Examination Committee Member National Centre for Animal Diseases Lethbridge, Canadian Food Inspection Agency	Research Scientist	PhD
Dr. Kim Stanford Internal External Examiner	Research Associate	PhD
Dr. Denice Bay External Examiner Max Rady College of Medicine Medical Microbiology & Infectious Diseases University of Manitoba	Associate Professor	PhD
Dr. Steven C. Mosimann Chair, Thesis Examination Committee	Professor	PhD

Abstract

Antimicrobial resistance is considered a One Health problem, affecting human, livestock, and environmental health. The overuse of antimicrobials in human and livestock settings exerts selective pressure, resulting in the proliferation of antimicrobial-resistant bacterial strains. *Enterococci* and *Escherichia coli* exist ubiquitously in One Health sectors and are therefore used as indicator bacteria in antimicrobial resistance surveillance programs. These bacteria carry antimicrobial resistance genes on mobile genetic elements that may transfer to other bacterial species, most importantly to pathogenic bacteria.

In this thesis, we characterized *Enterococcus* species recovered from One Health sectors and carbapenem-resistant bacteria in Canada. Phenotypic characterization was conducted using the disc diffusion method according to standard CLSI protocols, and whole sequencing using short and long-read platforms was used for genotypic characterization. Additionally, whole-genome comparative genomics was used to investigate the genomic relatedness among isolates from One Health sectors and the effect of antimicrobial usage on antimicrobial resistance in each sector.

We identified *Enterococcus hirae* as the prevalent *Enterococcus* spp. in bovine and swine production systems, while *Enterococcus faecium* was predominantly found in poultry. In human clinical isolates, *Enterococcus faecium* and *Enterococcus faecalis* were predominantly identified. In antimicrobial resistance profiling of *Enterococcus* spp., we found that tetracycline (*tetL*, *tetM*) and macrolide resistance (*ermB*) genes exist universally in *E. faecium* and *E. faecalis* from all One Health sectors. However, these resistance genes were predominantly found in bovine-sourced *E. hirae* compared to human settings.

Across livestock sectors, the analysis showed that multidrug-resistant *E. faecium* and *E. faecalis* isolates were more commonly found in poultry than in swine and bovine cattle. The reconstruction of a core-genome-based phylogenetic tree of *E. hirae*, *E. faecium*, and *E. faecalis* did not reveal segregation based on their sample source. However, in *E. faecium*, the majority of clinical and municipal wastewater isolates were mapped to the hospital-associated clade.

The study on carbapenem-resistant bacteria showed that the prevalence of carbapenem-resistant bacteria in beef production is rare, even after culture enrichment with carbapenem. The majority of carbapenem-resistant species recovered from bovine feedlot samples carried intrinsic resistance genes against carbapenem. Moreover, in the comparative genomic analysis of *Pseudomonas aeruginosa*, bovine-sourced isolates were phylogenetically different from human clinical isolates.

Preface

The thesis consists of five chapters and three published articles. Chapter 2-4 were published in peer reviewed journals

Author's Contributions

Chapter 2 (Genomic Characterization of Enterococcus hirae From Beef Cattle Feedlots and Associated Environmental Continuum)

This chapter was published in the Frontiers in Microbiology in 2022. [citation: Zaidi, S. E., Zaheer, R., Barbieri, R., Cook, S. R., Hannon, S. J., Booker, C. W., Church, D., Van Domselaar, G., Zovoilis, A., & McAllister, T. A. (2022). Genomic Characterization of *Enterococcus hirae* From Beef Cattle Feedlots and Associated Environmental Continuum. Frontiers in microbiology, 13, 859990. https://doi.org/10.3389/fmicb.2022.859990]. This work was accomplished by myself and my colleagues. In this study, I did whole genome sequencing of recovered isolates followed by their genomic characterization, data analysis, overall data/results analysis. I also wrote manuscript and generated all figures for publication. Following are contributions of other authors. Dr. Sherry J Hannon (Feedlot Health Management Services, Okotoks, AB, Canada) and Shaun R. Cook (Lethbridge Research and Development Centre, Agriculture and Agri-Food Canada, Lethbridge, AB, Canada) collected feedlot samples, metadata, and antimicrobial use data. Ms. Ruth Barbieri (Lethbridge Research and Development Centre, Agriculture and Agri-Food Canada, Lethbridge, AB, Canada), Shaun R. Cook, Dr. Rahat Zaheer (Lethbridge Research and Development Centre, Agriculture and Agri-Food Canada, Lethbridge, AB, Canada) isolated and characterized enterococci. Dr. Gary Van Domselaar (National Microbiology

Laboratory, Public Health Agency of Canada, Winnipeg, MB, Canada) provided access and managed the bioinformatics cluster facility and bioinformatics tools for genomic and comparative genomic analysis. Dr. Tim A. McAllister and Dr. Rahat Zaheer guided the study with conceptualization, methodology and writing the manuscript. Dr. Tim A. McAllister and Dr. Athanasios Zovoilis provided funding and supervision.

Chapter 3 (Comparative Genomic Analysis of Enterococci across Sectors of the One Health Continuum)

This chapter was published in Microorganisms in 2023: [citation: Zaidi, S. E., Zaheer, R., Poulin-Laprade, D., Scott, A., Rehman, M. A., Diarra, M., Topp, E., Domselaar, G. V., Zovoilis, A., & McAllister, T. A. (2023). Comparative Genomic Analysis of Enterococci across Sectors of the One Health Continuum. Microorganisms, 11(3), 727. https://doi.org/10.3390/microorganisms11030727]. This work was a collaboration between multiple Agriculture Agri-food Canada centers. Dr. Dominic Poulin-Laprade from Sherbrooke Research and Development Centre, Agriculture and Agri-Food Canada, Sherbrooke, QC J1M 1Z3, Canada provided metadata and whole genome sequencing data for all swine isolates. Dr. Muhammad Attiq Rehman and Dr. Moussa Diarra from Guelph Research and Development Centre, Agriculture and Agri-Food Canada, Guelph, ON N1G 5C9, Canada provided us meta-data and whole genome sequencing data for all poultry isolates. Dr. Edward Topp and Mr. Andrew Scott from London Research and Development Centre, Agriculture and Agri-Food Canada, London, ON N5V 4T3, Canada provided us metadata and whole genome sequencing data for a subset of environmental isolates (from Ontario). Whole genome sequencing data and metadata of bovine, human clinical, municipal waste water and natural water sources were generated in our lab at Lethbridge

Research and Development Centre Agriculture and Agri-Food Canada. For bioinformatic analysis, Dr. Gary Van Domselaar from National Microbiology Laboratory, Public Health Agency of Canada, Winnipeg, MB, Canada provided us access to the bioinformatics cluster facility and bioinformatics tools. In this project, whole genome analysis and comparative genomic analysis was done by me. In addition, I wrote manuscript and generated figures for publication. Dr. Tim A. McAllister and Dr. Rahat Zaheer guided the study with conceptualization, methodology and writing a manuscript. Dr. Tim A. McAllister, Dr. Athanasios Zovoilis and Dr. Rahat Zaheer provided funding and supervision.

Chapter 4 (Genomic Characterization of Carbapenem-Resistant Bacteria from Beef Cattle Feedlots)

This manuscript is published in the Antibiotics journal in 2023; [citation: Zaidi S-e-Z, Zaheer R, Thomas K, Abeysekara S, Haight T, Saville L, Stuart-Edwards M, Zovoilis A, McAllister TA (2023). Genomic Characterization of Carbapenem-Resistant Bacteria from Beef Cattle Feedlots. Antibiotics 12(6):960. https://doi.org/10.3390/antibiotics12060960.] This study was a collaboration between myself and my colleagues. Ms. Krysty Munns did sample collection, enrichment and bacterial isolation. Sujeema Abeysekara performed sequencing and CarbaNP testing. I did antimicrobial susceptibility testing, genomic DNA extraction of recovered isolates. In addition, I also generated all figures, analyzed sequence data and overall data/results analysis and wrote the first draft of the manuscript. Dr. Tim A. McAllister and Dr. Rahat Zaheer guided the study with conceptualization, methodology and writing. Dr. Tim A. McAllister, Dr. Athanasios Zovoilis and Dr. Rahat Zaheer provided funding and supervision. All authors have read and agreed to the published of the manuscript.

All figures in this thesis were created by myself through Biorender application. We also hold publication licence for figures submitted in the thesis and manuscripts. Some supplementary tables are submitted as external files due to their large size.

Acknowledgements

I am grateful to Almighty God for bestowing upon me the strength, wisdom, aptitude and opportunity to pursue and achieve for my goals.

As a graduate student, my journey has been an awe-inspiring yet challenging. Throughout these years numerous individuals have played a vital role in shaping my academic trajectory I would not be able to thrive my endeavour without their unwavering support.

First and foremost, I would like to extend my heartfelt appreciation to my supervisors Dr. Tim A. McAllister and Dr. Athanasios Zovoilis for granting me this invaluable opportunity and acquainting me with the realm of research. Their guidance, infectious enthusiasm, and amiable demeanor have allowed me to relish every aspect of my work. I am genuinely grateful to Dr. Rahat Zaheer for exceptional mentorship. I am thankful for her patience and responding every 'quick call?' message. I deeply value the confidence she has placed in me.

I am grateful to my husband for his unwavering encouragement and patience throughout this journey. None of this would have been possible without his support. I would like to extend my gratitude to my family for their continuous support and love.

I would also like to acknowledge the Major Innovation Fund of the Government of Alberta, the University of Calgary AMR One Health Consortium, the Beef Cattle Research Council (BCRC) Project and the Genomics Research and Development Initiative (GRDI) of the Government of Canada for generous financial support. Additionally, I express my gratitude to the University of Lethbridge for providing me with financial support through their teaching assistantship program.

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

List of Tables

List of Abbreviations

AMR	Antimicrobial resistance
ARG	Antimicrobial resistance genes
AST	Antimicrobial susceptibility testing
ATCC	American Type Culture Collection
BEA	Bile esculin azide
BLAST	Basic local alignment search tool
Bp	Base pair
CCS	Circular consensus sequencing
CLSI	Clinical and Laboratory Standards Institute
CIPARS	Canadian Integrated Program for Antimicrobial Resistance Surveillance
COGs	Collections of Clusters of Orthologous Genes
CRISPR	Clustered regularly interspaced short palindromic repeats
CRL	Continuous long read
EASSA	European Antimicrobial Susceptibility Surveillance in Animals
ESBLs	Extended spectrum β-lactamases
EUCAST	European Committee on Antimicrobial Susceptibility Testing
FAO	Food and Agriculture Organization of the United Nations
Gb	Giga bytes
GBD	Global Burden of Diseases
3GCs	3rd generation cephalosporins
gDNA	Genomic DNA
GLASS	Global Antimicrobial Resistance and Use Surveillance System
HiFi	High-fidelity
HGT	Horizontal gene transfer
HMM	Hidden Markov models
JVARM	Japanese veterinary antimicrobial resistance monitoring systems
kb	Kilo bases
µg	Microgram
ml	Microliter
MALDI-TOF	matrix-assisted laser desorption ionization–time of flight mass spectrometry
MDR	Multidrug resistant
MGEs	Mobile genetic elements
MLST	Multilocus sequence typing
MVTR	multiple-locus variable-number tandem-repeat analysis
NARMS	National Antimicrobial Resistance Monitoring System for Enteric Bacteria
NGS	Next generation sequencing
OIE	World Organization for Animal Health
OLC	Overlap-layout-consensus
ONT	Oxford Nanopore Technologies
PCR	Polymerase chain reaction
RFLP	Restriction fragment length polymorphism
SBS	Sequence by synthesis
ST	Sequence type
SNV	Single nucleotide variant
VRE	Vancomycin resistant enterococci
WGS	Whole genome sequencing
WHO	World Health Organization

Chapter: 1 Introduction

1.1 Global burden of antimicrobial resistance

Antimicrobials are chemical compounds that can inhibit microbial growth, including bacteria, parasites, viruses, and fungi, by targeting important cellular mechanisms. Antimicrobial resistance is defined as the ability of microbes to resist the effects of drugs or antimicrobials. Currently, more than 150 antimicrobials are approved in human and veterinary medicine to treat all known bacterial infections (1). Bacteria are defined as 'multidrug-resistant' if they show resistance to three or more antimicrobial classes. Globally, multidrug-resistant bacterial infections claimed the lives of 1.27 million people in 2019, with a projected increase to 10 million per year by 2050.(2). According to the study, the death rate per 100,000 due to multidrug-resistant bacterial infections was highest in Sub-Saharan Africa (98.9), followed by South Asia (76.8), Central Asia and Central and Eastern Europe (67.7), Latin America and the Caribbean (57.9), Southern Latin America, North America, Western Europe, and Australia (55.7), Southeast Asia, East Asia, and Oceania (47.1), and North Africa and the Middle East (42.0) (3). These regions were categorized based on their epidemiological similarities and geographical proximity. These regions were categorized on the basis of their epidemiological similarities and geographical proximity (4).

1.2 How antimicrobial resistance occurs and spreads

Antimicrobial resistance (AMR) is not a new phenomenon and it is an expected result of the interaction of many organisms with their environment. Bacteria lived in varied environments and required different genes for survival. They can accumulate diverse

metabolic and protective mechanisms to survive unfavorable environmental conditions (5). Many antimicrobial compounds are naturally produced by organisms including fungi and bacteria. The co-resident bacteria have evolved mechanisms to resist the action of antimicrobial compounds for their survival. Thus, these bacteria are often considered to be 'intrinsically' or 'naturally' or 'inherently' resistant to one or more antimicrobial compounds. Mechanisms of intrinsic resistance include lack of outer-membrane permeability (vancomycin resistance in Gram-negative and β-lactam resistance in Gram-positive) (6), lack of drug target sites (daptomycin resistance in Gram-negative) (7), the presence of multidrug resistance efflux pumps (RND efflux pump in *Pseudomonas aeruginosa*) (8, 9). However, when discussing the antimicrobial resistance, bacteria-harboring intrinsic resistance are not the focus of the problem. Rather, the expression of 'acquired resistance' in a previously susceptible bacterial population is a concern. Acquired resistance occurs as the result of gene mutations or via gene acquisition through horizontal gene transfer (HGT) (10). The prolonged and continuous exposure to antimicrobials compound in a particular ecosystem selects resistant bacterial populations in an ecosystem through Darwinian natural selection processes (Figure 1.1). As a result of this natural selection process, the susceptible bacterial population that was dominant in that ecosystem is replaced by a resistant population. The studies conducted on antimicrobial usage and its impact on the prevalence of resistant bacterial species indicated a positive association/correlation between antimicrobial usage and increasing resistant bacterial population in human and animal ecosystems (11, 12). Mobile genetic elements like plasmids, integrons, composite transposons, integrated conjugative elements, and bacteriophages can disseminate antimicrobial resistance genes (ARGs) among bacteria in

diverse environments (13). Bacteria may acquire ARGs through gene co-selection or selective pressure imposed by antimicrobials, resulting in ARGs acquisition through HGT events (Figure 1.1) (14). Co-selection of ARGs can occur through two mechanisms (i) co-resistance, where a selection of one gene supports the selection of another gene that usually does not offer a selective advantage to the compound of interest, for instance, co-selection of ARGs along with the heavy-metal or biocides resistance genes (15, 16) and (ii) cross-resistance where one resistance gene protects against a range of toxic chemicals (17, 18). Acquiring new DNA can affect the fitness of the host bacteria (19), but the origins and molecular mechanisms of these fitness costs are poorly understood. The cost of fitness can arise from the transfer of acquired gene (20), disruption of the bacterial genome (21), and metabolic cost associated with replication and gene expression and interaction between newly acquired protein, and host leading to disruption of cellular mechanisms. (22).

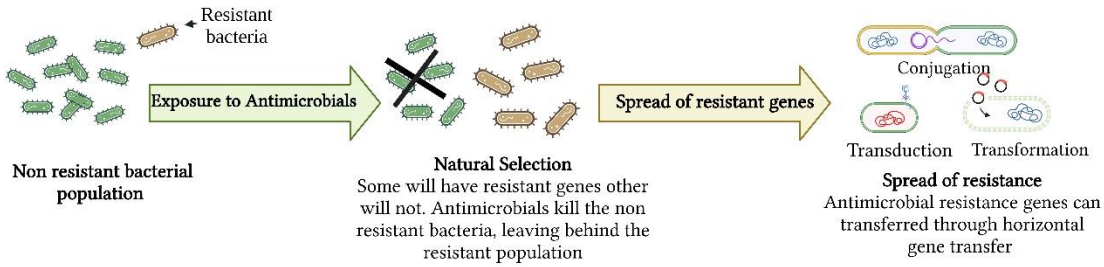

Figure 1.1 Exposure to antimicrobial compounds and spread of antimicrobial resistance. Prolonged and recurrent exposure to antimicrobial agents leads to the selection of bacterial populations that are resistant to these compounds. Consequently, these resistant can transfer their resistance gene to other susceptible bacterial population through horizontal gene transfer events.

1.3 Antimicrobial usage in One-health sectors and route of transmission

Antimicrobial resistance affects humans, animals, and the environments thus being considered as an important One-health problem. The possible route of AMR transmission is shown in Figure 1.2. The concept of One-health recognizes the interconnectivity of human, animal, and the environmental health and is endorsed at a global level by the World Health Organization (WHO), the Food and Agriculture Organization of the United Nations (FAO), and the World Organization for Animal Health (OIE) (23). It is estimated that about 10% of global antimicrobial use occurs in the United States (over 17 million/year). According to the FAO data of 2017-2018, 65% of medically important antimicrobials (approved for use in human and food-producing species) were consumed by livestock whereas 35% were consumed by humans.

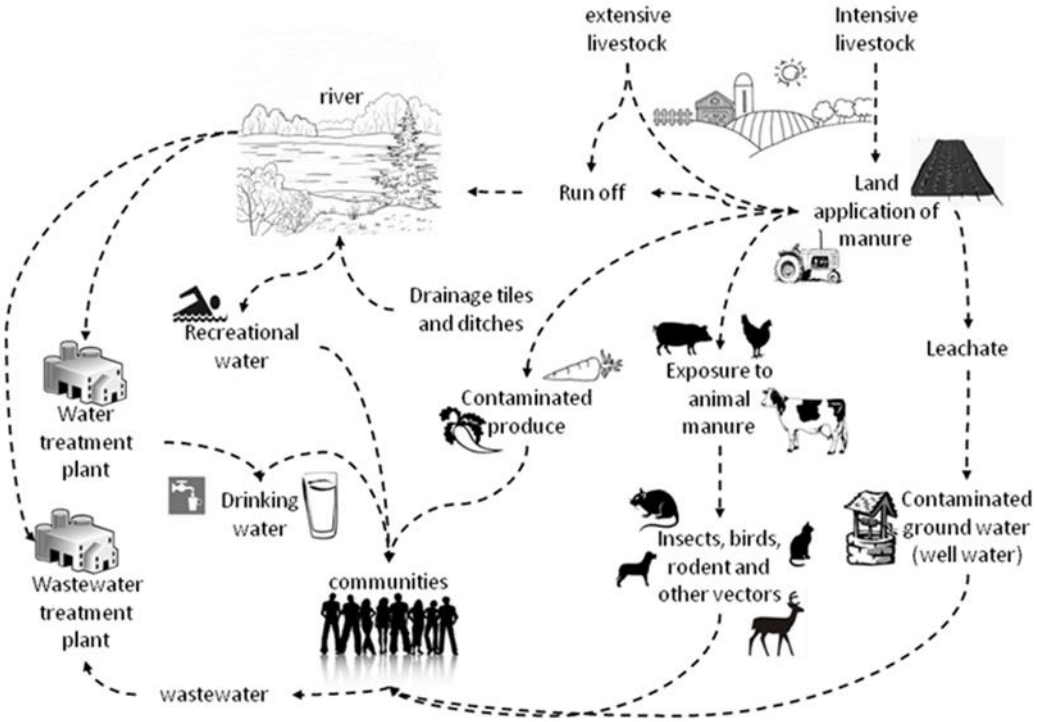

Figure 1.2 Possible routes of transmission of antimicrobial resistance across One-health sectors (24). Dashed lines indicate putative transmission paths. Antimicrobials consumed by animals in livestock select antimicrobial-resistant population. The resistant bacteria contaminate the animal fertilizer and ground/ natural water. Humans get exposed to resistant bacteria through the consumption of contaminated food, recreational water, and drinking water. The wastewater from hospitals, pets, and the communities, goes to wastewater treatment plants and after treatment enters the environment.

In human medicine, medically important antimicrobials are used primarily for the treatment of bacterial infections, but also in some cases to prevent the secondary infections (infections that occur after or because of another infection) during surgery and in high-risk immunocompromised patients undergoing chemotherapy (1, 25, 26). In the US, approximately 80-90% of antimicrobials were prescribed to out-patients (patients who donot require hospitalization) (636 antibiotic prescriptions per 1000 persons) in 2021. In these patients, penicillin was commonly prescribed (140 prescriptions per 1000 persons) (27). The excessive use of antimicrobials by humans is associated with its over the counter availability and inappropriate prescribing (26, 28).

In livestock, antimicrobials are used for disease treatment and prevention, for instance, mastitis in dairy cattle, bovine respiratory disease in beef cattle. Prophylaxis or metaphylaxis are practices where antimicrobials introduced to a mass population with the objective to prevent or decrease the disease burden by treating the entire population at a single point in time (29). Prophylaxis is defined as the administration of antimicrobial agents to a mass population without any clinical symptoms of disease, whereas in metaphylaxis antimicrobials are administrated to a susceptible population exhibiting clinical disease (30). The average annual global consumption of antibiotics in producing 1 kg of different meats is 172 mg/kg for pork meat, 148 mg/kg for chicken and 45 mg/kg for beef (31). Five major countries involved in the highest share of antimicrobial consumption during food-animal production are China (23%), the US (13%), Brazil (9%), India (3%) and Germany (3%).

Environment connects human and animal sectors (Figure 1.2). Exposure of antimicrobials in the environment occurs through several routes such as municipal and hospital waste,

animal husbandry, the manufacturing industry, runoff from agricultural fields containing livestock manure, and landfill leachates of antimicrobial discharge. According to a 2013 report, approximately 20-30% of patients in Europe received antimicrobials during their hospitalization period (32). The residual of the resistant bacterial population through urine and feces and the residual of antimicrobial compounds enter the hospital wastewater system. Hospital effluents are hotspots for horizontal gene transfer, facilitating the interspecies and intraspecies transfer of ARGs and virulence genes. Similar to humans, 30–90% of antimicrobials consumed in animals are released into manure, feces, and urine. The high concentration of antimicrobial residues in livestock settings contaminates the environment with antibiotic-resistant bacteria (33, 34). Therefore, the environment acts as a reservoir for mobile genetic elements that interact and spread to other sectors or human and animal hosts. Bioindicators in AMR surveillance programs

Bioindicators are defined as 'living organisms (microbes, animals and plants) that are used as a potential tool to monitor the changes (either positive or negative) in environmental health and their possible impact on human civilization' (35). In AMR surveillance programs, commensal gut bacteria are used as bioindicators.

Commensal gut bacteria colonize the gastrointestinal tract and live in a symbiotic relationship with the host (36). However, they act as opportunistic pathogens and can also cause infections in immunocompromised patients (37). Commensal bacteria are not the specific target of antimicrobials used in human and animal sectors however they get exposed to those antimicrobial compounds. The selective pressure due to this exposure results in the accumulation of AMR genes from the environment. Therefore commensal bacteria are considered as the reservoir of acquired ARGs (38, 39). Some commensal

bacteria including *Escherichia coli* and enterococci are ubiquitously found in One-health sectors including humans, animals and the environment. These bacterial species are used as 'indicator bacteria' in AMR surveillance programs (described in the next paragraph) to monitor the occurrence of AMR in different reservoirs throughout the food-chain as they (i) exist as commensal bacteria in the gut microbiota of healthy animals and humans, (ii) they can acquire AMR genes both via mutations in chromosomal genes and HGT of AMR genes, (iii) they have the potential to cause infections in both animals and humans and to transfer AMR genes to pathogenic bacteria of the same or other species through HGT (36).

Several AMR surveillance programs at the national level like the Swedres-Svarm program in Sweden (40), the RESAPATH program in France (41), the UK One Health program in England (42), the NethMap program in Netherland (43), the Danmap program in Denmark) (44), the National Antimicrobial Resistance Monitoring System for Enteric Bacteria (NARMS) https://www.cdc.gov/narms program in (the United states) and the Canadian Integrated Program for Antimicrobial Resistance Surveillance (CIPARS) program in Canada (45) and globally by the WHO: The Global Antimicrobial Resistance and Use Surveillance System (GLASS) (https://www.who.int/initiatives/glass/) recognize the roles that commensal gut microbiota can play in the occurrence, distribution, flow and diffusion of ARGs in human, animal, and environmental ecosystems. The majority of AMR surveillance programs like RESPATH, UK One Health report, NethMap and GLASS focus on *E. coli* as an indicator bacteria of AMR whereas programs like NARMS, Danmap, CIPARS, Swedres-Svarm include *Enterococcus* spp. in the surveillance programs along with *E. coli*.

1.4.1 Escherichia coli

Escherichia coli are gram-negative, rod-shaped, facultative anaerobes that are part of the Enterobacterales, a gamma division of the phylum Proteobacteria (46). In humans and other animals, including warm-blooded animals, insects, and reptiles. *E. coli* predominantly reside as commensal in the cecum and colon of the large intestine (47). *E. coli* can persist in the natural environment in soil, water, food, and sediments and are often used as an indicator of fecal contamination (48). Although 90% of *E. coli* strains are commensals, they can also act as pathogens in humans and animals (49). In humans, approximately 90% of hospital- and community-acquired urinary tract infections are caused by uropathogenic *E. coli*. *E.coli* can also cause infections like enteritis, septicemia, neonatal meningitis, and postsurgical peritonitis in humans (50-52). Multidrug-resistant *E. coli* were responsible for the majority of AMR-associated deaths in 2019 (>0.7 million deaths globally) (3). In livestock, *E. coli* are predominantly associated with diarrhea and are one of the primary pathogens associated with bovine mastitis in dairy cattle (53).

The genomic structure and size of *E. coli* are driven by evolutionary forces/pressure that reflect the lifestyle of the bacteria itself. The genome of *E. coli* may vary from 4.0-6.0 Mb (average 4200 number of genes) with a GC content of ~50% (54). The genetic content of pathogenic and commensal *E. coli* strains may differ by up to 20% (55). The bacterial genome can be classified into two categories; the core genome, and the flexible/dispensable/accessory genome (56). The core-genome is considered as the genetic backbone of a bacterial species. It consists of genes found in all strains of a bacterial species (>99% of species genomes). The core-genome consists primarily of housekeeping genes responsible for metabolic functions like replication, translation, and transcription (57). The

accessory genome represents genes present only in the genomes of a single or few strains of a species. In a gut microenvironment, the presence of diverse bacterial species increases the availability of new gene families. These new gene families can be acquired by *E. coli* through horizontal gene transfer, resulting in the expansion of the pangenome (58). These genes are typically responsible for facilitating adaptation to a specific environment. Mobile genetic elements like plasmids, integrons, bacteriophages, and transposons are components of the accessory genome that facilitate the transfer of accessory genes among the members of a bacterial community. The combined core and accessory genomes of a species constitute the pangenome, an open-pangenome represents the genetic diversity within a species (Figure 1.3). Characterization of the pan-genome of 1324 complete *E. coli* genomes showed that the pan-genome consisted of approximately 25,000 gene families. The softcore genome (genes that are present in >95% of genomes) of *E. coli* was more stable than the core genome (genes that are present in ≥99% of genomes). The core genome of *E. coli* consisted of ~3000 gene families (54). A similar pattern was observed in other comparative genomic studies of *E. coli* where the pangenome increased with the additions of genomes (~13,000 to ~23,000 genes), while the size of core-genome decreased (~2200 to ~800 genes) (59-62).

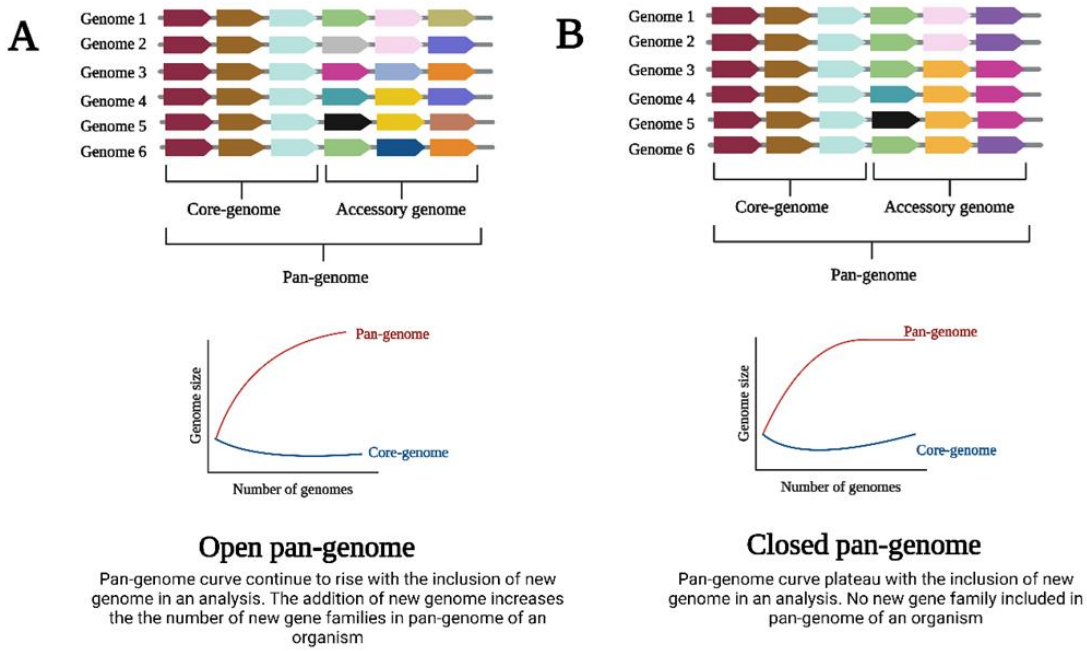

Figure 1.3 Pan-genome of bacterial species. Arrows represent genes in the genomes. genes found in all genomes of species constitute the core genome and genes found in a few genomes constitute the accessory genome. (A) In an open pan-genome, the graph between the number of genomes and genome size does not plateau as the addition of a new genome in an analysis increases the number of new gene families. (B) In a closed pan-genome, the graph between the numbers of genomes and genome size plateau as the inclusion of new genomes do not increase the number of the new gene family.

Phylogenetically, *E. coli* can be categorized into eight different groups (A, B1, B2, C, D, E, F, and G) (63, 64). The majority of phylotype group A and a small group of phenotype B1 and B2 were identified as commensal strains of *E. coli*, whereas environmental strains are usually categorized as phenotype C. Previous studies have shown that the prevalence of phylogenetic group A, B1, and B2 varies between humans and animals (wild, domestic mammals and reptiles), depending on host characteristics and environmental factors (65).

Commensal bacteria are defined as those bacteria that reside or colonized body surfaces. They co-evolved with their host however, in some immunocompromised conditions they can overcome the protective host response and cause infection (66). Whereas, pathogenic bacteria can cause mild to life-threatening infections in healthy individuals. Commensal strains of *E. coli* can also acquire virulence genes such as genes for biofilm formation as a means of adapting to diverse environments including human and animal gut. Studies have shown that some virulence genes provide *E. coli* with a selective advantage in colonizing the gut of a specific host. For example, strains harboring genes encoding adhesion, iron uptake, toxins, and protectins were more persistent in gut microenvironments in humans, dogs, and pigs than strains that lacked these traits (*E. coli* from the aquatic environment) (67-69). *E.coli* strains expressing intimin (adhesin) more readily colonize bovine rectal mucosa than those that lack this adhesion factor (70). Moreover, the prevalence of virulence genes may differ between humans and animals, where human commensal *E. coli* strains harbor more virulence genes as compared to *E. coli* isolated from the livestock Khi2 $p<0.0001$) (71).

E. coli has been used as an indicator of fecal contamination and to monitor AMR in humans, animals, and the environment for decades (72). Studies conducted in the late

1900s recognized the ability of *E. coli* to acquire and transfer ARGs between humans and livestock (73-76). *E. coli* readily acquires ARGs and MDR *E. coli* strains can outcompete sensitive *E. coli* strains in the gut via anaerobic metabolic processes, resulting in prolonged colonization and dissemination (77). The MDR *E. coli* induces an inflammatory response in the intestinal tract that results in the production of metabolic by-products that are utilized by MDR *E. coli* through anaerobic metabolic processes.

1.3.1.1 Resistance to medically important antimicrobials in E. coli

E. coli exhibiting resistance to critically important (category I) antimicrobials including 3rd generation cephalosporins (3GCs) and carbapenems are considered one the major challenges to human and animal health. *E. coli* can gain resistance against 3GCs by the production of extended-spectrum β-lactamases (ESBLs), and AmpC β-lactamases. ESBL producers *E. coli* are clinically relevant in veterinary medicine as they confer resistance against penicillin, aminopenicillins, and cephalosporins that may result in treatment failure. In epidemiological studies, the linkage between clinical ESBL producers *E. coli* and those recovered animals settings was not found (78). Carbapenem is considered for the treatment for MDR bacterial infections including ESBL-producing *E. coli* infections in humans and is thus referred to as a 'last resort' drug (79, 80).

Resistance to carbapenem is variable and is associated with the production of carbapenemase (carbapenem degrading enzyme), impermeability of porin proteins, utilization of carbapenem-insensitive transpeptidases, and increased activity of efflux pumps (81). The accumulation of more than one of these mechanisms results in a more pronounced effect on carbapenem activity (82, 83). Enzyme-mediated resistance is the most common mechanism of resistance (84). These enzymes hydrolyze all or almost all

beta-lactams, confer high levels of carbapenem minimum inhibitory concentrations (MICs), are encoded by genes that are horizontally transferable by plasmids or transposons and are commonly associated with genes encoding for other resistance determinants (85, 86). β-lactamases are broadly classified based on molecular characteristics, into four groups (A, B, C, and D) known as 'Amber's classification' (87). Molecular classes A, C, and D include the β-lactamases with serine at their active site, whereas molecular class B β-lactamases are all metalloenzymes with active-site zinc (88, 89). Carbapenemases are diverse and include three members (A, B, and D) of Amber's classification. Briefly, Class A includes some enzymes that are chromosomally mediated like non-metallocarbapenemase-A (NMCA), *S. marcescens* enzyme (SME), and imipenemase (IMI-1) while others are plasmid mediated such as *K. pneumoniae* carbapenemase (KPC), IMI-2, Guiana extended-spectrum (GES), and their derivatives. They can hydrolyze carbapenem effectively but inhibit by clavulanic acid (90). The most prevalent carbapenemases in this class are the KPC enzymes, which are now endemic in parts of the USA, UK, Greece, Italy, Israel, and China (90-94). Class B or metallo- β-lactamases are characterized by their ability to hydrolyze carbapenem and their resistance to β-lactamases inhibitors but susceptibility to EDTA. The most common families of this class are: Verona integrons-encoded MBL VIM, Imipenemase IMP, and New Delhi MBL (NDM) enzymes (95, 96). These are located within intergon structures as gene cassettes. The association of these integrons with transferable elements facilitates gene mobilization. Class D is an oxacillinase-type enzyme, characterized by its hydrolytic activity against oxacillin. They are usually found in *Acinetobacter spp.*, although there are reports of increasing prevalence in *Enterobacteriaceae* (97).

Although, resistance to carbapenem in *E. coli* from animal origins is low as carbapenems are not approved for use in veterinary medicine. However, it is increasing and carbapenemase-producing *E. coli* have been isolated from a variety of animals (livestock, wildlife, and companion animals) worldwide (98). Therefore, AMR surveillance needs to evaluate carbapenem resistance in livestock sectors. In this thesis, we investigated the prevalence of carbapenem-resistant *E. coli* in the beef production system.

1.4.2 Enterococcus spp.

Enterococci is of the four genera of the Enterococcaceae family of bacteria, belonging to the order Lactobacillales, class Bacilli of the phylum Firmicutes. This genus groups a great diversity of gram-positive bacteria, with an ovoid shape that do not form spores. They are a core member of the gastrointestinal microbiome of humans and other mammals and are frequently isolated from soil and water (99, 100). They are also one of the leading causes of nosocomial infections in humans (14%) (101, 102). The tolerance of *Enterococcus* spp. to a broad range of pH (4.6–9.9), temperature (10-45°C), desiccation, UV radiation, starvation, antiseptics, and disinfectants is the key to their widespread occurrence and persistence in diverse environmental conditions (103-106).

The genus *Enterococci* consists of 50 species that possess variable habitats, metabolic, phenotypic, and genotypic characteristics. *E. faecalis* and *E. faecium* are predominantly isolated from the mammalian digestive tracts of both humans and animals (101) and are responsible for the majority of infections in humans (107). Other species including *E. hirae*, *E. gallinarum*, *E. mundtii* and *E. casseliflavus* have also been isolated from human and animal environments. Due to their high prevalence and strong association with

nosocomial infections, *E. faecalis* and *E. faecium* have been the indicator species in most AMR surveillance studies.

The *Enterococcus* spp. genome range in size from 2.3 to 5.4 MB with GC content 34 to 45%. *Enterococci* display a high level of genomic plasticity driven by HGT, as evidenced by the open structure of their pan-genome. In 1964, enterococci were classified into five groups based on their sugar utilization, but this classification scheme has been replaced by a molecular-based classification in 2002 (108). Phylogenetically, enterococci are categorized into seven groups based on single nucleotide polymorphism analysis of the 16S rRNA gene (109). Comparative analysis of the metabolic pathways of *E. feacium* and *E. faecalis* identified 70 and 140 Clusters of Orthologous Groups of proteins (COGs) respectively, that were found exclusively in these species. This shows that considerable genomic variation exists among *Enterococcus* spp. at the species level (110).

Enterococci are intrinsically resistant to several antimicrobials (111) including penicillins, semisynthetic penicillins, ampicillin, and cephalosporins. These antimicrobials lack affinity to penicillin-binding proteins (PBPs) of *Enterococcus* sp (i.e., PBP4 in *E. faecalis* and PBP5 in *E. faecium*) required for their activity (101). They also possess a *lsa* gene that encodes for clindamycin resistance and exhibits low-level resistance to aminoglycosides. Cell-wall inhibitors such as β-lactams and aminoglycosides can act synergistically against enterococci and in combination they are often used to treat enterococcal infections in humans (112). These species can also readily acquire genes from the environment through HGT (113).

1.3.1.2 Resistance to medically important antimicrobials in Enterococci

Glycopeptides are critically important antibiotics in human medicine and are used to treat gram-positive bacterial infections (114). The emergence of vancomycin-resistant enterococci (VRE) in human clinical settings and livestock is of great concern. In Europe, VRE reservoirs were identified in livestock and humans and were shown to be linked to the use of avoparcin as a growth promoter in poultry (115). Vancomycin resistance is associated with the presence of the *van*-operon, which may be found on mobile genetic elements or integrated into the chromosome. Gentamicin is a member of the aminoglycosides, exclusively used in human medicine, and is categorized as a 'high important' antimicrobial. The acquisition of *aac(6')Ie-aph(2'')Ia* encoding bi-functional enzyme results in high-level gentamicin resistance in enterococci. This gene was found more in clinical enterococci compared to those recovered from livestock and is often associated with transposon Tn5281 (116).

1.4.3 Pseudomonas aeruginosa

Pseudomonas aeruginosa is a gram-negative, environmental bacteria. This bacterial species exhibits extensive metabolic adaptability, enabling it to thrive in an extraordinary range of niches (117). It is an opportunistic pathogen and can cause a wide range of infections in humans, animals, and plants. In humans, is of the top six infectious disease threats, causing infection in immunocompromised individuals. It is also considered as the major pathogen contributing to the burden of hospital-acquired infections associated with chronic lung infections.

P. aeruginosa is a challenging pathogen to treat due to the formation of biofilm during infection. In addition to this, it can naturally resist multiple antimicrobial compounds

through variable mechanisms including prevention of access to target, increased efflux, changes in antimicrobial the targets by mutation, modification, and protection of targets, and direct modification of antimicrobial compounds (117, 118).

Pseudomonas aeruginosa PAO1 was the 25th genome sequenced completely. The genome size of *P. aeruginosa* can vary between 5.5 to 7 Mbp (119). The *P. aeruginosa* genome consists of more than 550 transcriptional regulators, which presumably play exquisite roles in coordinating the colonization of a wide range of ecological niches (120).

1.3.1.3 Resistance to medically important antimicrobials in P. aeruginosa

The global pattern of AMR in *P. aeruginosa* varies substantially. The highest rate of AMR in *P. aeruginosa* occurs in North, Central, and South America, Western and Central Europe, China, India, and Southeast Asia (121). The *P. aeruginosa* lineages ST235 and ST175 are major contributors of hospital-acquired infection and have emerged as high-risk globally dispersed clones (122, 123). The resistance to last-resort drugs like polymyxin and carbapenems in *P. aeruginosa* is well-documented. The resistance to carbapenems in *P. aeruginosa* is associated with loss of porin protein due to mutation, increased efflux activity (MexXY efflux pump), and the production of carbapenemase enzyme to degrade carbapenem antibiotics. The mechanism of resistance is described in detail by Horcajada, J, and colleagues (124)

1.4 High-throughput whole-genome sequencing and genomic characterization

Conventionally, phenotypic characterization of bacteria-based methods including serotyping, antimicrobial susceptibility testing (AST), and mass-spectrometry was widely

used for AMR surveillance (125). Upon, the advent of DNA sequencing technologies, nucleic acid-based methods were used to complement phenotypic methods to achieve phenotypic and genotypic information on resistant bacterial strains. Nucleic acid-based techniques can be broadly categorized into non-amplification and amplification-based typing. The most common non-amplification-based typing methods are restriction fragment length polymorphism (RFLP) (126, 127), matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF) (128), and pulsed-field gel electrophoresis (PFGE) (129). The DNA amplification-based methods frequently for characterization are multiple-locus variable-number tandem-repeat analysis (MVTR), multilocus sequence typing (MLST), and virulence typing. One disadvantage of nucleic acid-based methods is that they provide limited genotypic information as they can only detect/amplify known regions in a genome. Considering the level of complexity of bacterial genome and frequency of ARG acquisition via HGT, it is difficult to study gene acquisition and thereby bacterial evolution, using nucleic acid-based technologies alone (130).

1.5.1 Whole Genome sequencing technologies

Whole genome sequencing (WGS) can provide high-resolution coverage of bacterial genomes. It can detect mobile genetic elements, ARGs, and virulence genes and it can be used for epidemiological typing. The ground-breaking Sanger sequencing method was the precursor to WGS, referred to as ‘first generation sequencing’ (131). Sanger sequencing depends on chain termination with the use of distinct fluorescent dyes to terminate the synthesis reaction. Next-generation sequencing (NGS) has revolutionized the field of genomics by significantly increasing the amount of genetic material that can be sequenced compared to Sanger sequencing. This technological advancement has had a profound

impact on various areas of research, including studies related to antimicrobial resistance (AMR). Most genomic studies focused on antimicrobial resistance now depend on whole-genome sequencing (WGS) (132, 133). Sequencing the entire genome through WGS platforms including short-read and long-read technologies enables comprehensive analysis of the genetic variations underlying AMR, including the identification of resistance-conferring mutations, mobile genetic elements carrying resistance genes, and their distribution within microbial populations (134) (Figure 1.4).

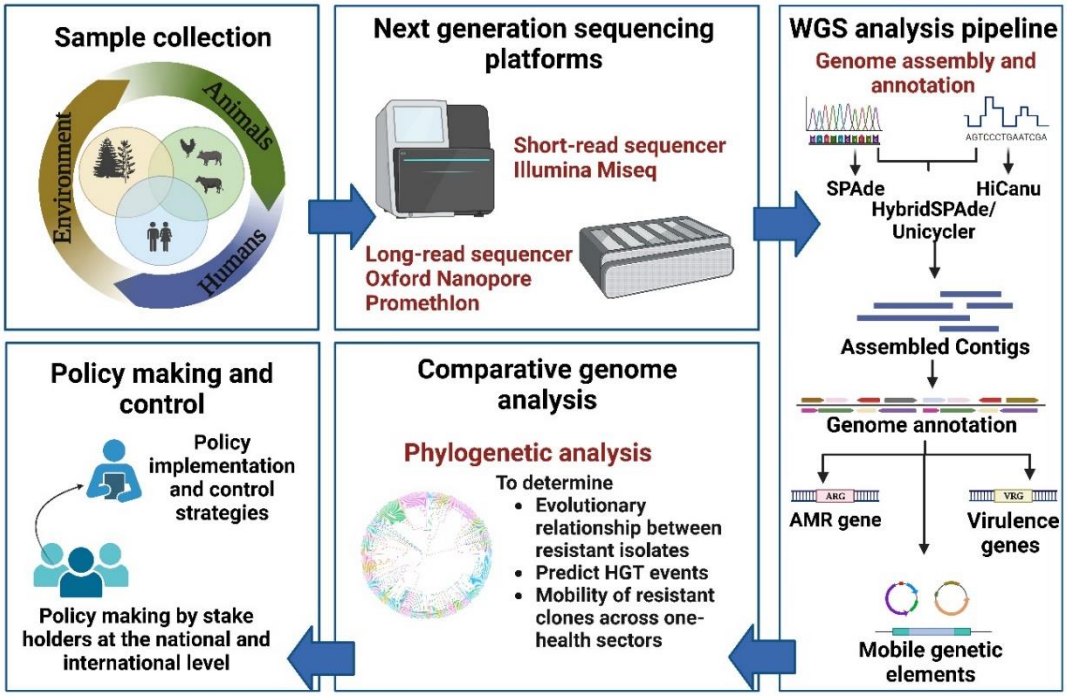

Figure 1.4 Pipeline for AMR surveillance using short and long read whole genome sequencing technologies. Sample are collected from One-health sectors including humans, animals and environment. Following genomic DNA extraction from isolated bacteria, whole genome sequencing care performed on short or long read sequencing technology platforms. Reads generated are assembled using short read (SPAde), long read (HiCanu) or hybrid (HybridSPAde/ Unicycler) assembling tools. The assemble contigs are annotated with Prokka. Antimicrobial resistance gene (AMR), virulence gene and mobile genetic elements are identified using specific tools (discussed below). The collection of genomes are used for comparative genomic analysis. The analysis and findings of surveillance programs are utilized by stake holders to make and implement control strategies at local, national and at international level.

1.4.1.1 Short-read sequencing technologies

Short-read sequencing technology or second-generation sequencers generate read lengths ranging from 36 to 300 base pair with >99% read accuracy (135). The accuracy of these methodologies has continuously improved the sequencing, throughput has increased and the overall cost of sequencing has declined (136). Single-read sequencing platforms include Ion-torrent (137), 454 pyrosequencing (http://www.454.com/) (138), SOLiD (139), and Illumina (http://www.illumina.com/). About 90% of all short-read sequencing data has been generated using the Illumina sequencing platform (140). Illumina short-read sequencing is based on the sequence by synthesis and has 4 models: Miniseq, Miseq, Hiseq, and Nextseq, where Miseq is the most widely used platform for bacterial genome sequencing. The workflow of Illumina sequencing consists of three steps as described by Kuchta et al. (141). (i) library preparation, in which genomic DNA (gDNA) is fragmented mechanically via sonication or enzymatically via transposases into ~200-bp lengths. Unique adaptors are attached to the 3' and 5' ends of DNA fragments and amplified by the polymerase chain reaction (PCR). (ii) clustering in which clonal clusters are generated through a bridge amplification process (iii) sequencing in which clonal cluster is sequenced using sequence by synthesis technology with DNA polymerases incorporating nucleotide bases into the DNA template. The incorporation of nucleotide is detected through fluorescence (142). Illumina sequencers provide high output (1.2–6,000 Gb) and high accuracy at a relatively low cost per base compared to long-read platforms.

1.4.1.1.1 Assembly pipelines

Illumina can generate single-end reads or pair-end reads. In pair-end sequencing, DNA fragments are sequenced from both ends, thereby, generating more sequence data. The

reads generated through Illumina can be assembled in two ways; either as reference-based or *de novo* genome assemblies. In reference-based assembly, the generated reads are aligned against a reference genome. There are numbers of software packages for reference-based genome assemblies, which differ in their algorithms and performance in terms of speed and scalability. Most reference-based assemblers depend on Burrows-wheeler transform algorithms like Bowtie (143), BWA (144), KART (145), and BWBBLE (146). The disadvantage of using reference-based assemblers is that the appropriate reference genome must be available for the assembly to be conducted (147).

Another approach for genome assembly is the *De novo* method. In this method, the genome is assembled based on overlapping sequences or contiguous sequences. *De novo* assembly algorithms can be categorized into three broad classes; overlap-layout-consensus (OLC), extension-based, and De Bruijn (or Eulerian) graph. The use of OLC is computationally intensive and are used by only a few assemblers like Edena (148). SGA (149) and FERMI (150). Extension-based methods as SSAKE (151), and JR-Assembler (152) are computationally efficient but are prone to sequencing error. *De Bruijn* (or Eulerian) graph (153) is used for short-read assemblies due to their accuracy and efficiency. In this approach, reads are partitioned into *k*-mers (*k*-mer is a sequence of *k* characters in a string or nucleotides in a DNA sequence) to form a graph node which is linked with a shared *k*-1 mer. Some *de Brujin*-based assemblers include Velvet (154), SPAde (155), SOAPdenovo (156), and ABySS (157), with SPAde (https://github.com/ablab/spades) being one of the most widely used tools in WGS studies (158).

1.4.1.1.2 Limitations of short-read sequencing

Short-reads generated through the Illumina platform have some limitations. Multiple copies of bacterial genomic elements including CRISPR (clustered regularly interspaced short palindromic repeats) array, transposons, insertion sequences, ribosomal genes, rhs (Rearrangement hot spot) toxins, secondary metabolite biosynthetic gene clusters, genomic islands, and complex resistance regions in plasmids may harbor repetitive sequences at different genomic locations that exceed the read length of Illumina platforms. These regions are challenging to assemble and frequently result in misassemblies and gaps that preclude the closure of a bacterial genome. As a result, genomes sequenced using Illumina platforms are often fragmented into hundreds of contigs. Furthermore, algorithms for *de novo* genomic assemblies struggle with the presence of intergenic repeats and GC biases created during library preparation (159, 160). Misassembled genomes can prevent the accurate identification of MGEs, gene copy number, and gene cluster(s). These limitations significantly affect the informational value of draft-quality genomes sequenced using short-read Illumina platforms.

1.4.1.2 Long-read sequencing technology

Lon-read sequencing as the name indicates generates longer reads compared to the short-read sequencing platforms. The long-read platforms target single DNA molecules in real-time, generating sequences of longer length (maximum read length of 100kb) in less time than short-read sequencers. Two long-read sequencing platforms, PacBio and Oxford Nanopore are the most commonly used long-read sequencing platforms, although additional technologies such as Infinity (based on sequence-by-synthesis technology by Illumina) are emerging (not launched yet). PacBio is based on the circular consensus

sequencing (CCS) method. The target dsDNA are attached to the ssDNA adaptor that is then sequenced multiple times creating one continuous long read sequence (CLR) (161). The most recent Sequel II system is based on high-fidelity (HiFi) sequencing technology and it can produce 30 kb sequences with a >99.99% accuracy. Through this technology, complex genomic structures including MGEs can easily be identified.

In this thesis, nanopore sequencing is used therefore, it is described in detail in later sections. Nanopore was first introduced by ONT in 2014. In Nanopore sequencing, the DNA molecule passes through a nanopore (A nanopore is a tiny, nanoscale-sized hole usually made from a protein or a synthetic material.). An electrolyte potential is created around the nanopore, creating an electric field. The DNA molecule passing through the pore disrupts the voltage flow across a channel, through a voltage detector. Unlike other platforms, the nucleotides or enzymes like polymerase and ligases for sequencing are not required, as a result, nanopore can tolerate temperature variation during sequencing (162). Available Nanopore platforms include MinION, GridION, and PromethION, which differ in the numbers of flow cells they contain and consequently sequence output.

Libraries must also be prepared for Nanopore sequencing, with a wide range of library preparation kits available that use either amplification or non-amplification methodology. Amplification-based kits are typically used if the quantity of DNA available is low. For library preparation, the DNA is sheared using G-tube into small DNA fragments (6-20kb). DNA damage is then repaired through an end-repair step, followed by the addition of a polyA tail to the 3' end of DNA for adaptor ligation. The DNA is then ligated using Leader (Y-adaptor) and a hairpin (HP-adaptor). The final step of library preparation is the His-

bead purification of the DNA to separate it from free nucleotides and ligase (https://nanoporetech.com/support/how-it-works).

The nanopore sequencing begins from the 5'-end of the Y-adaptor followed by template DNA, then it progresses to the HP-adaptor followed by the complementary strand. The motor protein unzips the DNA molecule when it approaches the complementary site of the Y-adaptor, enabling the template strand to pass through the pore. The HP-adaptor allows sequencing of the complementary strand in a similar manner. The change in electrical current is then detected by an Application-specific integrated circuit (ASIC) sensor which transfers it to a base caller for reading.

1.4.1.2.1 Assembly pipelines

The MinKNOW by Oxford nanopore technology (ONT) (https://github.com/nanoporetech/minknow_api) pipeline is used for data processing, acquisition, and analysis of these raw data generated from Nanopore sequencers. Similar to short reads, raw reads can be assembled either using a reference-based or *de novo* approach. The basic pipeline of Nanopore assembly consists of base-calling followed by read-to-read alignment. The generated pair-wise aligned reads are then used to generate a draft assembly that is mapped in a read-to-read-mapping step to generate the assembled genome. Some tools performed multiple tasks for instance HiCanu and Flye assemblers generate polished genomes therefore, the described pipeline can be altered based on the study approach and tools being used (Figure 1.4). The steps are as follows

(1) Base-calling is defined as conv. of raw current to raw reads or nucleotide sequence (Table 1.1). The reads are called either 1-directional or 2-directional if a single strand

or both strands, respectively are used by the basecaller. There is a wide range of basecallers available for this purpose including ONT (Metrichor, Nanonet, Albacore, Guppy, Scrappie, and Flappie), Nanocall (163), DeepNano (164), BasecRAWller (165), and Chiron (165).

(2) The raw reads generated through the basecaller are then mapped using an alignment tool to generate a consensus contig. This step is referred to as read-to-read mapping (Table 1.1). The overlap-layout consensus algorithm is an assembly method that utilizes a graphical representation of reads and their overlaps to construct a consensus sequence. In this algorithm, each read is represented as a node, and the overlaps between reads are represented as edges in a graph. The algorithm begins by comparing pairs of reads to identify overlapping regions. These overlaps are used to create edges between the corresponding nodes in the graph. By analyzing these overlaps, the algorithm determines the most likely alignment and ordering of the reads. As the algorithm progresses, it iteratively merges and extends the reads based on the identified overlaps. This process helps build longer contiguous sequences, called contigs, by stitching together the overlapping regions. To generate a consensus sequence, the algorithm analyzes the overlapping regions of the contigs. It considers the base calls from each read, taking into account their quality scores and potential errors or variations. By reconciling the overlapping bases, the algorithm generates a consensus sequence that represents the most accurate and reliable representation of the original DNA sequence. Overall, the overlap-layout consensus algorithm leverages the overlapping information between reads to construct a consensus sequence, allowing for the assembly of longer and more accurate contigs. Two alignment tools

recommended for this step are; GraphMap (166) and MiniMap2 (167). As compared to GraphMap, MiniMap2 is computationally efficient but less accurate. Some assemblers such as HiCanu (168), Flye (169), and NEXTdenovo (https://github.com/Nextomics/NextDenovo) do not require a read alignment tool, these tools do read alignment themselves.

(3) The alignment tool constructs an overlap graph, which serves as the basis for the assembler to generate a layout consensus sequence (Table 1.1). This graph represents the relationships and overlaps between the reads, forming the foundation for subsequent assembly steps. Two strategies for *de novo* assembly are; (i) correction, followed by assembly, and (ii) assembly followed by correction. In the correction-then-assembly approach, assemblers like HiCanu (168), and NextDenovo (https://github.com/Nextomics/NextDenovo) depend on first correcting errors in the raw reads before conducting the assembly process. These assemblers typically utilize error correction tools to refine the reads followed by the assembly process The aim is to generate more accurate and reliable assemblies. In the 'assembly then correction' method, assemblers like Miniasm (170), Flye (169), SMARTdenovo (171), and Ravon (172) conduct the assembly process first and then refine the assembly through subsequent error correction steps. These assemblers typically utilize the overlap graph constructed from the raw reads to generate an initial assembly followed by identification and errors or variations correction within the assembly using a correction tool, thereby enhancing its accuracy and completeness.

(4) If the quality of the draft assembly is low, it can be improved by adding a step where the generated draft assembly is mapped against raw long reads using an alignment tool such as GraphMap, MiniMap2, or BWA-MEM (Table 1.1).

(5) Polishing is the final step in the assembly as it identifies mismatches, indels and removes gaps in the final assembled genome. The available polishing tools are Nanopolish and Racon (Table 1.1). Polishing tools improve consensus sequence for a final draft genome.

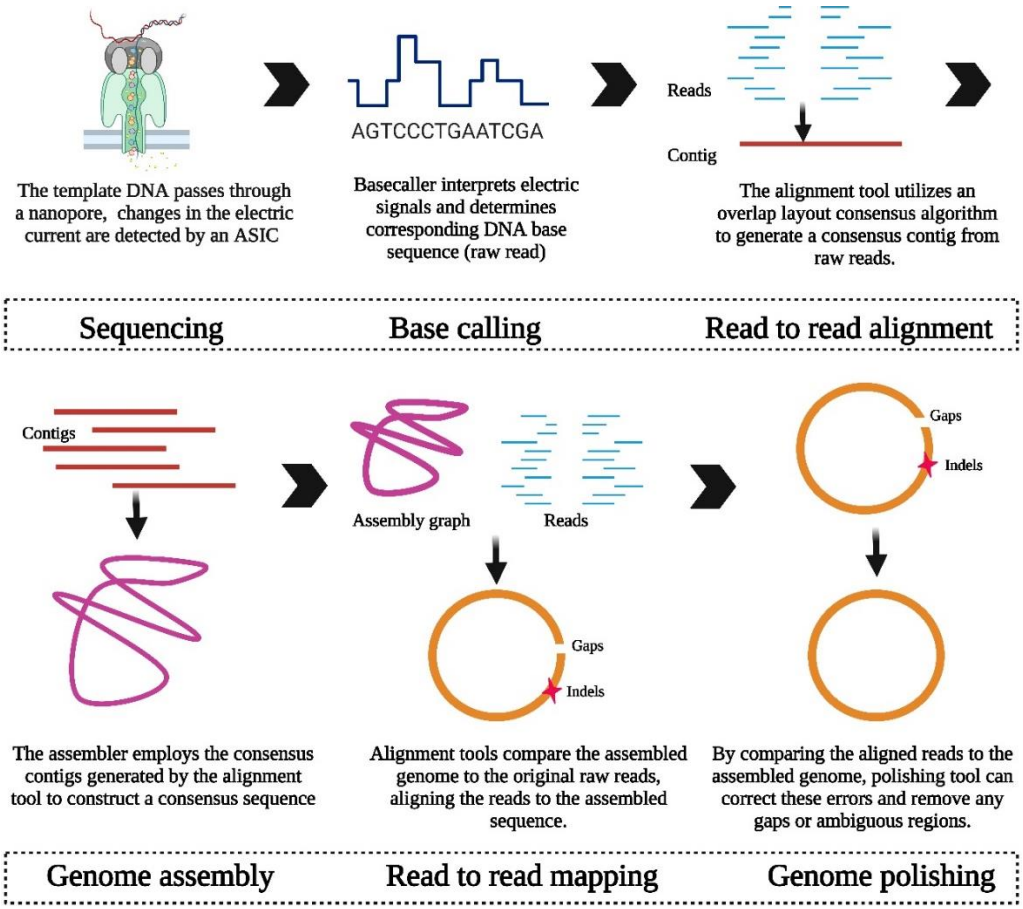

Figure 1.5 Whole genome assembly pipeline using long-read generated through Nanopore sequencer. The template DNA passes through a nanopore, and during this process, any changes in the electric current are detected by an Application-Specific Integrated Circuit (ASIC). Once the ASIC detects the variations in current, this information to a basecaller. The basecaller is responsible for interpreting the electrical signals and determining the corresponding DNA base sequence. The alignment tool utilizes an overlap layout consensus algorithm to generate a consensus contig from raw reads. This algorithm compares the raw reads and identifies regions of overlap between them, generating a consensus contig that represents the most reliable representation of the original DNA sequence. Using various algorithms and statistical models, the assembler resolves conflicts and resolves ambiguities to generate a consensus sequence. It considers factors such as read coverage, base quality, and potential variations within the contigs. After the genome is assembled, it can be further improved by mapping it against the raw reads using alignment tools. By analyzing the alignment results, the tools can identify discrepancies, such as misassembled or incorrectly ordered contigs, as well as regions where the assembly may be fragmented or contain errors. The polishing tool plays a crucial role in refining the assembled genome by removing indels (insertions and deletions) and gaps. It achieves this by mapping the assembled genome against the original raw reads. The polishing tool utilizes various algorithms and techniques to adjust the assembled genome based on the alignment information.

1.4.1.2.2 Limitation of long-read sequencing

Despite the fact that long-read sequencing has greatly improved the quality of whole genome sequencing, covering repeated regions in a genome. However, there are still some limitations. The largest limitation is its lower read quality with and error rate of 15% for Nanopore long read sequencing vs error rate of less than 1 % for short read sequencing (173, 174). Compared to short read sequencing, long read sequencing is costly limiting the large-scale sequencing projects. The higher cost is mainly associated with the need for specialized instruments, reagents, and computational resources. Long-read sequencing requires longer DNA fragments. The DNA samples be more prone to fragmentation and damage during sample preparation and sequencing, leading to a loss of genomic information and affect the quality and accuracy of the generated sequences.

Table 1.1 Tools for constructing bacterial genomes using long-read sequencing technologies generated through Nanopore platforms.

Tools	Algorithm	Description	Input Data	Availability (developer)
Basecallers				
Metrichor	HMM [1] and Viterbi decoding	Raw current is converted into event (each event corresponds to movement of k-mer through the pore). HMM and Viterbi decoding algorithm are used to model event space and decoding the base sequence, respectively.	1D[2] and 2D[3] reads	Cloud based (Oxford)
Nanocall			1D reads generated by R7.3	Open-source (https://github.com/mateidavid/Nanocall) (MIT)
DeepNano	RNN[4]	Bidirectional recurrent neural approach to model statistical characterization of event and nucleotide sequence prediction	1D and 2D reads	Open source, offline (https://github.com/jeammimi/deepnano)
Guppy			1D and duplex	Current "official" ONT basecaller
BasecRAWller	Unidirectional RNN	Use two unidirectional RNN, initial neural network predict the event boundary between segments, then these segments are transformed into discrete event through segmentation algorithm then second neural network decodes events into base sequences	Raw signal	Berkeley lab (https://www.osti.gov/biblio/1572483)
Chiron	Segment free deep neural network or Event-free	Directly translate current measurement into a base sequence. It is based on 5-mer segmented sequence and decodes final base using Viterbi decoding algorithm	Raw signal	https://github.com/haotianteng/Chiron
Alignment tool				
Graph-Map	OLC[5]	It uses progressive approach. First, reads are divided into k-mers then construct a	Raw reads	https://github.com/isovic/graphmap

[1] Hidden Markov model
[2] One-directional
[3] Two-directional
[4] Recurrent neural network
[5] Overlap layout consensus

		Hash-table. Two sets of information are generated with incorporation of each *k*-mer (i) *k*-mer strings and (ii) position in read. The reads are then aligned based on overlapping regions.		
MiniMap2		Reads are divided into k-mers and minimizer approach to construct a good alignment.	Raw reads	https://github.com/lh3/minimap2
Assemblers				
HiCanu	Celera with OLC	It pipeline consist of error correction, trimming and assembly. It can also perform read alignment and polishing steps.	Raw reads	https://github.com/marbl/canu
Miniasm	OLC based approach with read correction and consensus generation	It concatenates read sequences to generate the final sequences. Polishing using Racon is recommended.	All-vs-all read self-mappings (usually from MiniMap2)	https://github.com/lh3/miniasm
SMARTdenovo	OLC based approach without error correction	The assembly consist of four steps; overlapping, trimming, layout and consensus	All-vs-all read alignment	https://github.com/ruanjue/smartdenovo
Ravon	OLC approach	It builds an assembly graph from reads, preprocessed by pile-o-gram followed by polishing with Racon	Raw reads	https://github.com/lbcb-sci/raven
Flye	Repeat graph approach	It take approximate sequence matches to build an assembly. It produced polished genome	Raw reads	https://github.com/fenderglass/Flye
NextDenovo	String graph based approach	The assembly pipeline involves error correction and assembly	All-vs-all read alignment	https://github.com/Nextomics/NextDenovo
Polishing				
Nanopolish	HMM-based	It uses raw reads mapped it against draft genome to evaluate and maximize the likelihood of each base with HMM approach. BWA-MEM for read mapping is recommended.	Draft genome	https://github.com/jts/nanopolish

Racon	Partial order alignment graph	It finds consensus sequence between draft genome and raw reads. MiniMap2 for read mapping is recommended	Draft genome	https://github.com/isovic/racon

1.4.1.3 Hybrid Assembly

To overcome the limitation of both short-read and long-read technologies (discussed in previous sections), hybrid-sequencing was introduced (175). In hybrid assembly, both short-read and long-read data are used to generate a complete gap-free genome.

1.4.1.3.1 Hybrid genome assembly pipelines

Hybrid assemblies can be achieved in one of three ways; (i) assembly of short-read corrected raw long-reads into contigs through assemblers like PBcR (176), and MaSuRCA (Maryland Super Read Cabog Assembler) https://github.com/alekseyzimin/masurca (177), (ii) raw long reads assembled first then polished using short-read data, and (iii) short-reads assembled first then long-reads are used to build longer contigs using assembler like hybridSPAde (178), Unicycler (179), DBG2OLC (*de Bruijn* graph to an overlap-layout-consensus) (180), and Wengan (181).

Unicycler, and HybridSPAde assemblers are commonly used for the hybrid assembly of bacterial genomes. Both these assemblers depend on SPAde to generate an assembled genome. This generated assembly graph is then aligned with long-reads to resolve ambiguities and to create longer contigs.

1.4.1.3.2 Limitation of hybrid genome assembly

Although, the hybrid genome assembly pipeline approach offers numerous advantages in genome construction, it has some limitations. First, it requires both short and long-read sequences that increase overall sequence cost and complexity. Combining these two technologies required expertise and infrastructure. Integration and data analysis from two

different platforms can challenging and may require sophisticated computational approaches. In addition to this, compared to single-platform data, hybrid genome assembly generates complex and large datasets that require more computational resources and time for analysis. Short and long-read sequencing platforms have their limitation that can be inherited by hybrid genome assembly that can introduce additional challenges in genome assemblies. The integration of data from two different sequencing technologies can introduce errors and inconsistencies during the assembly process.

In conclusion, many different platforms are available for WGS analysis and all of them have some limitations. Therefore it is important to consider these limitations when utilizing these technologies and to carefully assess the trade-offs between available platforms based on study objectives. Moreover, it is also important to stay updated with advancements in WGS technologies. The evolution of these technologies may shift the balance between limitations and advantages, leading to more opportunities and a better understanding of bacterial genomics.

1.5.2 Gene annotation and characterization

Assembled genome is used to garner additional information through gene annotation. There are a wide variety of computational tools and pipelines available for gene annotation. The standard approach is sequence-to-sequence or sequence-to-model based searches such as through BLAST (Basic Local Alignment Search Tool) (182). In this method, the sequence is scanned against protein models and/or domain families (BLASTx). Another more sensitive approach is profile-based identification using Hidden Markov Models (HMM). The most commonly used annotation tool in whole genome studies is PROKKA, which utilizes both sequence-to-sequence and profile-based methods for annotation (183).

The computational tools for genome annotation have been reviewed by Ejigu GF. et al. in 2020 (184).

The accurate identification of AMR, virulence factors, and MGEs is crucial in surveillance programs. With the development of NGS, the number of prediction tools have been introduced Similar to annotation tools, these tools also depend on BLAST or/and HMM-based methods for identification. Table 1.2 summarizes the available tools for ARGs, virulence genes, MGEs, and plasmid identification.

Table 1.2 Bioinformatics prediction tools available for bacterial resistome, mobilome and virulome prediction.

Prediction tools	Description	URL
CARD [6]	A primary AMR knowledge resource and database, provides genotype and phenotype prediction from curated publication and sequences. The curated data are organized using ontologies; MO[7], RO [8], OBO [9], NCBI Taxon [10] (). The graphical web tool RGI [11] provided by CARD can do annotation of query genome.	Open-source, accessed from http://arpcard.mcmaster.ca (last updated on September 2022)
ARG-ANNOT (Antibiotic Resistance Gene-ANNOTation)	Utilizes local BLAST program in Bio-Edit software without a web server to detect ARGs and resistance associated point mutations in a genome. the prediction is made based on published data and online resources for ARGs data.	https://www.mediterranee-infection.com/acces-ressources/base-de-donnees/arg-annot-2/ (last updated on July 2019
ResFinder	Provides information of ARGs from completely and partially sequenced genome and can only identify acquired genes. PointerFinder in ResFinder 4.0 project can now also predicts mutations based resistance.	https://cge.cbs.dtu.dk//services/ResFinder/
STAR-AMR	Scans bacterial contigs using BLAST against ResFinder, PlasmidFinder databases to identify AMR and plasmids, respectively in a query genome. It also performs MLST typing.	https://github.com/phac-nml/staramr (last updated in 2022)
PATRIC (The Pathosystems Resource Integrated System)	Provides multiple analysis including genome assembly and annotation, protein family comparison, ARGS, virulence gene identification, phylogenetic analysis and met-data information. It collects data from public genome and databases like CARD and ARDB	https://legacy.patricbrc.org/
AMRFinderPlus	It is developed by NCBI to identify AMR genes, mutations associated with resistance based on reference gene database and curated collection of Hidden Markov Model.	https://github.com/ncbi/amr

[6] The comprehensive antibiotic resistance database
[7] Model ontology
[8] The CARD relations ontology
[9] Open Biological and Biochemical ontology
[10] NCBI-organismal taxonomy ontology
[11] Resistance Gene Identifier

ARGminer (Antibiotic Resistance Gene miner database)	It is an online platform to ensemble data from other databases including CARD, ARDB, MEGARes, ResFinder, SARG, ACLAME and PATRIC.	https://bench.cs.vt.edu/argminer/#/home
ResFams	It is a protein families database, linked to their HMMs associated with AMR prediction function. It is relatively more sensitive than BLAST-based tools but is computationally expensive.	http://www.dantaslab.org/resfams
ABRicate	It can identify AMR, virulence genes, plasmids with multiple databases including NCBI, CARD, ARG-ANNOT, Resfinder, MEGARES, EcOH, PlasmidFinder, *Ecoli*_VF and VFDB.	https://github.com/tseemann/abricate (last updated on March 2020)
PlasmidFinder	It is a NIH[12]-based tool, developed to predict complete or partial plasmids in assembled bacterial contigs.	https://github.com/kcri-tz/plasmidfinder (last updated on July 2022)
HyAsP (Hybrid Assembler for Plasmids)	It is plasmid prediction tool that depends utilizes both reference-based and denovo methods, using information of plasmid genes and read depth, respectively.	https://github.com/cchauve/HyAsP
Mob-Recon	Based on reference-based prediction of plasmids using collection of known relaxases and replicases.	https://github.com/phac-nml/mob-suite
PlasmidSeeker	It is a *k*-mers based tool to identify known plasmid in an unassembled bacterial genome.	https://github.com/bioinfo-ut/PlasmidSeeker
IslandViewer 4	A webserver for prediction and visualization of genomic islands in bacterial genome. It has integrated three genomic island prediction tools; IslandPath-DIMOB, SIGI-HMM and IslandPick.	https://www.pathogenomics.sfu.ca/islandviewer/ (last updated on December 2022)
ICEFinder	An online tool for predicting integrated and conjugated elements in a bacterial genome.	https://bioinfo-mml.sjtu.edu.cn/ICEfinder/index.php
ISFinder	A database and prediction tool for insertion sequences.	https://www-is.biotoul.fr/about.php
PHASTER ((PHAge Search Tool Enhanced Release)	A web accessible tool for the identification of prophage in bacterial genome and plasmid sequence.	http://phaster.ca/ (last updated December 2022)

[12] National Institutes of Health

1.5.3 Impact of WGS on surveillance programs

The key objectives of AMR surveillance programs (briefly discussed above) are to determine the epidemiology of AMR bacterial species, the prevalence of resistance in One-health sectors, the detection of unknown resistance phenotype and genes associated with it, the spread of particular types of resistant strains in different sectors and the association of the AMR with outbreaks in human and livestock sectors. Whole sequencing technologies allow comprehensive detection of the bacterial resistome[13], virulome, and mobilome, and thus address they address key goals of AMR surveillance program.

In a WGS-based surveillance program, phenotypic characterization methods are often undertaken to enable the examination of phenotypic and genotypic correlations of an AMR clone. Phenotypic resistance is usually determined through AST profiling methods such as minimum inhibitory concentration (MIC) testing using agar dilution or micro-broth dilution techniques.

Whole genome sequencing AMR data can be correlated with AST data using one of the approaches i.e. rule-based and model-based. In these approaches, the AMR phenotype is predicted based on the clinical breakpoints for each antimicrobial as set by the Clinical & Laboratory Standards Institute (CLSI) and the European Committee on Antimicrobial Susceptibility Testing (EUCAST) for each bacterial species. In the rule-based method, AMR genes or *k*-mers are analyzed using raw reads or the assembled genome using software (as described above). Whereas in the model-based approach, the model is trained to predict phenotypic resistance based on genomic data utilizing the information from a set

[13] Resistome is a collection of total ARGs and their products in pathogen, antibiotic producer and environmental bacteria, contributing to AMR.

of genomes with known phenotypes using machine learning or statistical models (185). In the majority of studies, gram-negative bacteria were sequenced using Illumina© platforms. In these studies, phenotypic AMR profiles were predicted from WGS sequencing data with >90% accuracy. However, there was an exception, where phenotypic trimethoprim and streptomycin resistance were predicted with 80% and 75%, specificity, respectively. In the case of gram-positive bacteria, resistance against cell-wall inhibitors and fluoroquinolones can be predicted using the WGS platform with 90% and 80% accuracy, respectively. (186).

One of the objectives of the AMR surveillance programs discussed in the previous section is to understand the relative risk of AMR dissemination across the One Health Continuum. Mobile genetic elements represent a unique challenge as they play a significant role in the dissemination of ARGs among different ecosystems (13). For example; the majority of ARGs in *E. coli* are associated with plasmids (e.g. IncF, IncI, IncH, and IncA/C (187). These plasmids can carry multiple ARGs including ESBL genes, carbapenem resistance genes, genes encoding for aminoglycoside-modifying enzymes, and plasmid-mediated quinolone resistance (PMQR) genes. The global spread of cephalosporin ARG (*blaCTX-M-15*) in humans is associated with IncFII plasmid in ST131 and ST405 *E. coli* clones (188, 189). The phylogenetic analysis had shown that *E. coli* harboring IncI2 plasmids that carry ESBL genes can be easily transferred to pathogenic *Salmonella* strains and other Enterobacterales (190, 191). The resolution of complex MGEs structures through hybrid genome assemblies can be help predict the potential transmission of ARGs across One-health sectors.

The physical linkage of ARGs on MGEs promotes the co-selection of resistance determinants (192). Whole genome sequencing is an important tool to study these

associations among different ARGs. It has been known that selective pressure by heavy metals in the ecosystem co-selects ARGs associated with antimicrobial compounds, along with heavy metal resistance determinants (193). An example of this is the identification of physical linkage between *tcrB* (responsible for copper resistance), the *vanA*-cluster (glycopeptides resistance), and *ermB* in the genome of *E. faecium* isolated from copper-exposed pigs (194). Similar to this, exposure to one antimicrobial may also co-select for other ARGs that are on the same MGEs. An example of this is identification of six different plasmids harboring carbapenem ARGs along with other ARGs associated with resistance to aminoglycosides, fosfomycin, and sulphonamides (195). The reconstruction of the complete genomic structure of these MGEs through the WGS data provided significant information on the spread of AMR in One-health sectors. Moreover, in some cases, ARGs located on different MGEs may also be together during conjugation, referred to as co-transfer. One example is co-transfer of the colistin resistance gene (*mcr-1*) and *ESBLs* gene in *E. coli*. These genes were found on separate different plasmids in donor *E. coli* strains and were transferred together to a recipient under the selective pressure of ampicillin (196). The above-defined events where AMR genes co-transferred in a single conjugation event are the big challenges faced by clinicians as they are left with few antimicrobials that can be used for disease treatment. Large databases generated from WGS provide more in-depth knowledge of how this event occurs.

Through WGS analysis, the route of transmission and evolutionary relationship between isolates from different settings was investigated. For instance; Wang et al. performed a whole genome phylogenetic analysis of carbapenem-resistant *E. coli* (*blaNDM*-positive *E. coli*), identifying clonal commonality among strains from chicken farms, slaughterhouses,

supermarkets, and humans. They identified *E. coli* isolates belonging to MLST types ST10 and ST156, suggesting that the presence of these resistant clones in humans, livestock, and the environment could play an important role in the spread of AMR across different sectors (197).

Access to WGS data through public databases is possible and can be utilized to determine the epidemiology of ARG. One such example is the identification of the plasmid-borne mobile colistin resistance (*mcr-1*) gene in *E. coli* recovered from animals, retail meat, and humans from China in 2016 (198). Soon after this report, *mcr-1* was annotated in *E. coli* genomes and metagenomes that were already present in public databases (199, 200). Surveillance studies conducted in England and Denmark in 2016 used WGS to study the prevalence of *mcr-1* in human and animal settings. The sequence analysis showed that the risk of *E. coli* carrying *mcr-1* is low in Europe across a One-health continuum (201). Moreover, a new *mcr*-type gene, *mcr-5* identified as part of Tn*6452* transposon in *E. coli* from livestock and humans suggests that HGT along the food chain may be occurring (202, 203). Another example is the identification of *fexA* and *B*, novel exporter genes that confer resistance to chloramphenicol and florfenicol (204). In human medicine, chloramphenicol is effective against vancomycin-resistant enterococci (204), whereas florfenicol is used to treat respiratory tract infections in cattle and pigs (205). Whole genome sequencing analysis showed a high degree of sequence similarity between *fexA* and *B* genes found in *E. faecalis* and *E. faecium* from humans and livestock origin (206). Furthermore, these genes were frequently found on MGEs along with the oxazolidinone resistance determinants (*poxtA* and *optrA*), illustrating the important role that MGEs play in the dissemination of ARGs across different environments (207, 208).

In this thesis, we characterized indicator bacteria both phenotypically and genotypically. A whole genome sequencing pipeline had been established for the characterization of *Enterococci* and carbapenem-resistant bacteria under One-health concept.

1.5 Study objectives

The current thesis was built on three studies with the following objectives; **Study 1:** In a previous AMR surveillance study conducted by our lab at AAFC in Alberta, *Enterococcus spp* was used as an indicator bacteria. This study had the following objectives; the first study conducted was built on the recovered *Enterococcus hirae* recovered from the One-health sectors. The first objective of this study was to characterize *Enterococcus hirae* recovered under the One-health continuum. The second objective was to investigate the genomic relatedness of *E. hirae* across the environmental continuum through comparative genomic analysis. The third objective of this study was to identify genes that may account for the predominance of *E. hirae* within beef cattle production systems using pan-genome analysis. **Study 2:** In this study was based on *Enterococcus faecium* and *Enterococcus faecalis* isolates recovered from One-health sectors from four provinces (Alberta, British Columbia, Ontario, and Quebec) in Canada. This study had the following objectives: the first objective was to investigate the prevalence of *Enterococcus* species recovered from swine feces. The second objective was to investigate ARGs, virulence genes, and mobile genetic elements in *E. faecium* and *E. faecalis* isolates recovered from the One-health sectors. The third objective of this study was to conduct a comparative genomic analysis on *E. faecium* and *E. faecalis* sourced across various sectors of the One Health continuum. **Study 3:** This study was based on the recovery of carbapenem-resistant bacteria from the beef production system. It had the following objectives: the first objective was to

investigate the occurrence of carbapenem-resistant bacteria in the beef production system in Alberta. The second objective was to characterize carbapenem-resistant isolates at the phenotypic and genotypic levels. The third objective was to investigate the genomic relatedness among isolates recovered from beef-production systems and other One-health sectors through comparative genomic analysis.

Chapter 2: Genomic Characterization of *Enterococcus hirae* From Beef Cattle Feedlots and Associated Environmental Continuum.

Frontiers in microbiology vol. 13 859990. 27 Jun. 2022, doi:10.3389/fmicb.2022.859990.

Author's Contributions

RZ, and TAM designed the study; SH, CWB arranged for collection of feedlot samples, metadata and antimicrobial use data; RB, SRC, RZ and AC isolated and characterized enterococci; RB performed AST and RZ, AC, RB isolated genomic DNA; GVD, provided and managed the bioinformatics cluster facility and bioinformatics tools; SZ, RZ analyzed sequence data; SZ generated figures and analyzed overall data/results and wrote first draft of the manuscript; TAM, AZ provided funding and supervision. All authors participated in editing and reviewing the manuscript and approved the final manuscript.

2.1 Abstract

Enterococci are commensal bacteria of the gastrointestinal tract of humans, animals and insects. They are also found in soil, water and plant ecosystems. The presence of enterococci in human, animal and environmental settings make these bacteria ideal candidates to study antimicrobial resistance in the One-health continuum. This study focused on *Enterococcus hirae* isolates (n= 4601) predominantly isolated from beef production systems including bovine feces (n=4117, 89.5%), catch-basin water (n=306, 66.5%), stockpiled bovine manure (n=24, 0.5%), and natural water source near feedlots (n=145, 32%), and few isolates from urban wastewater (n=9, 0.2%) denoted as human-associated environmental samples. Antimicrobial susceptibility profiling of a subset (n=1319) of *E. hirae* isolates originating from beef production systems (n=1308) showed

high resistance to tetracycline (65%) and erythromycin (57%) with 50.4% isolates harbouring Multidrug resistance, whereas urban wastewater isolates (n=9) were resistant to nitrofurantoin (44.5%) and tigecycline (44.5%) followed by linezolid (33.3%). Genes for tetracycline (*tetL, M, O , S/M* and *O/32/O*) and macrolide resistance *erm(B)* were frequently found in these beef production isolates. Antimicrobial resistance profiles of *E. hirae* isolates recovered from different environmental settings appeared to reflect the kind of antimicrobial usage in beef and human sectors. Comparative genomic analysis of *E. hirae* isolates showed an open pan-genome that consisted of 1427 core genes, 358 soft core genes, 1701 shell genes and 7969 cloud genes. Across species comparative genomic analysis conducted on *E. hirae, Enterococcus faecalis* and *Enterococcus faecium* genomes revealed that *E. hirae* had unique genes associated with vitamin production, cellulose and pectin degradation, traits which may support its adaptation to the bovine digestive tract. *E. faecium* and *E. faecalis* more frequently harboured virulence genes associated with biofilm formation, iron transport, and cell adhesion, suggesting niche specificity within these species.

2.2 Introduction

Antimicrobial resistance (AMR) is recognized as one of the major global health challenges of the 21st century. The interconnected microbiomes between humans, animals and the environment contribute to the emergence, acquisition and spread of AMR (209). A One-health approache provides an in-depth knowledge of the evolution of AMR by focusing on significant biological elements including microorganisms involved in emergence and dissemination, hosts (human and animals) and associated environments that may facilitate AMR spread (210). Gram positive enterococci are core members of the gastrointestinal

microbiota of humans and animals and are frequently isolated from soil and water (99, 100). Enterococci often carry antimicrobial resistance genes (ARGs) as they compete within complex microbial communities and are exposed to antimicrobials in clinical settings and during livestock production (211, 212). Furthermore, depending on the species, enterococci exhibit intrinsic resistance to several antibiotics including cephalosporins, anti-staphylococcal penicillins, aztreonam, aminoglycosides, lincosamides and streptogramins (111). Enterococci are typically commensals, but they can cause nosocomial infections in humans including septicemia, endocarditis, and urinary tract infections (213). There are over 50 species of enterococci with *E. faecalis* and *E. faecium* most frequently linked to human infections. Occasionally, other species including *E. hirae, E. avium, E. durans, E. gallinarum, E. casseliflavus and E. raffinosus* may also be associated with infections in people (214-216). Due to their widespread occurrence and persistence in the environment, enterococci are considered indicators of fecal contamination (103, 104) and also serve as key indicator bacteria for AMR surveillance systems in humans and animals (105, 106).

Studies have indicated that *E. faecium* and *E. faecalis* are more prevalent in human-associated environments, whereas *E. hirae* are prevalent in beef cattle production systems (217). *E. hirae* only accounts for 1% of enterococcal infections in humans (218), and is mainly linked to pyelonephritis (219-221), endocarditis (222-224) and biliary tract infections (225, 226). The choice of treatment for *E. hirae* infections is similar to *E. faecalis* and *E. faecium* where ampicillin, gentamicin and vancomycin are commonly used (219).

The focus of this study was to investigate the genomic relatedness of *E. hirae* across the environmental continuum, and to identify the genetic nature of AMR in *E. hirae*.

Furthermore, we applied a pan-genome analysis to identify genes that may account for the predominance of *E. hirae* within beef cattle production systems.

2.3 Methodology

2.3.1 Bacterial isolates

A total of 8430 *Enterococcus* strains were isolated in a One-health surveillance study from different segments of the environmental continuum using samples collected from beef production systems (i.e., feedlot cattle feces, catch-basin water, manure), natural water sources, urban wastewater, and human clinical samples (217). Bovine fecal samples came from four feedlots in southern Alberta for two years (March 2014-April 2016).

Wastewater samples came from catch basins accumulating runoff from the feedlots under study. Natural surface water samples came from up-stream and down-stream regions of feedlots. Urban wastewater samples and clinical isolates were representatives of human ecosystems. Urban wastewater samples came from two wastewater plants located in southern Alberta. *Enterococcus* spp. recovered from patients with clinical infections were obtained through the Division of Medical Microbiology, Calgary Laboratory Services (now Alberta Precision Laboratories, Alberta Health Services) (217). This study focuses on *Enterococcus hirae*, collected as the most prevalent species from beef production system (n=4601 isolate) (217). Figure 2.1 represents the prevalence of *E. hirae* isolates in the sampled sources.

Enterococci were recovered in parallel from two different media types including Bile Esculin Azide (BEA) agar without antibiotic and BEA supplemented with 8 µg/mL erythromycin, followed by species identification. *E. hirae* were identified via multiplex

PCR targeting *groES-EL*, and muramidase genes (217). As *E. hirae* was absent among clinical *Enterococcus* isolates (n=1892; Figure 2.1), complete genomes (n=3) of clinical *E. hirae* were retrieved from NCBI database for comparative genomic analysis (Appendix 1; Table S1).

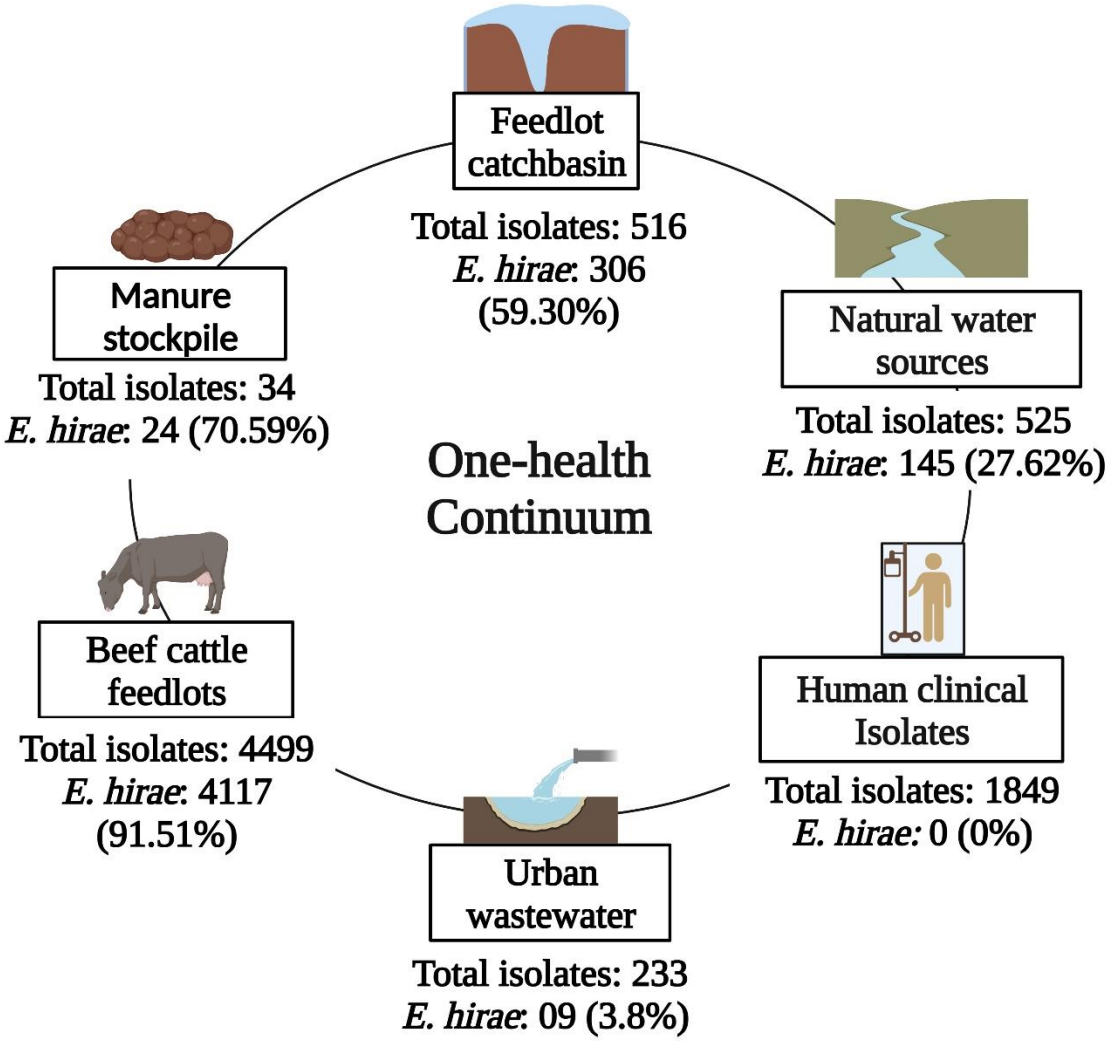

Figure 2.1 Prevalence of *Enterococcus hirae* (n= 4601) identified across One-health sectors. The beef production system isolates were recovered from bovine feces, manure stockpile and feedlot catch basin. The natural water isolates were recovered from up-and downstream natural water sources located in close proximity to the beef production system. Human clinical isolates were recovered from clinical samples from sterile and nonsterile sites. Urban waste water isolates were recovered from wastewater treatments plants located in Alberta.

2.3.2 Antimicrobial susceptibility testing

Antimicrobial susceptibility testing was performed on a randomly selected subset (n=1319, 29%) of *E. hirae* isolates using the disk diffusion method, as per the Clinical and Laboratory Standards Institute (CLSI) documents M02-A12 and M100-S24. A panel of twelve antibiotics was used for testing based on their common usage for treatment of human enterococcal infections that included those drugs of critical importance (levofloxacin, linezolid, quinupristin/dalfopristin, teicoplanin, vancomycin and tigecycline), high importance (erythromycin, ampicillin, gentamicin and streptomycin) and medium importance (nitrofurantoin, and tetracycline). *Staphylococcus aureus* ATCC 25923 and *E. faecalis* ATCC 29212 were used as reference quality controls (217). The BioMic V3 imaging system (Giles Scientific, Inc., Santa Barbara, CA, USA) was used to read zones of inhibition. Isolates were categorized based on CLSI interpretive criteria, except for tigecycline for which EUCAST interpretive criteria (The European Committee on Antimicrobial Susceptibility Testing, 2014) were used.

2.3.3 Whole-genome sequencing

Whole genomic sequencing of a subset of *E. hirae* isolates (n=286), including isolates originating from bovine feces (n=168), feedlot catch-basin (n=62), bovine manure stockpiles (n=8), natural water sources (n=42), and urban wastewater (n=7) was performed using next-generation sequencing technology. Briefly, genomic DNA was extracted using the DNeasy Blood and Tissue Kit (Qiagen, Montreal, QC, Canada) with modifications (217), followed by DNA quality assessment and quantification using a Nanodrop 2000 spectrophotometer and a Qubit Fluorometer with PicoGreen (Thermo Fisher Scientific, Mississauga, ON, Canada). Isolates were sequenced on an Illumina MiSeq platform using

the MiSeq Reagent Kit V3 to generate 2 × 300 bp paired-end reads. Raw read FASTQ files were assessed for the quality of sequence data using FastQC (Galaxy v.0.72+galaxy1) (227) and *de novo* assemblies were performed using Shovill v.3.11.1 (228). Assembled contigs were then annotated by Prokka to identify all gene-coding sequences (183).

2.3.3.1 AMR determinants, virulence and plasmid detection

Assembled genomes were screened for the presence of AMR determinants, virulence genes and plasmids using ABRicate (https://github.com/ tseemann/abricate/) against the NCBI Bacterial Antimicrobial Resistance Reference Gene database (NCBI BioProject ID: PRJNA313047), VirulenceFinder (PMID: 34850947), and PlasmidFinder databases (229), respectively. Intact prophage were identified using PHASTER tool (230).

2.3.3.2 Comparative genomic analysis

A total of 289 genomes including 286 assembled genomes from this study and 3 complete *E. hirae* genomes of clinical isolates retrieved from NCBI database (strain: 708, accession: NZ_CP055232.1; strain: FDAARGOS_234, accession: NZ_CP023011.2 and strain: 13344, accession: NZ_CP055229.1) were subjected to phylogenomic analysis. A core-genome phylogenomic tree was constructed using the (SNVPhyl v 1.0) pipeline (217). Briefly, all paired end reads were mapped against reference genome *E. hirae* (strain R17; Genbank accession: CP015516.1) to produce read pileups (SMALT v.0.7.5; https://www.sanger.ac.uk/tool/smalt-0/). The read pileups were evaluated for mapping quality (minimum mean mapping quality score of 30), coverage cut offs (15X minimum depth of coverage), and a single nucleotide variant (SNV) abundance ratio of 0.75 to generate a multiple sequence alignment of SNV containing sites. The final maximum likelihood based phylogeny was generated by PhyML using unfiltered SNV alignment.

Phylogenomic trees and associated metadata were visualized using Interactive Tree Of Life (iTOL) v5 tool (231).

Comparative genome analysis was done using the Roary v3.12.0 pipeline with default parameters (232). Genes identified by Prokka were used to construct pan-genomes. A pan-genome of 289 *E. hirae* isolates was reconstructed to identify core and accessory genes present in *E. hirae*. Furthermore, comparative analysis was performed between *E. faecium* and *E. faecalis,* which are predominantly associated with humans infections, and *E. hirae.* For this purpose, a small subset of *E. hirae* isolates (n=16) representative of the various sources and phylogenetic clades were randomly selected. Similarly, a subset of *E. faecium* (n=26) and *E. faecalis* (n=24) isolates were randomly selected on the same bases from our previous study (217). The Phandango interactive viewer tool (233)was used to interpret pan-genome data obtained from Roary analysis. This tool utilizes two of the Roary output files: one is a gene absence and presence matrix file that creates a heat map based on the number of genes present or absent in each isolate and a Newick-formatted tree file of accessory genomes used to plot a relatedness dendrogram of the accessory genes present in all isolates.

A pan-genome plot was generated using ggplot2 package of R Studio V.V. 1.4.1103 (R Studio Inc, Boston, MA, USA) based on two Roary output files (the number of conserved genes and the number of total genes). The number of conserved genes represented the size of the core genome. The number of total genes represented both the core and accessory genomes, creating a curve based on the pan-genome completeness. The pan-genome of an organism is considered ‘closed’ if the curve is predicted to plateau, or ‘open’ if the curve

is predicted to continue to rise. In contrast to a closed genome, the number of new gene families in an open genome increases with the inclusion of new genomes in the analysis.

Discriminatory genomic signatures between *E. hirae*, *E. faecium* and *E. faecalis* were identified using Neptune v1.2.5 with default parameter (234)The signature discovery process using Neptune identifies sequences that are sufficiently common to a group of target sequences (inclusion group) and sufficiently absent from non-targets (exclusion group) using probabilistic models. Analyses was done using *E. hirae* genomes as the inclusion group and *E. faecium* and *E. faecalis* as independent exclusion groups, respectively. The genomic signature found in ≥90% of isolates in the inclusion group were selected and annotated using Prokka (183).

2.4 Results

2.4.1 Antimicrobial susceptibility testing

Phenotypic susceptibility testing was conducted on 1319 *E. hirae* isolates originating from bovine feces, feedlot catchbasin water, stockpiled bovine manure and natural and urban wastewater sources. (Figure 2.2). Fifty-one different resistance profiles were identified with the most frequent being resistant to doxycycline and erythromycin (364/1319, 27.6%) followed by resistance to doxycycline alone (242/1319, 18.3%) (Appendix 1; Table S2). Across all tested isolates, 14.1% (186/1319) were Multidrug resistant (resistant to ≥ 3 tested antimicrobials). Antimicrobial susceptibility profiles of all tested isolates are presented in Appendix 1; Figure S1.

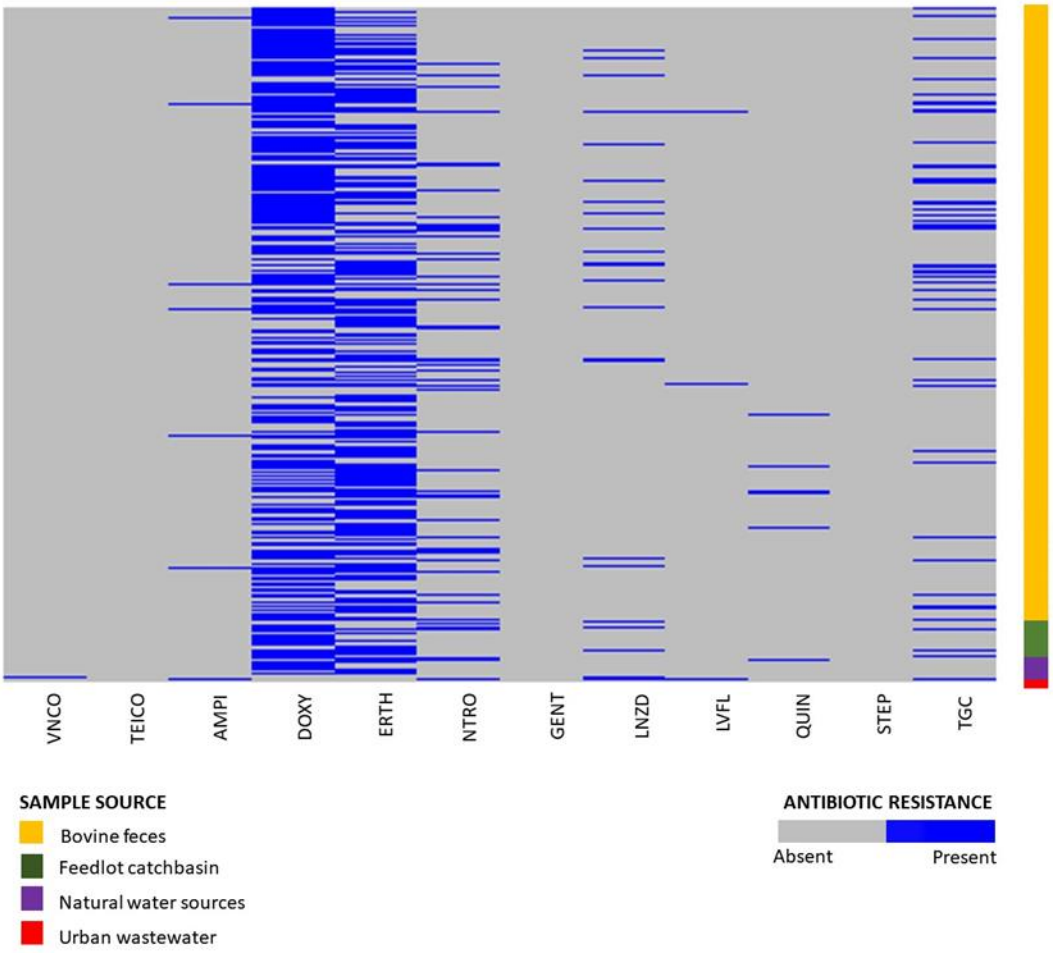

Figure 2.2 Phenotypic resistance profiles of *Enterococcus hirae* isolated from beef production systems (n=1264). This includes isolates from bovine feces, feedlot catch basin, natural water source (n= 45) and urban wastewater sample (n=9). (VANCO=Vancomycin, TEICO= Teicoplanin, AMPI=Ampicillin, DOXY=Doxycycline, ERTH= Erythromycin, NTRO= Nitrofurantoin, GENT= Gentamicin, LNZD= Linezolid, LVFL= Levofloxacin, QUIN= Quinolones, ATEP= Streptomycin, TGC= Tigecycline).

Isolates recovered from BEA plates without erythromycin: Out of 1319 total isolates tested for antimicrobial susceptibility, 666 isolates were recovered from BEA plates without erythromycin. From these, isolates originating from beef production systems (i.e., bovine feces, catch basin and stockpiled bovine manure; n=632) exhibited a high prevalence of resistance to tetracycline (376/632, 59.4%), followed by macrolides (200/632, 31.6%), nitrofurantoin (102/632, 16.1%), tigecycline (76/632, 12.0%), linezolid (40/632, 6.32%), ampicillin (9/632, 1.42%), quinupristin/dalfopristin (8/632, 1.26%), vancomycin (1/632, 0.15%) and teicoplanin (1/632, 0.15%) (Appendix 1; Figure S1).

The natural water source isolates recovered from BEA plates without antibiotics (n=28) also showed a high prevalence of tetracycline resistance (22/28, 78.5%), followed by macrolides (8/28, 28.5%), nitrofurantoin (5/28, 17.8%) and tigecycline (1/28, 3.57%). Resistance against quinupristin/dalfopristin, linezolid, ampicillin, and fluoroquinolones was not detected (Appendix 1; Figure S1).

Among 233 *Enterococcus* spp. isolates recovered from urban wastewater [16], only nine isolates were identified as *E. hirae*. Six of those were recovered from media without erythromycin. Four of the six isolates exhibited resistance to tigecycline (4/6, 66.6%) followed by nitrofurantoin (3/6, 50%), linezolid (3/7, 48.85%), fluoroquinolones (2/6, 33.3%), vancomycin (1/6, 16.6%) and ampicillin (1/7, 14.2%) (Appendix 1; Figure S1).

Isolates recovered from BEA plates with erythromycin: A total of 652 isolates from erythromycin plates were selected for phenotypic antimicrobial testing. Of these isolates 632 originated from beef production (i.e., bovine feces, catch basin and stockpiled bovine manure). Tetracycline resistance (437/632, 69%) was the most prevalent resistance in the beef isolates from production systems, followed by macrolides (525/632, 83%), tigecycline

(67/632, 10.6%), nitrofurantoin (60/632, 9.5%), linezolid (33/632, 5.2%), quinupristin/ dalfoprstin (13/632, 2.0%), ampicillin (9/632, 1.42%), fluoroquinolones (5/632, 0.79%) and gentamicin (1/632, 0.15%) (Appendix 1; Figure S2).

The isolates recovered from natural water sources (n=17) showed high prevalence of resistance to macrolides (15/17, 88.2%) followed by tetracycline (13/17, 76.4%), tigecycline (2/17, 11.76%), quinupristin/ dalfoprstin (1/17, 5.88%), and nitrofurantoin (1/17, 5.88%). None of the isolates were resistant to linezolid, ampicillin, or fluoroquinolones (Appendix 1; Figure S2).

A total of three *E. hirae* isolates were recovered from urban waste water sources on erythromycin plates. Two of those isolates showed macrolide resistance (2/3, 66.6%), followed by tetracycline (1/3, 33.3%), nitrofurantoin (1/3, 33.3%), quinupristin/ dalfoprstin (1/3, 33.3%) and streptomycin (1/3, 33.3%). These isolates were sensitive to teicoplanin, ampicillin, vancomycin, gentamicin, tigecycline, fluoroquinolones and linezolid (Appendix 1; Figure S2). Overall 16.8% of isolates (110/652) recovered from erythromycin plates showed intermediate resistance to erythromycin.

2.4.2 Whole-genome sequencing

Of the *E. hirae* isolates tested for antimicrobial susceptibility, 286 randomly selected isolates were used for whole-genome sequencing. The size of *E. hirae* genomes as interpreted from the assembled sequence read data ranged from 2307753 bp to 3200875 bp with GC content of 36.7%. Detailed assembly statistics are provided in Appendix 1; Table S3.

2.4.3 AMR determinants, virulence and plasmid detection

AMR determinants: Assembled genomes (n=286) were screened for the presence of AMR determinants using the Abricate tool v.1.0.1 (https://github.com/tseemann/ABRICATE) along with the NCBI AMR gene database. Ten different ARGs, including aminoglycosides ARGs *aac(6')-Iid, ant(6)-Ia* and *aph(3)-III*, streptothricin *sat4,* tetracycline ARGs *tet (L, M, O, S/M, (O/32/O)* and macrolide ARG *erm(B)* were identified across the examined genomes (Figure 2.3). Overall, nineteen different resistance genotypes were identified with the most frequent being *aac(6')-lid-tet(L)-erm(B)* (87/286, 30.41%) followed by *aac(6')-lid-tet(L)- tet(M)-erm(B)* (64/286, 22.37%) (Appendix 1; Table S4). The aminoglycoside resistance gene *aac(6')-lid* was identified in all but two of the *E. hirae* genomes (284/286, 99.30%).

Of the 286 sequenced isolates, 238 were recovered from beef production systems (i.e., bovine feces, feedlot catchbasin and stockpiled bovine manure). *tet(L)* (199/238, 83.61%) was the most prevalent ARG identified in these isolates, followed by *erm(B)* (179/238, 75.21%), and the tetracycline resistance genes, *tet (M)* (73/238, 30.67%), *tetO* (36/238, 15.12%), *tet(O/32/O)* (13/238, 5.46%) and *tet(S/M)* (03/238, 1.26%).

Similar to beef production system isolates, *E. hirae* isolates recovered from natural water sources located near feedlots showed a high prevalence of *tet(L)* (38/41, 92.68%) followed by *erm(B)* (27/41, 65.85%). Occasionally, *tet(M)* (4/41, 9.75%), *tet(O)* (4/41, 9.75%) and *tet*(*O/32/O*) (2/41, 4.87%) were also present in these isolates.

Among seven *E. hirae* isolates recovered from urban wastewater samples, the streptomycin resistance gene *ant(6)-la* was present in two isolates (2/7, 28.57%) . The

kanamycin/neomycin *aph(3')-III* and streptothricin *sat4* resistance genes were both found in a single urban wastewater isolate (1/7, 14.28%). Tetracycline resistance gene(s) were not found in any of these isolates, whereas *erm(B)* was only detected in one isolate (1/7, 14.28%).

tet(L) and *erm(B)* were found together in 63.63% of total isolates (182/286) indicating a strong correlation. Similarly *tet (L)* and *tet (M)* coexisted in 24.12% of isolates (69/286) and in most cases were found on a same the contig (60/69, 87%) in assembled genomes.

In the studied *E. hirae* isolates, genotypes generally associated with quinolone resistance (i.e., presence of quinolones resistant gene (*qnr*) or DNA gyrase and DNA topoisomerase IV genes mutations) and linezolid resistance (i.e., mutations in the 23S ribosomal RNA gene or presence of resistance genes including *cfr*, *cfrB*, *optrA* and *poxtA*) were not identified.

Virulence factors: Within the 286 *E. hirae* isolates, we identified nine different virulence genes associated with biofilm formation *(bopD)*, capsular polysaccharides biosynthesis *(cpsA, cpsB, cap8E)*, hyaluronic acid production (*hasC*), proteolytic activity/chaperones *(clpP)*, fibrinogen adhesions protein *(fss3)*, bile salt hydrolase *(bsh)*, and listeria adhesion protein *(lap)*. All isolates carried *cap8E*, *clpP*, *cpsA*, *cpsB*, *bopD*, and *lap* genes, while *hasC*, *bsh* and *fss3* were found in 98.95% (283/286), 89.86% (257/286) and 2.44% (7/286) of total isolates. These genes were also identified in publicly available clinical *E. hirae* genomes from humans. Pili protein-encoding gene *ebpC* was only detected in one of the human clinical isolates retrieved from NCBI. Detailed information is provided in Appendix 1; Table S5.

Plasmid identification: Among all *E. hirae* isolates (n=286), 16% carried plasmids. Seven different plasmids (rep1, rep2, rep11, rep17, rep18, repUS7 and repUS12) were identified. Among these, rep2 and rep17 were recovered from all sample types except urban wastewater. In contrast, rep1, rep18, and repUS7 were recovered from urban wastewater samples. Two out of fifteen rep17 plasmids carried *erm(B)*, whereas one out of twenty rep2 plasmids carried *tet(L).* Of six repUS12 plasmids, five carried *tet (L)* and were recovered from bovine feces.

Prophage identification: A total of 30 genomes were randomly selected from all sample sources to identify bacteriophage using PHASTER. All isolates contain at least one prophage ranging from 7 to 48 kb in size. Twenty-four intact prophage sequences were identified, with 95.5% identified as members of the family Siphoviridae. None of the identified prophages harboured ARGs.

2.4.4 Comparative genomic analysis

Core genome phylogenomic analysis was conducted on the 286 *E. hirae* isolated in this study and the three *E. hirae* genomes retrieved from the NCBI public sequence repository. The *E. hirae* isolates clustered into six different clades, with no obvious segregation by source (Figure 2.3).

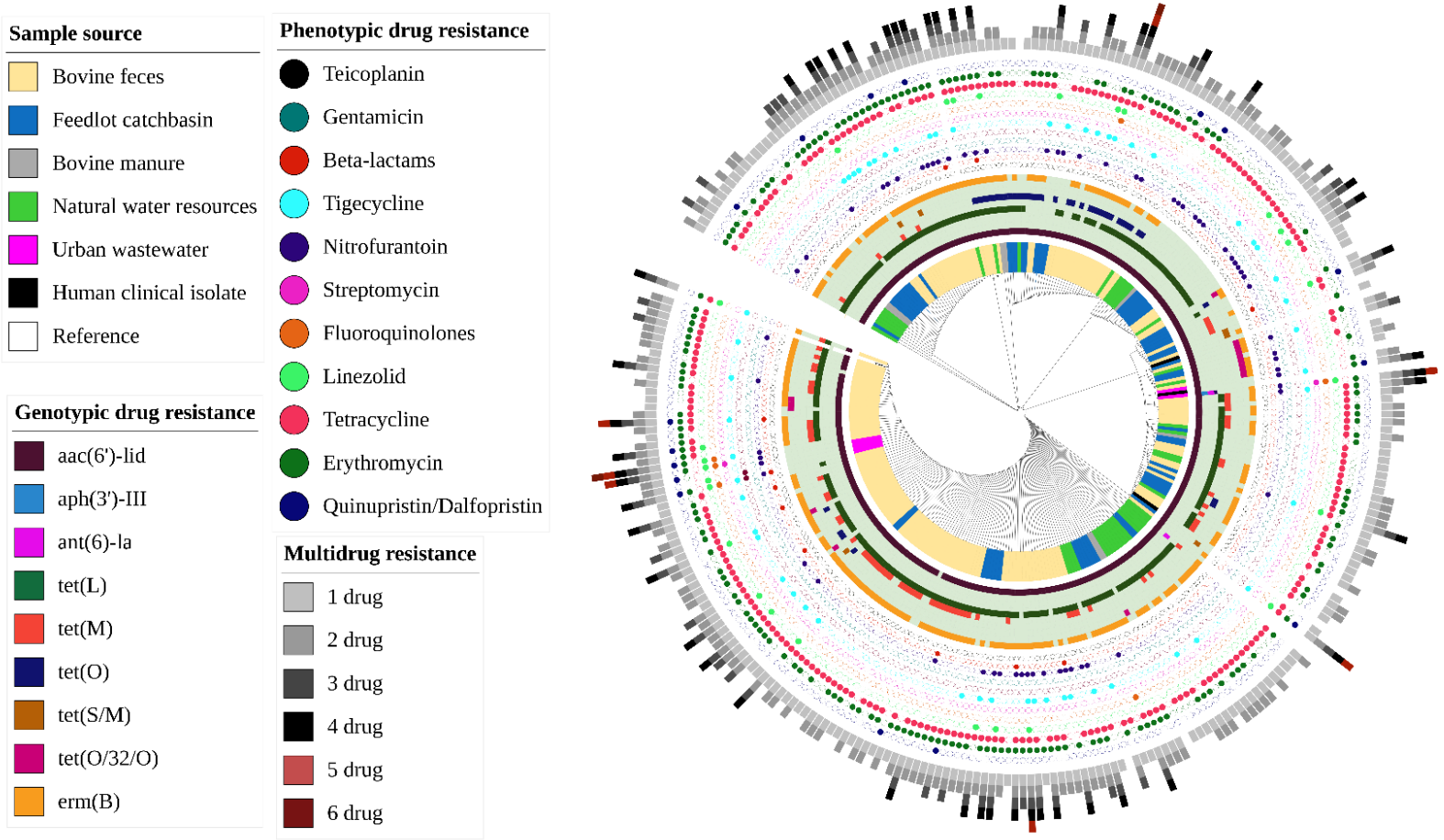

Figure 2.3 Core genome phylogenetic tree based on analysis of single-nucleotide polymorphisms (SNPs) of *Enterococcus hirae* genomes (n = 291) isolated from different environmental settings including beef production systems and human-related isolates. The genomes were compared using *E. hirae* OG1RF genome (GenBank accession # NZ_CP015516.1 /CP015516.1) as a reference

Pan-genome analysis of *E. hirae* isolates identified 1427 core genes (99% to 100% of strains), 358 soft core genes (95% to 99% of strains), 1701 shell genes (15% to 95% of strains) and 7969 cloud genes (0% to 15% of strains) (Figure 2.4). The pan-genome of *E. hirae* is open as the number of accessory genes progressively increased with increasing genomes (Figure 2.5). In addition, the gene presence and absence heat map showed that the accessory genome constituted a large part of the pan-genome, indicative of a high level of genomic diversity within this species (Figure 2.6).

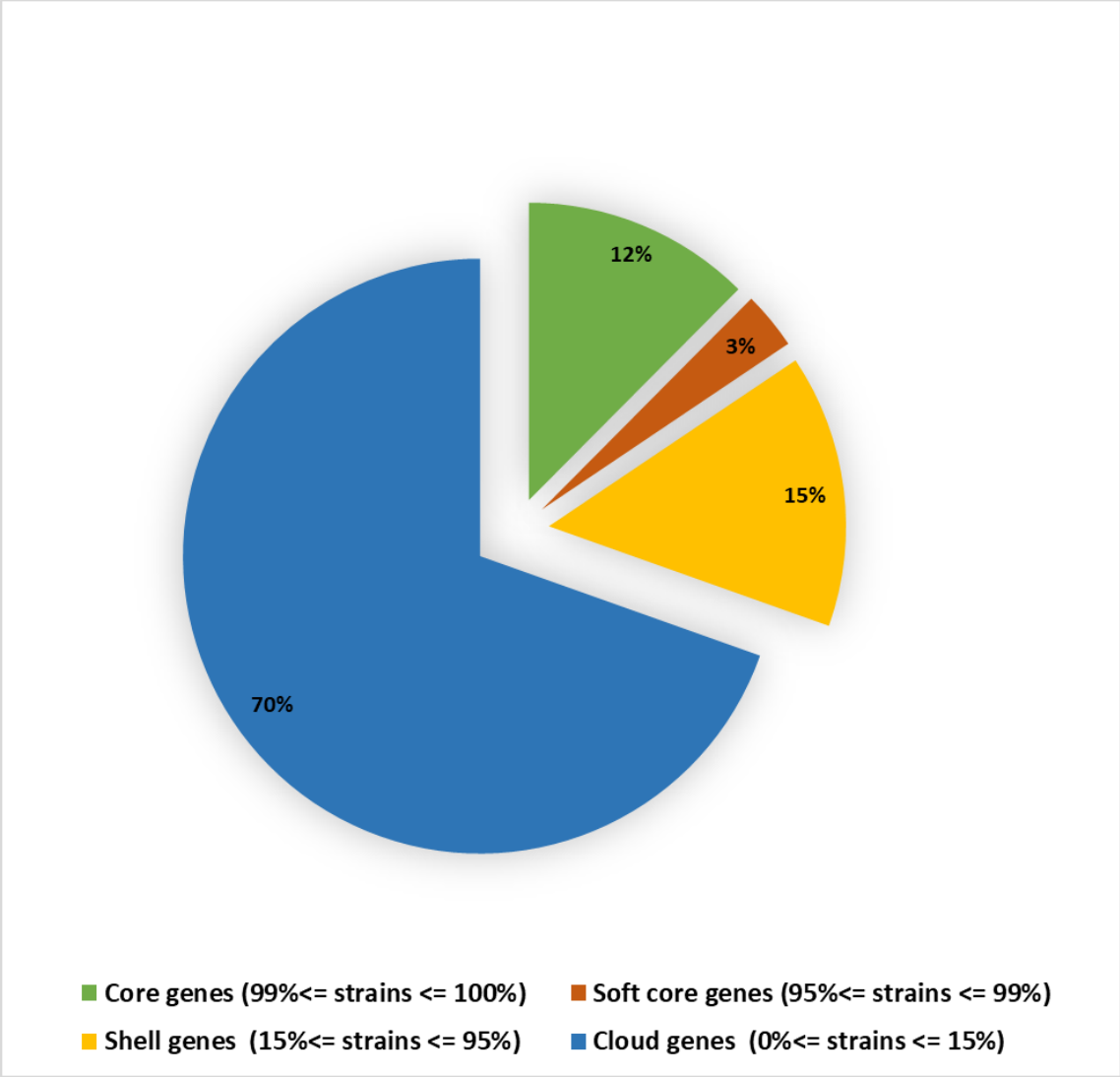

Figure 2.4 Genomic statistics and pan-genome estimation of 291 *Enterococcus hirae* genome. The pan-genome was constructed using Roary pipeline.

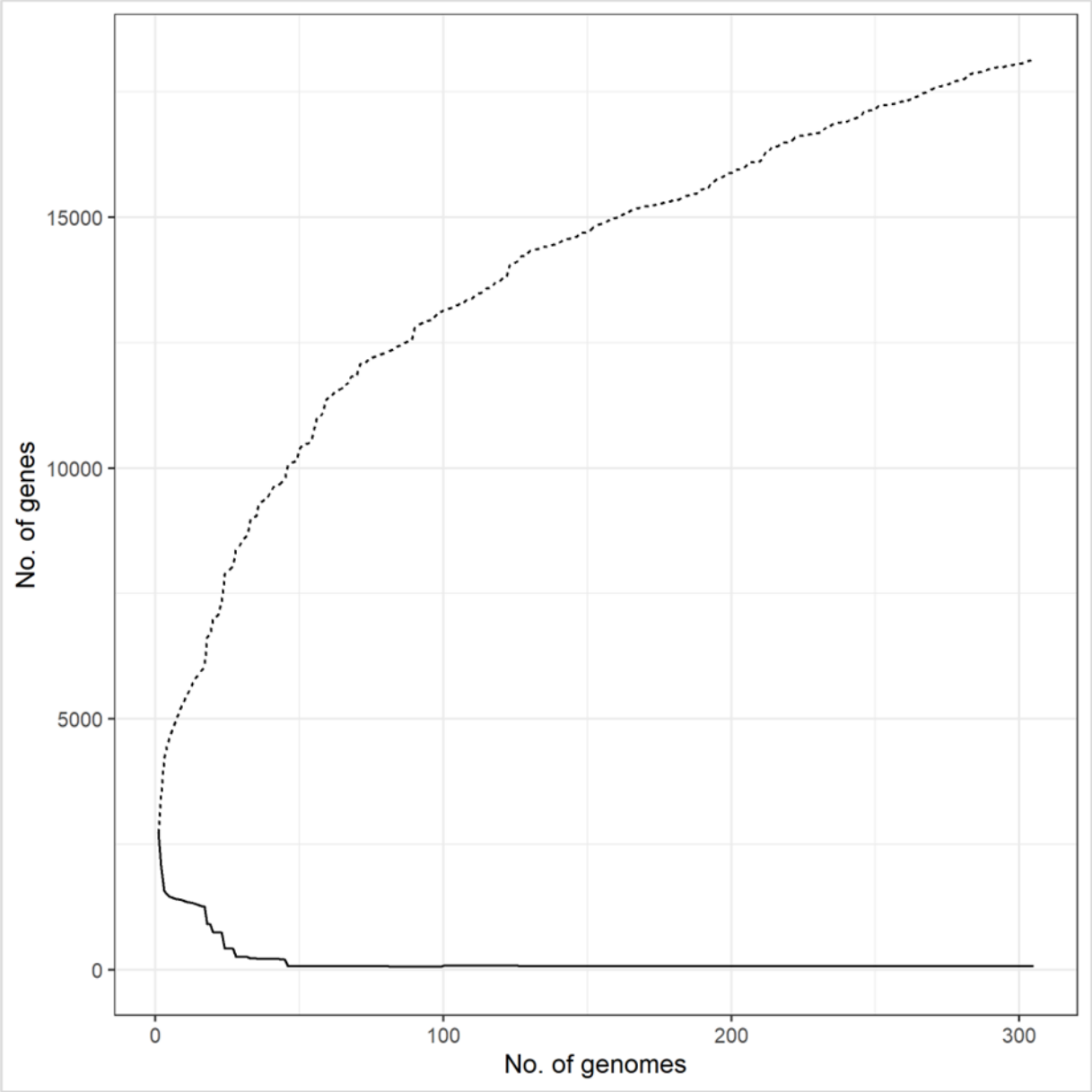

Figure 2.5 Development of pan- and core genomes of *E. hirae*. The graph illustrated the open nature of the pan-genome. the graph was constructed against total number of genes and number of genomes using output files generated by Roary.

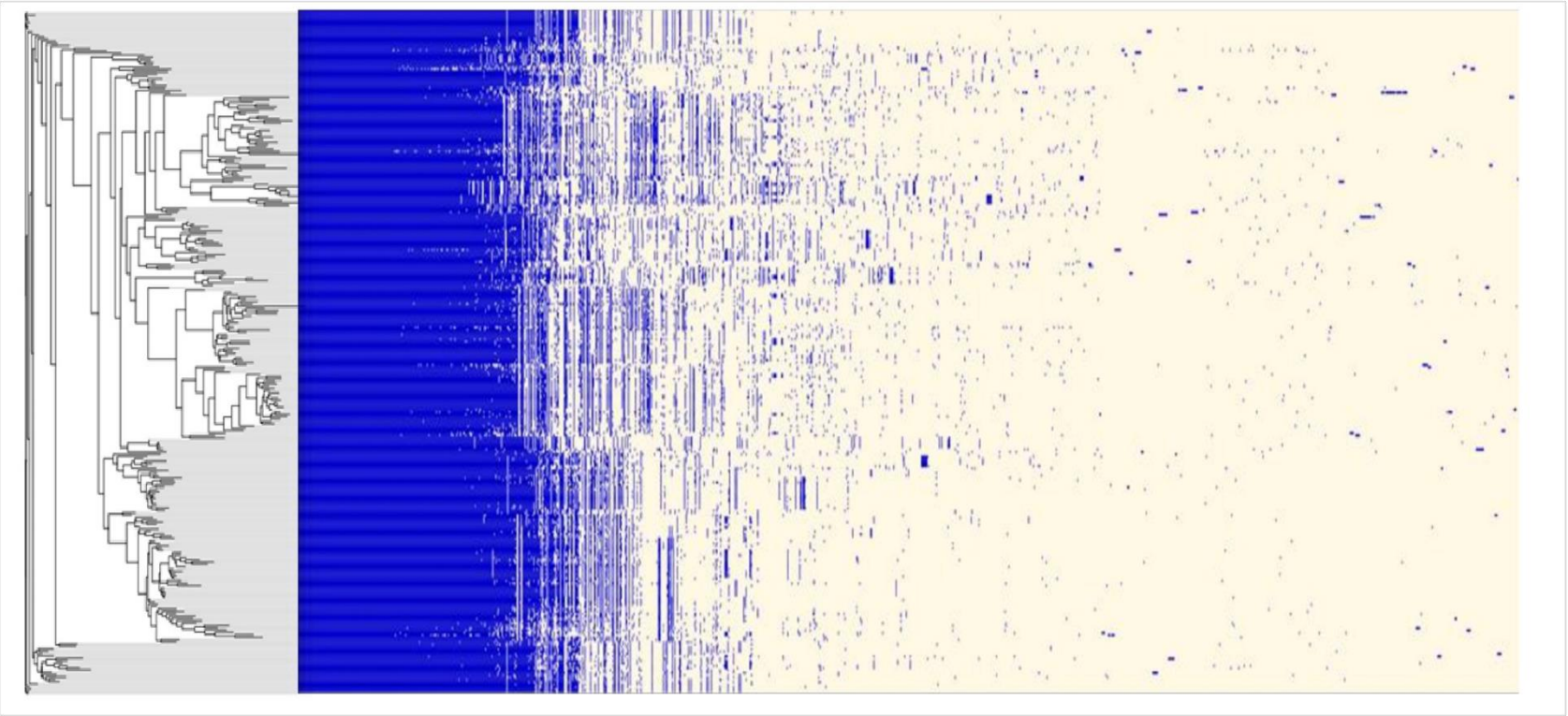

Figure 2.6 Heat map representing absence or presence of genes in *E. hirae* isolates. The tree generated was generated based on accessory genome using Phandango tool.

Cross-species comparative analysis of *E. hirae*, *E. faecium* and *E. faecalis* genomes, highlighted the genomic diversity within *Enterococcus* spp. as the total core genome shared between three species is small (Figure 2.7). Furthermore, it also demonstrate that each species distinct genomic traits indicated by gene absence and presence heat map (Figure 2.7).

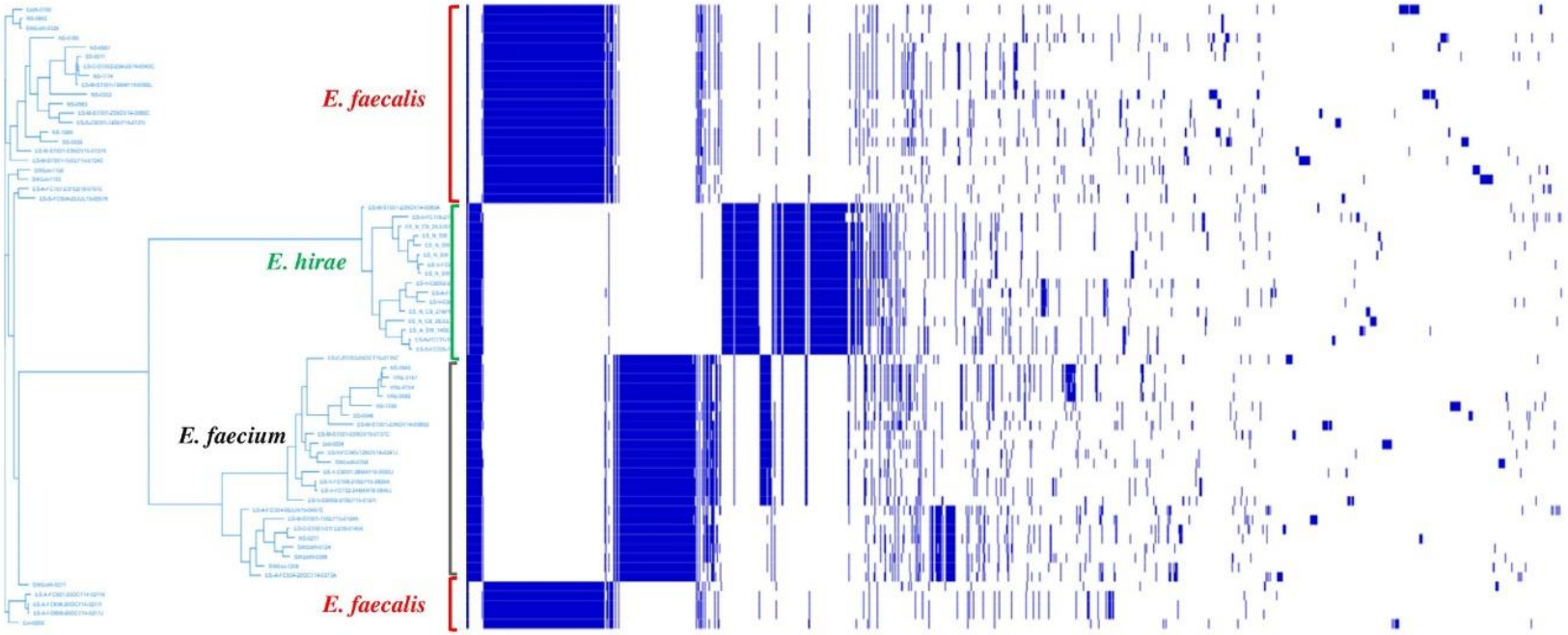

Figure 2.7 Heat map representing absence or presence of select genes and phylogenetic genetic tree generated from accessory genes of *Enterococcus hirae* (n=16), *Enterococcus faecium* (n=26) and *Enterococcus faecalis* (n=24).

A total of 1,069 discriminatory genomic signatures were recognized in *E. hirae* compared to *E. faecalis* (n=808) and *E. faecium* (n= 261). Most of these genomic signatures encoded for unknown hypothetical proteins (454/1069, 42.46%).

Neptune analysis is capable of identifying inter-species genomic variation, as discriminatory loci were identified in all three species. These loci coded for genes required for the synthesis of aromatic amino-acids including chorismate synthase (*aroC*), cyclohexadienyl dehydrogenase (*tyrC*), genes for sugar transport including permease (*yteP*), transcriptional regulator (*mtlR)*, lichenan permease IIC component (*licC*) and lipoprotein (*lipO*).

Signatures found exclusively in *E. hirae* genomes included genes associated with the phosphotransferase system for galactitol (*gatA, B* and *C*) and fructose (*fruA*, *frwA* and *frwD*), peptidoglycan synthesis (*rodA ftsW, mur E, F, J* and *Y*), teichoic acid synthesis (*tag H, G,* and *U*), coenzyme A biosynthesis (*coaD*), vitamin B12 synthesis (*nrdZ*), chitin degradation (*chiA*), capsule synthesis (*epsE, pglF, ywqD,* and *ywqC*), riboflavin synthesis (*ribBA*, *D, E* and *H*), vitamin B6 synthesis (*yvgN*), vitamin uptake transporter (*queT*), gamma-aminobutyric acid (GABA) production (*glsA2, gadC* and *amt*), cardiolipin biosynthesis (*clsA*), bacitracin export (*bceA* and *B*) , xenobiotic degradation (*nylA*) and iron transport (*yqgN, feuC, feuB, fepC* and *yfiY*).

Furthermore, genes that may be associated with antimicrobial resistance were also identified in *E. hirae* genomes such as those encoding for multidrug transporters (*marA, mepE msrR,* and *yxlF*), doxorubicin resistance (*drrA*), sulfonamide resistance (*bcr*) and

penicillin-binding protein (*pbp*). Early secreted antigenic target (ESAT) system genes (*eccC, essB*, and *esxA*) were also found in *E. hirae* genomes.

Compared to *E. hirae*, 160 and 944 discriminatory genomic signatures were identified in *E. faecium* and *E. faecalis* genomes respectively, with most of these genes associated with various aspects of cellular metabolism, such as molybdopterin biosynthesis (*mog*, *modB* and *modB*); cadmium, zinc and cobalt transport (*cadA*); copper export (*copY* and *copA*); phosphotransferase system for glucitol/sorbitol (*srlA*, *srlE* and *srlB*), sorbose (*sorB, sorF* and *sorA*), mannose (*manX*, and manZ), cellubiose (*celA*), mannitol (*mtlA*, *mtlF* and *mtlD*), maltose (*malX*), ascorbate (*ulaC* and *ulaA*) and beta-glucosides (*bglF*). In addition, genes encoding *E. faecalis* and *E. faecium* pathogenesis were also identified, such as biofilm formation genes (*brpA*, *icaA* and *lytR*); virulence genes, including unsaturated chondroitin disaccharide hydrolase (*ugl*); anthrax toxin regulator positive (*atxA*); hemin transport system (*hmuU* and *hmuT*); sialic acid TRAP transporter small permease (*siaQ*); carnitine transport system (*opuCB* and *opuCA*); arginine/ornithine system (*argR* and *arcD1*); genes encoding the adhesions, including gelatinase (*gelE*) and collagen (*cna*); and genes encoding for antimicrobial resistance, including penicillin-binding protein (PbpE, PbpX and PbpF), MDR protein (YkkC, YkkD and Stp), tetracycline repressor protein (TetR), fluoroquinolones export protein (Rv2688c), macrolide export protein (MacB) (Appendix 1; Table S7).

2.5 Discussion

Enterococci are ubiquitous Gram-positive bacteria. They colonize gastrointestinal tracts of most multicellular eukaryotic organisms including humans, animals and insects and aid in digestion and gut metabolic pathways (235-245). In addition, they are also found in food, plant and water ecosystems (246-251). Enterococci are remarkably resilient to broad pH ranges, temperature variation and osmotic pressure, traits that contribute to their broad distribution in nature (252-254). Resistant bacterial populations are selected by the exposure of commensal gut microorganisms such as enterococci to antimicrobials that are used for disease treatment and prevention (255). The ubiquitous nature of enterococci may facilitate the dissemination of antimicrobial resistance genes between different environments. For this reason, it is imperative to identify antimicrobial resistance determinants and their role in the spread of antimicrobial resistance (256). Here we focused on *E. hirae* isolates recovered from a One-health surveillance study (217). The genomic relatedness of *E. hirae* was examined across various sampling matrices of the continuum and AMR determinants that contribute to antimicrobial resistance were identified. Furthermore, we examined the genomic traits of *E. hirae* that may facilitate their growth in the cattle gut as compared with other human-associated *Enterococcus* species.

As described previously, *E. hirae* is highly prevalent in cattle and thus can be readily isolated from bovine feces, bovine manure and feedlot catch-basin water samples (217, 257-259). The number of *E. hirae* isolates recovered from urban wastewater was low (3%) and most of the *Enterococcus* spp. from this source were identified as either *E. faecalis* or *E. faecium*. Similarly, only *E. faecalis* or *E. faecium* were identified among human clinical

isolates, confirming that *E. hirae* is generally not associated with human infections. However, *E. hirae* have occasionally been isolated from human cases of septicemia (260), endocarditis (261, 262), urinary tract infections (221, 226), spondylodiscitis (263) and acute pancreatitis (264). The rarity of this species among clinical enterococci isolates suggests that this species may not be as virulent as *E. faecalis* and *E. faecium*. This finding is also evident from our comparative genomic analysis where virulence genes were more frequently identified in *E. faecalis* and *E. faecium* but not in *E. hirae*.

The phenotypic resistance profiles of 1319 *E. hirae* isolates showed that antimicrobial use and resistance phenotype were linked within a particular environment. For example, macrolides and tetracyclines are commonly used in the beef cattle production systems for disease treatment and prevention including prophylaxis / metaphylaxis (256, 265, 266). Isolates recovered from bovine feces, feedlot catch basin, stockpiled bovine manure and natural surface water in the vicinity of the feedlots showed high occurrence of resistance to these antibiotics. Although only nine *E. hirae* isolates were recovered from urban wastewater, their resistance profiles indicated linkage with drugs commonly used to treat human infections including nitrofurantoin, tigecycline and linezolid.

Genotypic resistance profiles of *E. hirae* corroborated to their phenotypic profiles, where tetracycline and macrolide resistance genes were predominantly present in isolates recovered from the beef cattle production system and natural water sources. This is consistent with previous studies where tetracycline and macrolide resistance genotypes were prevalent in beef production systems (267-269). Tetracycline resistance was associated with the presence of *tetL, M,* and *O*. Two mosaic tetracycline genes *tetS/M* and *tetO/32/O* were also identified. *tetL* confers resistance via an efflux mechanism, while *tetM*,

tetS/M, O and *O/32/O* encode for ribosomal protection proteins (270-274). These genes are mostly found on transposable elements that are often linked with chloramphenicol and macrolide resistance determinants (275). Macrolide resistance was associated with the presence of *erm(B)*, which confers resistance against macrolide-lincosamide-streptogramin antibiotics (276). . Others have also found *erm(B)* in *E. hirae* (277, 278), as well as in *E. faecalis* and *E. faecium* isolated from chicken (279, 280), turkey (281, 282), pigs (283), fermented food (284) and clinical isolates (285-287). The macrolide resistance gene and tetracycline resistance genes in *E. hirae* appear to be identical to those in *E. faecalis* and *E. faecium* (217, 288). Considering that all of these species carry similar AMR determinants, the possibility of horizontal gene transfer across species seems probable (289). Studies have identified the presence of pheromone responsive plasmids in *E. faecium* and *E. faecium* that either encode vancomycin resistance or facilitate the transfer of plasmids carrying vancomycin ARGs into recipient cells (290, 291). These plasmids can also transfer between *Enterococcus* species, as the pMG1 plasmid has been shown to transfer between *E. faecium* and *E. faecalis*, and from *E. faecium* to *E. hirae* (92).

Aminoglycoside gene *acc(6')-lid* is known to be intrinsic in *E. hirae* (292); it was detected in all but two genomes, likely as a result of gene coverage and assemblage issues. It is not surprising that vancomycin resistance genes were not identified in *E. hirae*, as this drug is not approved for veterinary use in North American cattle. Our result is consistent with a previous study where vancomycin resistance genes were not identified in *Enterococcus* spp. isolated from bovine feces (288). The virulence genes identified in *E. hirae* were mostly associated with biofilm formation and polysaccharide biosynthesis, as described by others (293). The 10 virulence genes that were identified in *E. hirae* were similar to those

in *E. faecalis* and *E. faecium*, but many more (i.e., 49) virulence genes were found in *E. faecium* and *E. faecalis* (217).

Pan-genome analysis is an important comparative analysis tool that allows linkages between genetic changes and specific phenotypes as it describes core- and accessory genomes as well as species-specific genes (294). The core genome constituted only 64% of the total genome in *E. hirae*. Both horizontal and vertical transfer of genes, including those that confer antimicrobial resistance, play a significant role in shaping the pan-genome of a bacterial species (295). The pan-genome of *E. hirae* was considered "open" as there was no sign of saturation and it would be expected to expand with the addition of new genomes as illustrated by the pan-genome curve (Figure 4B). The high presence of cloud genes reflects the heterogeneity of the pan-genome of *E. hirae.* The existence of *E. hirae* in diverse environments may increase the chance of gene acquisition, in contrast to other *Enterococcus* species that may live in more specific environments that require less genomic variation for survival (296).

Gram-positive bacteria have sophisticated cell wall structures that ensures bacterial structural integrity and cellular viability and is also a major component of the host defence system (297, 298). For this reason, numerous studies have been conducted to explore components of cell wall synthesis pathways as potential targets for drug therapy. Genes involved in cell wall synthesis were identified as discriminatory genomic signatures between *E. faecalis* and *E. hirae*. Identification of these different signatures [peptidoglycan synthesis genes (*Mur E, F,* and *Y*), penicillin-binding protein (*pbpE* and *pbpX*), teichoic acid synthesis genes (*tag H, G,* and *U*) and enterococcal polysaccharide antigen (*eps E, D,*

M and *N*)] may identify targets that offer more specific drug development against *E. faecalis* and *E. faecium* (299).

Members of gut microflora compete with each other for nutrient availability. Therefore, the ability of one bacterial species to utilize multiple nutrients for energy generation provides an advantage over other species. We found genes involved in the synthesis cobalamin (vitamin B12), pyridoxine (vitamin B6), riboflavin (vitamin B2), biotin (vitamin B7) and folic acid exclusive to the *E. hirae* species compared with other enterococci analysed in this study. With vitamins being undeniably important for both bacterial and mammalian host, gut bacteria associated with their production directly contribute to the development and welfare of the host and thus may have a privilege of existence in cattle microbiome. In *E. hirae* genomes, multiple phosphotransferase systems (PTS) for fructose, galactitol, mannose, sorbose, glucose, N-acetyl glucosamine and cellobiose were identified. The presence of these PTS allows colonization of bacterial population in the gut (300). In addition, compared to *E. faecalis* and *E. faecium*, *E. hirae* harboured genes that were predicted to contribute to the synthesis of bacterial cellulose. These findings indicate that cellular metabolism genes identified in *E. hirae* may contribute to fitness within the cattle gut, accounting for its high prevalence in beef cattle.

Analysis of the annotated genomes indicated that *E. hirae* possessed genes coding for the production of antimicrobial agents like bacilysin, subtilosin, and narbonolide. Bacilysin is a dipeptide antimicrobial with antifungal and antibacterial activity (301). Subtilosin belongs to the lantibiotics class of bacteriocins and has anti-biofilm activity (302, 303). These bacteria are also capable of producing gamma-aminobutyric acid (GABA), an inhibitory neurotransmitter. GABA may increase feed intake in cattle, and reduce anxiety

and pain (304, 305). These findings suggest that *E. hirae* may have probiotic properties that could benefit the gastrointestinal environment of cattle (306). Previously, *E. hirae* has been employed as a probiotic bacteria in freshwater fish (307). Recent studies have also demonstrated that *E. hirae* may have confer probiotic properties within the intestinal tract of cattle (308, 309).

One of the goals of this study was to identify the niche-specific genes in *E. faecalis* and *E. faecium* that may contribute to virulence and infection. Several virulence factors that contribute to the pathogenesis of *E. faecalis* and *E. faecium* have been reported (310, 311). Several virulence genes were unique to *E. faecium* and/or *E. faecalis* and were not found in *E. hirae*. Members of the SlyA/ MarA family of proteins are associated with virulence gene regulation, promote biofilm formation and act as cell adhesions. The presence of genomic signatures corresponding to these genes may in part account for the higher prevalence of *E. faecalis* and *E. faecium* infections than *E. hirae* infections in humans (312-314). Lipoproteins facilitate intake of nutrients and are often associated with ABC transporters that are linked to pathogenesis. This supports our findings as genes encoding lipoproteins mapped with ABC transport systems for manganese, arabinose and methionine (315, 316). Furthermore, in *E. faecalis* and *E. faecium*, we also identified an arginine-ornithine antiporter which could contribute to cell fitness by facilitating arginine uptake. A study conducted to investigate the role of arginine-ornithine antiporter in *Streptococcus suis* reported that intercellular survival of this pathogen within epithelial cells was compromised in the absence of the antiporter (317).

The potential of *E. hirae* as an opportunistic pathogen cannot be ignored, as it is occasionally recovered from both humans and animals clinical samples (318-322). Despite

a higher prevalence of virulence genes in *E. faecium* and *E. faecalis*, some virulence genes were also identified in *E. hirae* like genes associated with the ESX (or Type VII) secretion system, bicyclomycin resistance, capsule biogenesis, quorum sensing system and an ABC transporter for iron import (323-327). In addition, a lipoprotein gene associate with the iron transport system have been identified and is thought to play a role in *E. hirae* establishing opportunistic infection (316).

In conclusion, *E. hirae* has a tremendous ability for survival and adaptation. It has acquired resistance to antimicrobials used in beef production systems. In addition, cellular metabolism genes involved in vitamin biosynthesis, multiple ABC and PTS transport systems, chitin degradation and cellulose synthesis provide selective advantage and facilitate intestinal colonization of the cattle gut. As *E. hirae* appears to be uniquely adapted to cattle hosts, this likely limits the extent to which it transfers genes to bacteria that are important in human health. Regardless, the absence of resistance to critical antimicrobials in *E. hirae* gives credibility to limiting use of these drugs in feedlots, and suggests that prudent management of antimicrobials in feedlot settings is an important practice.

Chapter 3: Comparative Genomic Analysis of *Enterococci* across Sectors of the One Health Continuum.

Microorganisms vol. 11, 3 727. 11 Mar. 2023, doi:10.3390/microorganisms11030727.

Author's Contributions

RZ and TAM, designed the study, DP-L provided data for swine isolates, MAR and MD provided data for poultry isolates, AS and ET provided data for a subset of environmental isolates and GVD provided and managed the bioinformatics tools and computing environment. SZ and RZ analyzed sequence data; SZ generated figures and analyzed overall data/results and wrote first draft of the manuscript; and TAM, AZ and RZ provided funding and supervision. All authors participated in editing and reviewing the manuscript and approved the final manuscript.

3.1 Abstract

Enterococci are Gram-positive bacteria often isolated from a variety of environments, including soil, water, plants and the intestinal tract of humans and animals. Although considered commensals in humans, *Enterococcus* spp. are important opportunistic pathogens. Due to their presence and persistence in diverse environments, *Enterococcus* spp. are ideal for studying antimicrobial resistance (AMR) from One Health perspective. We undertook a comparative genomic analysis of the resistome, virulome and mobilome of 246 *E. faecium* and 376 *E. faecalis* recovered from swine, beef cattle, poultry, dairy cattle, human clinical samples, municipal wastewater and environmental sources. Comparative genomics of *E. faecium* and *E. faecalis* identified 31 and 34 different antimicrobial resistance genes (ARGs), with 62% and 68% of the isolates having plasmid-associated ARGs, respectively. Across the One-health continuum, tetracycline (*tetL* and

tetM) and macrolide resistance (*ermB*) were commonly identified in *E. faecium* and *E. faecalis*. These ARGs were frequently associated with mobile genetic elements along with other ARGs conferring resistance against aminoglycosides [*ant(6)-la*, *aph(3')-IIIa*], lincosamides [*lnuG*, *lsaE*], and streptogramins (*sat4*). The core *E. faecium* genome formed two main clades, clade 'A' and 'B', with clade A isolates primarily originating from humans and municipal wastewater and carrying more virulence genes and ARGs related to category I antimicrobials. Overall, despite differences in antimicrobial usage across the continuum, tetracycline and macrolide resistance genes persisted in all sectors.

3.2 Introduction

Antimicrobial resistance (AMR) is defined as the ability of the bacterial cell to avoid cell damage by antimicrobials (328). Some bacteria are naturally resistant to certain antimicrobials through intrinsic or inherent traits. Antimicrobial resistance genes (ARGs) conferring intrinsic resistance are mostly passed through clonal inheritance and are rarely transferred within or among bacterial populations. However, some ARGs can be acquired and associated with mobile genetic elements (MGEs) including plasmids, transposons, and integrative and conjugative elements. These ARGs can be transferred to other bacteria through horizontal gene transfer (329) and thus contribute to the spread of AMR in different ecosystems (330). Exposure of bacteria to antimicrobials can facilitate ARG acquisition and the proliferation of resistant populations within ecosystems (28). In animal production, sub-therapeutic administration of antimicrobials through feed and water to treat or prevent infectious diseases is one example of a practice that can increase AMR. Indeed, the imposed selective pressure can exacerbate AMR in gut microbiomes as a number of bacterial members that carry ARGs on MGEs (77) may facilitate their

dissemination, including transfer to pathogenic bacteria. Therefore, multiple organizations, including the Canadian Integrated Program for Antimicrobial Resistance Surveillance (CIPARS), European Antimicrobial Susceptibility Surveillance in Animals (EASSA), Japanese veterinary antimicrobial resistance monitoring systems (JVARM), and the National Antimicrobial Resistance Monitoring System for Enteric Bacteria (NARMS) in the United States are monitoring antimicrobial resistance in food animals and assessing their role in the dissemination of AMR to bacteria associated with humans.

Enterococci are commensal bacteria within the gastrointestinal tract of humans and animals (101). They can also be recovered from broader natural environments, including soil, water and plants. Some enterococcal species, particularly *Enterococcus faecalis* and *Enterococcus faecium*, are considered human pathogens as they are frequently associated with bacteremia, septicemia, meningitis, endocarditis, and urinary tract and wound infections (331). The presence of *Enterococcus* spp. in different ecosystems makes them an ideal species to study AMR from a One-health perspective. We investigated the prevalence and nature of *Enterococcus* species recovered from swine feces and undertook a comparative analysis of *E. faecium* and *E. faecalis* genomes sourced across various sectors of the One Health continuum. More specifically, we evaluated (i) profiles of ARGs, MGEs and virulence factors of these genomes (ii) association of MGEs with ARGs, and (iii) phylogenetic relatedness of isolates collected across sectors.

3.3 Methodology

3.3.1 Enterococcus recovery from swine feces and whole genome sequencing

In 2017 and 2018, fecal samples were collected from sows, weaning and finishing pigs raised on commercial antimicrobial-free farms, as well as conventional farms using

penicillin prophylaxis in Quebec, Canada. Isolates were collected at the same time as *Enterobacterales* isolates were collected in a previous study (332). Presumptive *Enterococcus* isolates were recovered from collected samples on Bile Esculin Azide (BEA) agar with and without erythromycin (8µg/mL) as described previously (217) and a total of 41 isolates were confirmed to be *Enterococcus* species following PCR with Ent-ES-211-233-F and Ent-EL-74-95-R primers and Sanger sequencing of the PCR product (217). Confirmed isolates were subjected to short-read Illumina sequencing. Genomic DNA was extracted using a Maxwell 16 Cell SEV DNA purification kit (Promega, Madison, WI) as per manufacturer's instructions, followed by DNA quantification using a Quant-it High-Sensitivity DNA assay kit (Life Technologies Inc., Burlington, Canada). One nanogram of gDNA was used for genomic library construction using an Illumina NexteraXT DNA sample preparation kit and the Nextera XT Index kit (Illumina Inc., Vancouver, Canada) according to manufacturer's guidelines. All libraries were sequenced on an Illumina Miseq platform generating 2x300 base-paired end reads with 600-cycle MiSeq reagent kit v3 (Illumina©).

3.3.2. Collection of *Enterococcus faecium* and *Enterococcus faecalis* genomes

A total of 622 *E. faecium* and *E. faecalis* genomes were included for comparative genomic analysis. These genomes originated from three sources: (i) swine isolates from this study (n = 18), (ii) a collection of genomes recovered from environmental and livestock isolates from Ontario (n = 66) and (iii) previously published data from poultry (n = 32) (333) and One-health continuum (n = 506) (217) studies. The number and the origin of *E. faecium* and *E. faecalis* genomes included in the analysis are summarized in Table 3.1. *E. faecium*

and *E. faecalis* genomes were categorized into four groups/sectors based on their origin: (i) clinical, (ii) municipal wastewater, (iii) livestock and (iv) environment.

Table 3.1 Collection of *Enterococcus faecium* and *Enterococcus faecalis* genomes included in the comparative genomic analysis and antimicrobials used in livestock.

Sources of genome		Number of genome		Antimicrobial usage	Location (Year of sample collection)	Reference
		E. faecium **(n=246)**	***E. faecalis*** **(n=376)**			
Municipal waste water (MW)		56	110	-	Alberta (March 2014-April 2016)	(217)
Clinical isolates (CL)		36	149	-		
Livestock (LV)	**Bovine cattle**	57	33	Conventional (tetracycline, macrolides), natural (antibiotic-free)		
	Dairy cattle	-	22	NA	Ontario (2004)	This study
	Swine	-	06	NA		
		12	06	Conventional (penicillin), antibiotic free (organic, certified-humane, AGRO-COM)	Quebec (2017-2018)	
	Poultry	23	09	Bambermycin, bacitracin, salinomycin and β-lactams	British Colombia (2005-2008)	(333)
		-	05	NA	Ontario (2004)	This study
Environment (EV)	**Natural water sources**	46	19	-	Alberta (March 2014-April 2016)	(217)
	River water	16	07	-	Ontario (2004)	This study
	Domestic animals	-	03	NA		
	Wild animals	-	07	-		

NA – Not available

3.3.3 Genome assembly and data analysis

All enterococcal genomes included in this study were *de novo* assembled using the Shovill pipeline v.1.1.0 (https://github.com/tseemann/shovill). Illumina adapters were removed using Trimmomatic v.0.36.5 (334). All reads were then *de novo* assembled into contigs by SPAdes v.3.11.1 (335). Assembly was evaluated by QUAST v.v. 5.2.0 (336). The contigs were then annotated using Prokka v.1.13.1 (337).

The annotated genomes were screened for the presence of antimicrobial resistance and virulence genes using ABRicate v.1.0.1 (https://github.com/tseemann/ABRICATE) against the NCBI Bacterial Antimicrobial Resistance Reference Gene Database (NCBI BioProject ID: PRJNA313047) and the VirulenceFinder database (PMID: 34850947) (338), respectively. All contigs were screened for the presence of plasmids using the Mob-recon v.v. 3.0.0 (https://github.com/phac-nml/mob-suite) (339).

E. faecium (n=246) and *E. faecalis* (n=376) genomes were used for comparative genomics (Table 3.1). The core-genome phylogenomic trees were constructed using the SNVphyl pipeline v. 1.2.3. The phylogenetic tree was generated by aligning paired-end Illumina reads against the respective reference genomes of *E. faecalis* (strain ATCC 47077/OG1RF; CP002621.1) and *E. faecium* (strain DO; CP003583.1) using SMALT (v. 0.7.5; https://sourceforge.net/projects/smalt/). The generated read pileups were then subjected to quality filtering (minimum mean mapping quality score of 30), coverage cut-offs (15X minimum depth of coverage), and a single nucleotide variant (SNV) abundance ratio filter of 0.75 to obtain a multiple sequence alignment of SNV-containing sites. This SNV alignment (with no SNV density filtering) was used to create a maximum likelihood

phylogeny using PhyML v. 3.0. The generated Newick file was visualized using Interactive Tree Of Life (iTOL) v. 6 (231).

Additionally, for *E. faecium* genomes, a *groEL*-based tree was constructed to investigate whether the genomes could be assigned to previously described hospital (clade A) or community (clade B) clades (340). The extracted *groEL* gene sequence was aligned with *E. faecium* strain 75 V68 (Clade A) and *E. faecium* strain 81 (Clade B) using MAFFT v. 7.490. The analysis included the *E. hirae* R17 (accession CP015516.1) *groEL* gene as an outgroup. The maximum-likelihood tree was then created with IQtree v. 2.1.4.

Multilocus sequence typing (MLST) was also used to study the population structure and evolution of bacterial species. *E. faecium* and *E. faecalis* sequence types were assigned through the MLST scheme of each respective species using PubMLST tool (http://cge.cbs.dtu.dk/services/MLST/) (341).

3.4 Results

3.4.1 Enterococci recovered from swine feces

3.4.1.1 Species identification

Of the *Enterococcus* spp. recovered from fecal samples, 14 isolates were from sows, 15 isolates were from weaners and 12 isolates were from finishers. Six different enterococcal species were identified [*E. hirae* (n=15), *E. faecium* (n=12), *E. faecalis* (n=6), *E. saccharolyticus* (n=3), *E. villorum* (n=3), and *E. asini* (n=2)] (Figure 3.1).

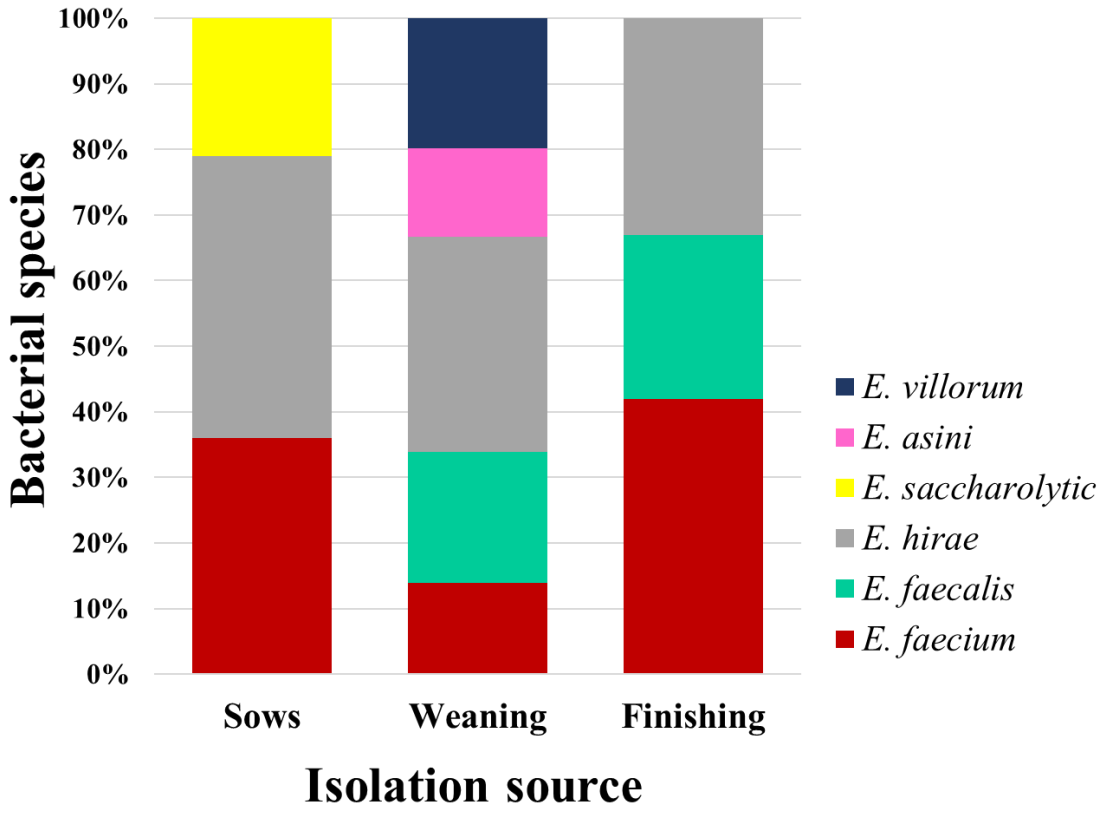

Figure 3.1 Prevalence of *Enterococcus* species recovered from fecal samples collected from sow (n=14), weaning (n=15) and finishing (n=12) pigs

3.4.2 Genome characterization

Across all isolates, 27 different ARGs/determinants were identified (Figure 3.2). Overall, 39% of identified enterococcal species were multidrug-resistant (MDR, resistant to ≥3 antimicrobials). MDR isolates were confined to three species: *E. faecalis* (67 %), *E. hirae* (47%) and *E. faecium* (41%) (Table 3.2). The most common ARGs in *E. faecium*, *E. faecalis* and *E. hirae* belonged to aminoglycoside (*aph(3')-IIIa, ant(6)-Ia*), tetracycline *(tetL , tetM),* macrolide (*ermB*), and streptothricin (*sat4*) drug classes.

Nine out of 27 ARGs conferred intrinsic/inherent resistance, including *msrC* (100%), *eat(A)* (100%) and *aac(6')-li* (41.6%) in *E. faecium*; *lsa(A)* (100%) and *dfrE* (100%) in *E. faecalis*, *aac(6')-lid* (66.6%) in *E. hirae*, *dfrF* (100%) and *vanC*-operon (100%) in *E. saccharolyticus*, and *aac(6')-Entco* (100%) in *E. villorum* and *E. asini.* The three genes, *aacA-ENT1, dfrG and aacA-ENT2,* were only identified in *E. faecium* (16.6%), *E. faecalis* (33.3%) and *E. hirae* (33.3%), respectively.

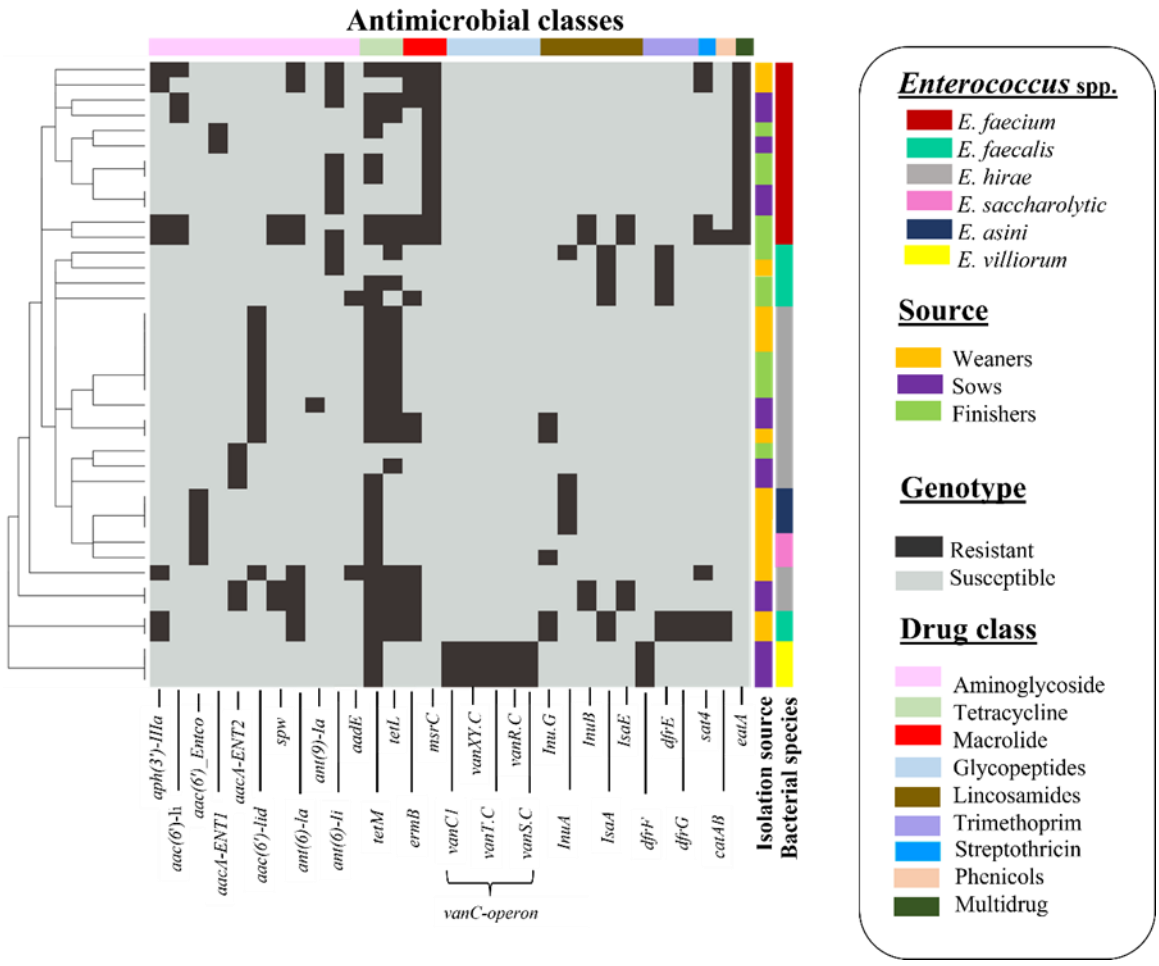

Figure 3.2 Heatmap representing antimicrobial resistance gene profiles of *Enterococcus* isolates recovered from fecal samples collected from sow (n=14), weaning (n=15) and finishing (n=12) pigs.

Table 3.2 Antimicrobial resistance genes profiles, plasmids harboring AMR genes and virulence genes identified in enterococcal species recovered from swine feces.

Enterococcal species	*&Antimicrobial resistance genes profile (Number of genomes)	Plasmids (Accession number) (total)	Antimicrobial resistance genes found on plasmid	Virulence Genes
E. faecalis	*aph(3')-IIIa, ant(6)-la, tetL, tetM, ermB, lnu(G), dfrG, sat4, catA8* (n=2)	pBEE99 (NC_013533) (n=2)	All ARGs	Adhesive matrix molecules: *ace, fss1 and fss2* Biofilm formation: *bopD* Capsule formation: *cpsA-E* & *cpsG-K* Cytolysis: *cylA, cylB, cylI, cylL, cylM, cylR1*, cylR2 & cylS Endocarditis and biofilm-associated pili: *ebpA-C* & *srtC* Putative transporter protein: *efaA* Hyaluronidase: *EF0818* & EF3023) Gelatinase and serine protease: *fsrA-C, gelE,* & *sprE* Aggregation proteins: *prgB/asc10*
	tetL, tetM (n=1)	pSWS47 (NC_022618.1) (n=1)	All ARGs	
	aadE, tetM, ermB (n=1)	None	None	
	tetL, lnu(A) (n=1)	None	None	
E. feacium	*aph(3')-IIIa, spw, ant(6)-Ia, tetL, tetM, ermB, lnu(B), lsa(E), sat4, catA8* (n=1)	pM7M2 (NC_016009) (n=4)	*tetL, tetM*	Adhesive matrix molecules: *acm, scm* & *sgrA* Biofilm formation: *bopD, clpC, clpE* & *clpP* Bile salt hydrolysis: *bsh*
	aph(3')-IIIa, ant(6)-Ia, tetL, tetM, ermB, sat4 (n=1)			
	tetL, tetM, ermB (n=1)			

	tetL, tetM (n=1)			Capsule formation: *cap8F, cpsA, cpsB* & *hasC* Pili formation: *srtC*
	aph(3')-IIIa, spw, ant(6)-Ia, tetL, tetM, ermB, lnu(B), lsa(E), sat4 (n=1)	pLAG (KY264168.1) (n=1)	*ant(6)-Ia, tetM, tetL, lnu(B), lsa(E)*	
	aph(3')-IIIa, ant(6)-Ia, ermB, sat4 (n=1)	none	none	
	tetM (n=3)	none	none	
E. hirae	*aph(3')-IIIa, ant(6)-Ia, aadE, tetL, tetM, ermB, sat4* (n=1)	p3 (CP006623) (n=1)	*aph(3')-IIIa, ant(6)-Ia, ermB, sat4*	Biofilm formation: *bopD* & *clpP* Hydrolysis of bile salt: *bsh*
		pBC16 (U32369) (n=1)	*tetM*	
	spw, ant(6)-Ia, tetL, tetM, ermB, lnuB, lsaE (n=2)	pEf37BA (MG957432) (n=2)	All ARGs	
	tetL, tetM, ermB, lnuG (n=2)	pDO1 (CP003584) (n=2)	*tetL, tetM, ermB*	
	ant(9)-Ia, tetL, tetM (n=1)	pM7M2 (NC_016009) (n=1)	*tetL, tetM*	
	tetL, tetM (n=7)	pM7M2 (NC_016009) (n=3)	*tetL, tetM*	
		pCTN1046 (CP007650) (n=1)	*tetM*	
		pBC16 (U32369) (n=1)	*tetL*	

	tetM, lnuA (n=1)	(CP029969) (n=1)	*lnu(A)*	
E. asini	*tetM, lnuG* (n=1)	None	None	Adhesion associated gene: *fss3*
	tetM (n=1)	None	None	
E. villorum	*tetM, lsaA* (n=3)	None	None	None
E. saccharolyticus	*tetM* (n=3)	None	None	Adhesion associated gene: *fss3*

* Antimicrobial drug classes and resistance genes: Aminoglycoside (*ant(9)-Ia, aph(3')-IIIa, ant(6)-Ia, aadE, spw*); Tetracycline (*tetL, tetM*); Macrolide (*ermB*), Lincosamides ARG (*lnuA, lnuG, lsaA, lsaE*), chloramphenicol (catA8), Trimethoprim (*dfrG*)

& All ARGs except for those shown in column 4 were mapped on chromosome.

A total of 35 plasmids were identified in *Enterococcus* spp. [*E. faecalis* (n=10), *E. faecium* (n=12) and *E. hirae* (n=13)] (Table 3.2). Among these, eleven plasmids harbored ARGs [*E. faecalis* (n=2), *E. faecium* (n=2) and *E. hirae* (n=7)] (Table 3.2). A total of 34 and 13 virulence genes were identified *in E. faecalis* and *E. faecium*, respectively. Most virulence genes were associated with cytolysis, biofilm and capsule formation (Table 3.2).

The core-genome *E. faecium* phylogenetic tree formed two distinct clades, where all genomes except two recovered from sows and finishers, were found in one clade. *E. faecalis* also clustered into two clades, where one clade contained exclusively genomes from weaners. As for *E. hirae,* one clade contained all genomes except two isolated from finishers (Figure 3.3).

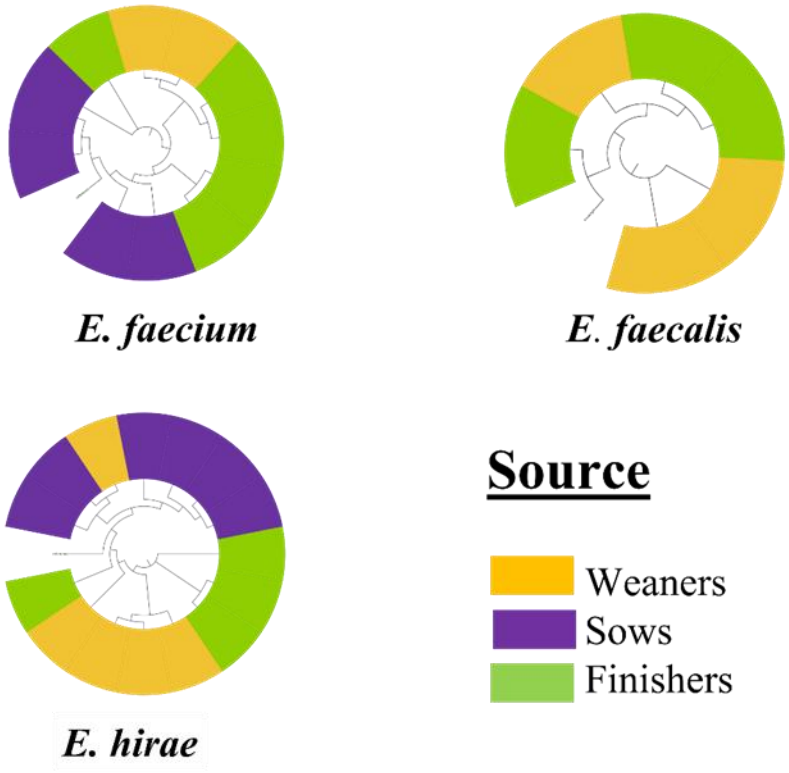

Figure 3.3 Core genome based phylogenetic tree of *E. faecium* (n=12), *E. faecalis* (n=6) and *E. hirae* (n=15) recovered across pig production stages. The trees were constructed using SNVPhyl tree and visualized using iTOL.

3.4.3 Comparative genomic analysis of E. faecalis and E. faecium across the One Health continuum

3.4.3.1 Livestock production

Comparative genomic analysis of *E. faecium* (n=91) and *E. faecalis* (n=81) collected from cattle, poultry and swine was performed to investigate similarities and differences in resistome, virulome and mobilome profiles as well as phylogenetic relatedness across the production sectors.

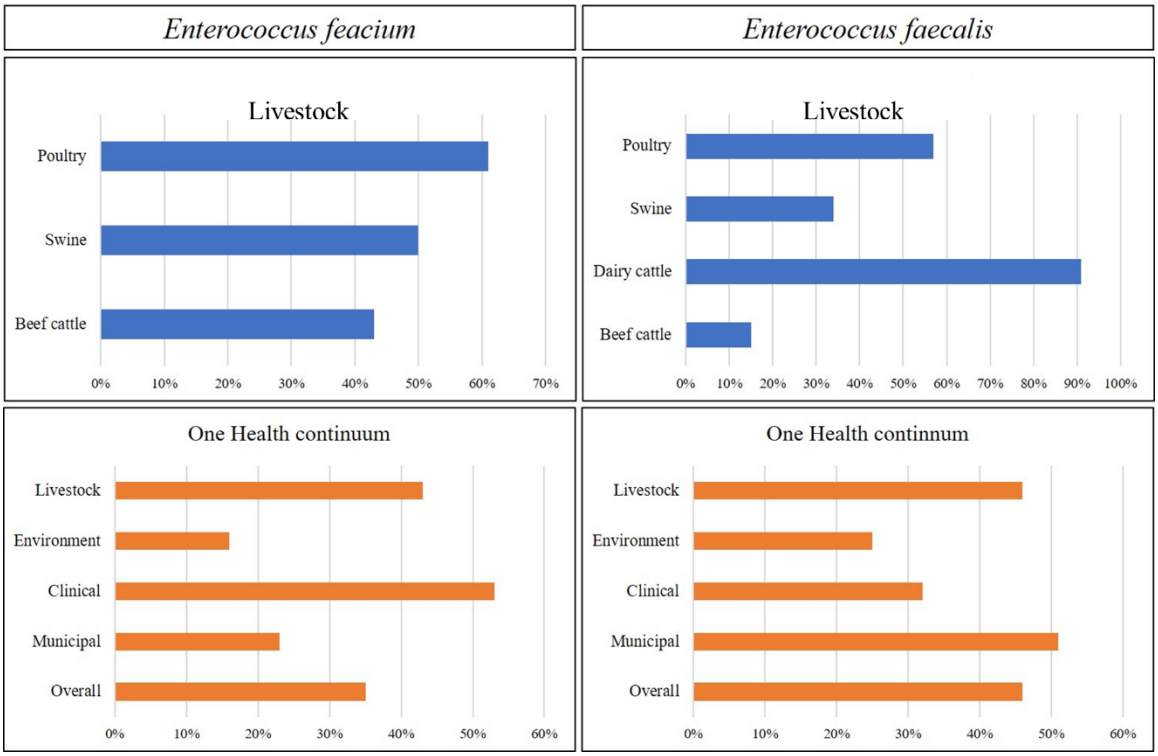

Figure 3.4 Multidrug resistant *Enterococcus faecium* and *Enterococcus faecalis* from overall continuum. Livestock sector consist of beef cattle , dairy cattle, poultry and swine. One-health continuum consists of livestock (beef cattle, dairy cattle, poultry and swine), environment (natural water source, river water, domestic and wild animal fecal samples), Municipal wastewater, human clinical isolates. Overall represents all isolates in included in this study. The percentage was calculated as total number of MDR isolates in particular sector divided by total number of isolates in that sectors and multiplying by 100.

Overall, 48% of *E. faecium* genomes from livestock were MDR (resistant to ≥3 antimicrobials) (Figure 3.4). Among livestock, *E. faecium* from poultry were the most frequently MDR (61%), followed by swine (50%) and beef cattle (43%) (Figure 3.4). Among *E. faecium* of bovine origin, two ARG profiles [(*ermB, tetL, tetM*) and (*ant(6)-Ia, spw, ermB, lnuB, lsaE, tetL, tetM*)] were most frequent (Appendix 2; Table S1). Isolates harboring *dfrE* were frequently identified in all sectors. Two ARG profiles [(*dfrE, tetL, tetM*) and (*dfrE, ermB, tetL, tetM*)] were present in both swine and poultry, while one profile (*dfrE, ermB* and *tetM*) was common to bovine and poultry isolates. Across livestock, chloramphenicol (*fexA* and *catA*) and oxazolidinone-resistant determinants (*optrA*) were exclusively found in *E. faecium* from cattle, whereas the *vanC*-operon was unique to poultry isolates. Aminoglycosides ARGs [*ant(6)-Ia, ant(9)-Ia, aph(3')-IIIa* and *spw*] were more prevalent in *E. faecium* isolated from poultry as compared to other sectors (Figure 3.5). In contrast, tetracycline ARGs (*tetL* and *tetM*) were found more frequently in *E. faecium* from cattle than in poultry and swine. Moreover, *E. faecium* isolates from cattle and poultry shared similar ARGs associated with macrolide-lincosamide-streptogramin (MLS) resistance (*ermA, ermB, lnuB, lnuG, lsaG* and *sat4*). In *E. faecium* from swine, only four ARGs associated with MLS resistance (*ermB, lsaG, mefA* and *sat4*) were identified. Across livestock, *ermB* (57%) was most prevalent in isolates from cattle. In contrast, the trimethoprim-resistant determinant *dfrE,* was found in all *E. faecium* genomes recovered from swine and 82.6% from poultry. Compared to other sectors, *drfE* and *dfrG* were infrequently associated with *E. faecium* isolated from cattle.

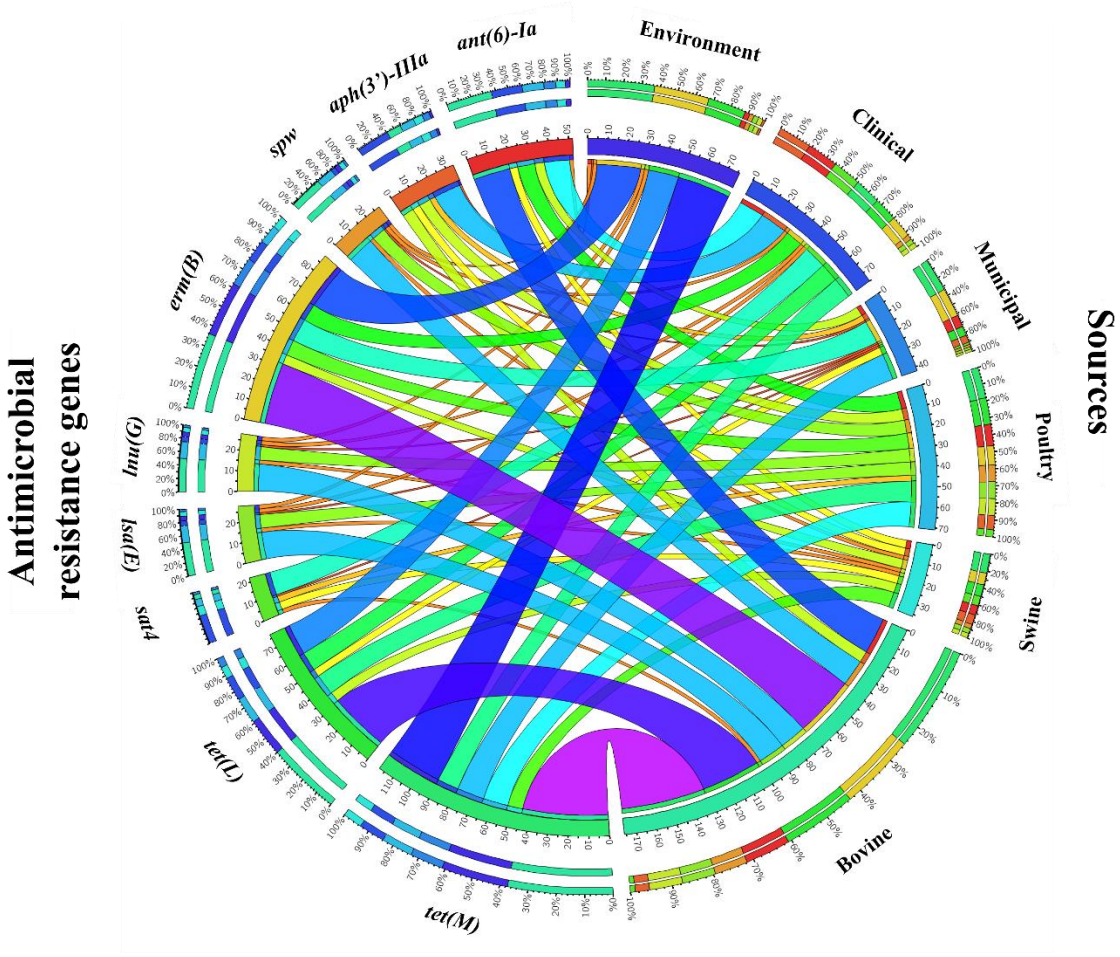

Figure 3.5 Comparative genomic analysis of 246 *E. faecium* genomes across the One-health continuum. A Circos plot depicts the relationship between commonly found ARGs and One-health sectors. The variables (ARGs and genome isolation source) are arranged around the circle and distinguished by different colors: The percentage of ARGs across various sectors is indicated by proportional bars (http://circos.ca/).

Mobilome analysis of *E. faecium* genomes showed that >60% of ARG-carrying plasmids were associated with isolates from cattle (Appendix 2; Table S2). Among these, pL8-A and pM7M2 were also found in poultry and swine isolates, respectively. MLST profiling identified 33 different genomic sequence types (STs) across enterococci genomes, with 13 STs exclusive to beef cattle. In swine, only 3 STs were identified (ST94, ST133, ST272). In *E. faecium* from poultry, 10 STs were identified, with ST154 being the most common. None of the STs was shared across all livestock species (Appendix 2; Table S3).

The virulome of *E. faecium* did not vary across livestock species. The majority of virulence genes, including those responsible for biofilm formation (*bopD*, *clpC*, *clpP*), bile-salt hydrolysis (*bsh*), capsule formation (*cap8F*, *cpsA*, *cpsB* and *hasC*), MSCRAMM-like proteins (*sgrA*) and pili formation (*srtC*) were found in >70% of the genomes of *E. faecium* from livestock. Two genes, *ebpA* and *lap* (encoding biofilm-associated pili), and a *Listeria* adhesion protein were identified in one poultry isolate (Appendix 2; Table S4).

Overall, 46% of *E. faecalis* were MDR with most MDR associated with isolates from dairy cattle (91%) followed by poultry (57%), swine (34%) and beef cattle (15%) (Figure 3.4). One ARG profile (ermB, tetM, tetL) was found across all livestock species (Appendix 2; Table S5). The ARG profile *ant(6)-Ia, aph(3')-IIIa, ermB, tetL, tetM* was present in 50% of poultry and 100% of *E. faecalis* genomes from dairy cattle (Figure 3.6). Similar to E. faecium, the oxazolidinone resistance gene (*optrA*) was occasionally (7% of genomes) present in E. faecalis isolated from cattle. The trimethoprim ARG (*drfE*) was mapped to 17% and 3% of *E. faecalis* isolates from swine and cattle, respectively, but was absent in poultry isolates. Chloramphenicol resistance profiles differed across sectors, as *catA8* was found in isolates from swine, whereas *catA7* was found in isolates from dairy cattle and *catA7* and *fexA* in isolates from beef cattle. Similarly, the profile of aminoglycoside ARGs also varied across livestock species (Figure 3.6). Aminoglycoside ARGs were most prevalent in isolates from dairy cattle, followed by poultry, swine and beef cattle. Two ARGs, *ant(6)-la and aph(3')-IIIa,* were prevalent across livestock species, whereas *aph(2'')-Ih* and *ant(9)* were unique to isolates from dairy and beef cattle, respectively. The ARG str, was found only in isolates obtained from beef cattle and poultry. Similarly, aadE was found only in isolates from swine and beef cattle. Tetracycline resistance determinants (*tetL*, *tetM*) were found in isolates across livestock sectors (Figure 4.4A; Appendix 2; Table S5).

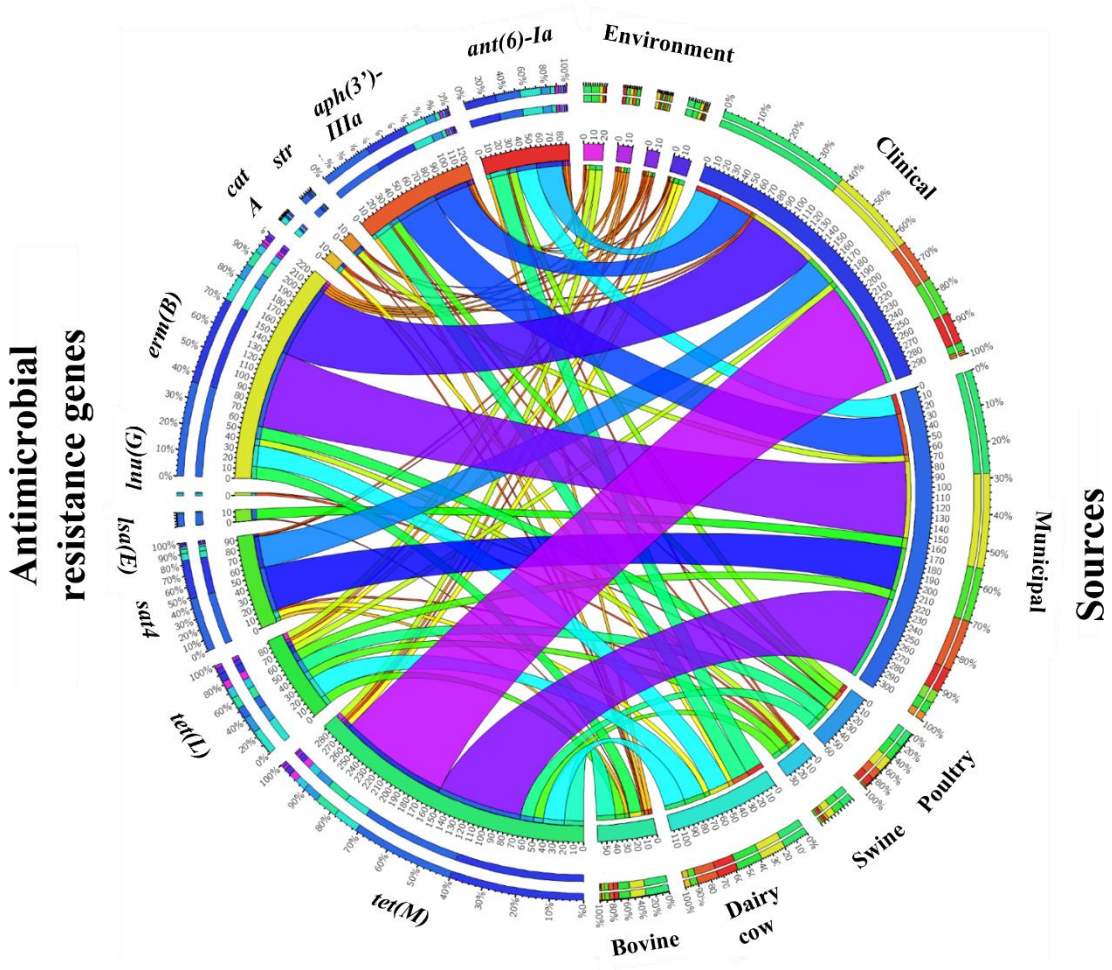

Figure 3.6 Comparative genomic analysis of 376 genomes *E. faecalis* genomes across the One-health continuum. A circos plot depicts the relationship between commonly found ARGs and One-health sectors. The variables (ARGs and genome isolation source) are arranged around the circle and distinguished by different colors: The percentage of ARGs across various sectors is indicated by proportional bars (http://circos.ca/)

Like *E. faecium*, plasmid profiling of *E. faecalis* found that 70% of isolates possessed plasmids that carried ARGs (Appendix 2; Table S6). Four ARGs carrying plasmids (DO plasmid, pCTN1046, p6742_2, pEf37BA and pBC16) were found in both *E. faecium* and *E. faecalis*. Across livestock species, 29 STs were identified, with ST 59 shared between swine, bovine and dairy cattle isolates (Appendix 2; Table S7). Virolome profiles of *E. faecium* genomes were similar across livestock species (Appendix 2; Table S8). Twenty-seven of 39 virulence genes were mapped to several isolates collected across the livestock sectors (40–100% of genomes). Genes encoding cytolysin (*cylA, cylB, cylI, cylL, cylM, cylR1, cylR2* and *cylS*) and the aggregation substance (*asa1*) were found in only one isolate from swine.

3.4.3.2 One Health continuum

Across the continuum, 35% of *E. faecium* were MDR, with most MDR found in clinical (CL) isolates (53%), followed by livestock (LI) (48%), municipal wastewater (MW) (23%) and environmental (EV) isolates (16%) (Figure 3.4). The ARG profile: *dfrE, ermB, tetM* was most common among MDR *E. faecium* from LI, EV and MW (Table S1). Aminoglycoside resistance genes were most prevalent in clinical genomes, followed by LI, MW and EV (Figure 3.5). Three aminoglycoside resistance genes, *ant(6)-Ia*, *aph(3')-IIIa* and *spw*, were found across the One Health continuum, with *ant(6)-Ia* and *aph(3')-IIIa* being frequently mapped to plasmids (73% and 61%, respectively). These genes were found together in 91% of genomes. The bifunctional gene, *aac(6')-Ie/aph(2'')-Ia*, was found only in CL (5/36, 14%) and MW (3/56, 5.3%) isolates. Genomes harboring *aac(6')-Ie/aph(2'')-Ia* were associated with five different plasmids (Appendix 2; Table S6). This gene was exclusively associated with an IS256 insertion element, except for one plasmid

associated with IS6 and IS1216 in combination with *ermB* and *dfrG*. Chloramphenicol resistance was found in LI and MW isolates but not among those from other sources. The ARG *fexA* was associated with *Tn*554 on plasmid pFSIS1608820, and *catA* was mapped to two plasmids in MW isolates (Appendix 2; Table S2). ARGs conferring resistance to trimethoprim were more prevalent in CL, followed by MW, LI and EV. Compared to CL isolates, where *dfrF* and *dfrG* were more prevalent, *dfrE* was found in EV, LI and MW isolates. In all but one *dfrG*-positive genome, *fosX* was found in an antisense direction to *dfrG* at an intergenic distance of ~3.2 kb. Macrolide-lincosamides-streptogramin-resistant genotypes were prevalent in LI, followed by CL, EV and MW.

Four ARGs conferring macrolide resistance (*ermA*, *ermB*, *ermT* and *mefA*) were identified across the continuum. The ARG *ermB*, was associated with plasmids 73% of the time. Moreover, in isolates from CL and LI, *ermB* along with the aminoglycoside ARGs *sat4*, *aph(3')-IIIa* and *ant(6)-la* were associated with *Tn*3 transposons. Similarly, *ermA* was also identified on plasmid pL8-A along with *ermB* and *ant(9)-la*. The ARG *ermA* was also found on plasmid pFSIS1608820 with *ant(9)-Ia, cfr, optrA, ermA, and fexA*. In contrast, *ermT* mapped only to plasmid p121BS. The lincosamide-resistant genes *lnuB* and *lsaE* were found together on 87% of plasmids. Glycopeptide resistance was found in clinical and poultry genomes, where *vanA* was found in pV24-3 and pF856 plasmids (Appendix 2; Table S2).

The core-genome-based phylogenomic tree of *E. faecium* formed two clades that were completely superimposed with the A and B clades identified by the *groEL* gene maximum-likelihood tree (Figure 3.7). *E. faecium* genomes did not group based on sample source, except for the clinical isolates in clade A. Furthermore, clade A harbored more virulence

genes and ARGs than clade B. Multilocus sequence typing of *E. faecium* genomes identified 72 different STs (Appendix 2; Table S3), with ST117 and ST17 being exclusive to human clinical isolates. Across the continuum, 37 virulence genes were identified, of which 15 were found in genomes from all sectors (Appendix 2; Table S4).

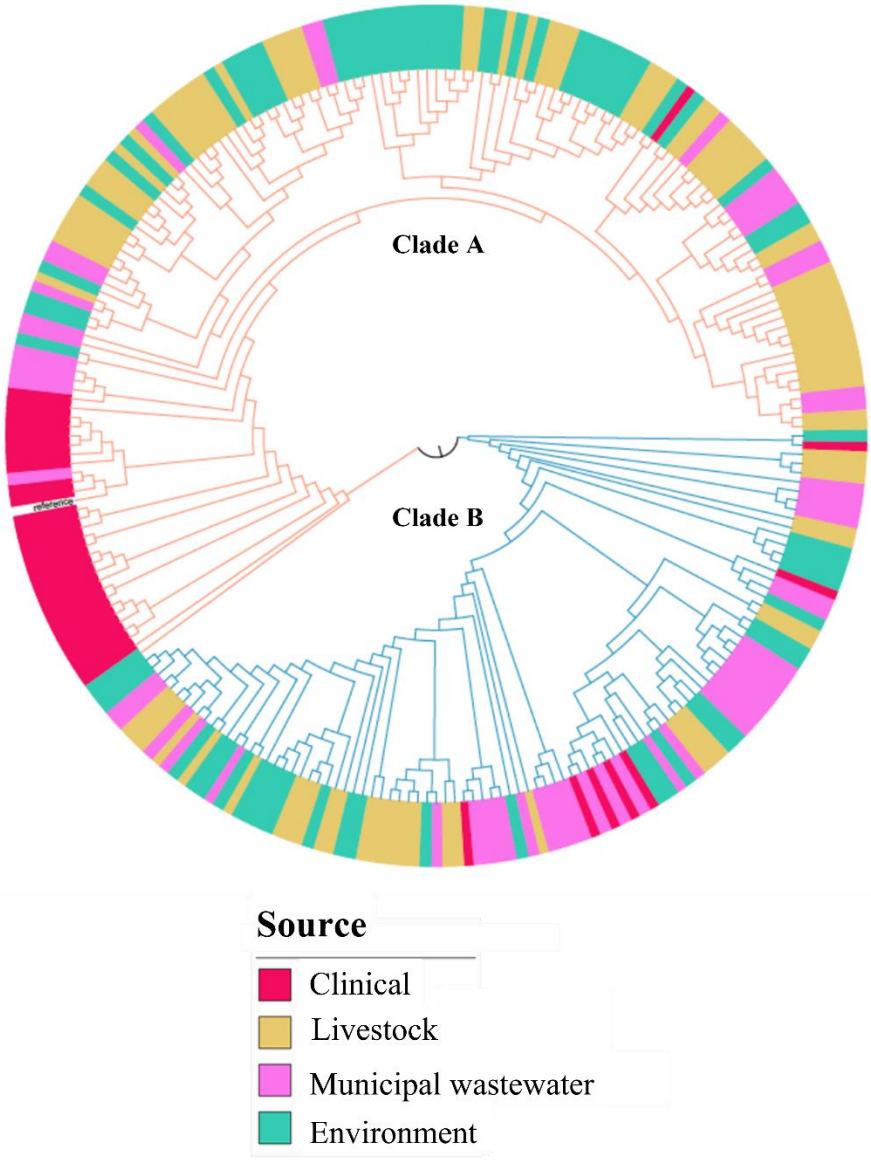

Figure 3.7 Maximum likelihood core-genome phylogenetic tree of 246 *E. faecium* genomes across the One-health continuum. The *Enterococcus faecium* DO genome (CP003583.1) was used as a reference genome. The *gro-EL* gene-based *E. faecium* tree was overlaid on the core-genome *E. faecium* tree. Genomes were characterized based on their source of isolation into four groups: livestock, clinical, municipal wastewater and environmental.

Overall, 40% of *E. faecalis* were MDR, with MDR isolates being most frequent in MW (51%) followed by LI (46%), CL (32%) and EV (25%) (Figure 3.4). Across all sectors, *ant(6)-Ia, aph(3')-IIIa, ermB, tetL, tetM* were frequently identified in MDR *E. faecalis* genomes (Appendix 2; Table S5). A total of 51 plasmids carrying one or more ARGs were identified (Appendix 2; Table S6). Among these plasmids, two were conjugative plasmids (related to AY855841 and CP028721), and two were identified as mobilizable plasmids (related to CP028286 and CP028836). Aminoglycosides ARGs were more prevalent in MW, followed by LI, CL and EV (Figure 3.6). Across all sectors, eight aminoglycosides ARGs were identified, with five (*ant(6)-Ia, aph(2'')-Ih, aph(3')-IIIa,* and *str*) found in all sectors. Similar to *E. faecium*, *ant(6)-Ia* and *aph(3')-IIIa* were frequently found together (61 genomes) and mapped to plasmids (71% and 75% of isolates, respectively). Chloramphenicol resistance genes were more prevalent in LI, followed by EV, CL and MW. Five ARGs (*catA7, catA8, catP, cat-TC* and *fexA*) were identified, with *catA7, catA8* and *fexA* present in all sectors. These three genes were always associated with plasmids (Appendix 2; Table S6). Trimethoprim ARGs (*dfrF/G*) were identified more in CL as compared to other sectors, with *dfrF* found in >60% of CL genomes (19% on a plasmid). Across all sectors, MLS resistance was more prevalent in MW, followed by LI, CL and EV. Three ARGs responsible for macrolide resistance (*erm A, ermB* and *msr*) were identified, with *ermB* present in 60% of all genomes and frequently associated with plasmids (75%). One *ermB*-carrying plasmid, CP024844 was found exclusively in CL and MW genomes (40% *ermB*-positive isolates). Lincosamide ARGs were not found in EV genomes, whereas in CL genomes, only *lnuB* was identified. Tetracycline resistance was found more in LI genomes, followed by EV, CL and MW. Five different tetracycline ARGs

were identified (*tetM, tetL, tetO, tetS,* and *tetW*), with *tetM* mapping to 76.5% of the genomes. Compared to *tetM* (18%), *tetL* (85%) was more frequently found on plasmids. Moreover, in 85% of *tetM*-positive plasmids, *tetL* was found together in close proximity with *tetM*. One *tetM*- and *tetL*-carrying plasmid, pS7316, was also prevalent in isolates from LI, CL and EV. Oxazolidinones resistance ARGs were found only in EV and LI, which were more prevalent in EV than LI. In EV, two ARGs (*optrA* and *cfrC*) were identified, whereas in LI, only *optrA* was found.

Across the continuum, the core-genome-based *E. faecalis* phylogenomic tree formed two main clades, where one clade contained the majority of CW and MW genomes (Figure 3.8). MLST profiling of *E. faecalis* identified 75 different STs (Appendix 2; Table S7), where 48 STs were source-specific (CL=17, LI=14, EV=8, MW=9). We identified 40 virulence genes across all *E. faecalis* genomes, with 28 shared across all sectors (Appendix 2; Table S8).

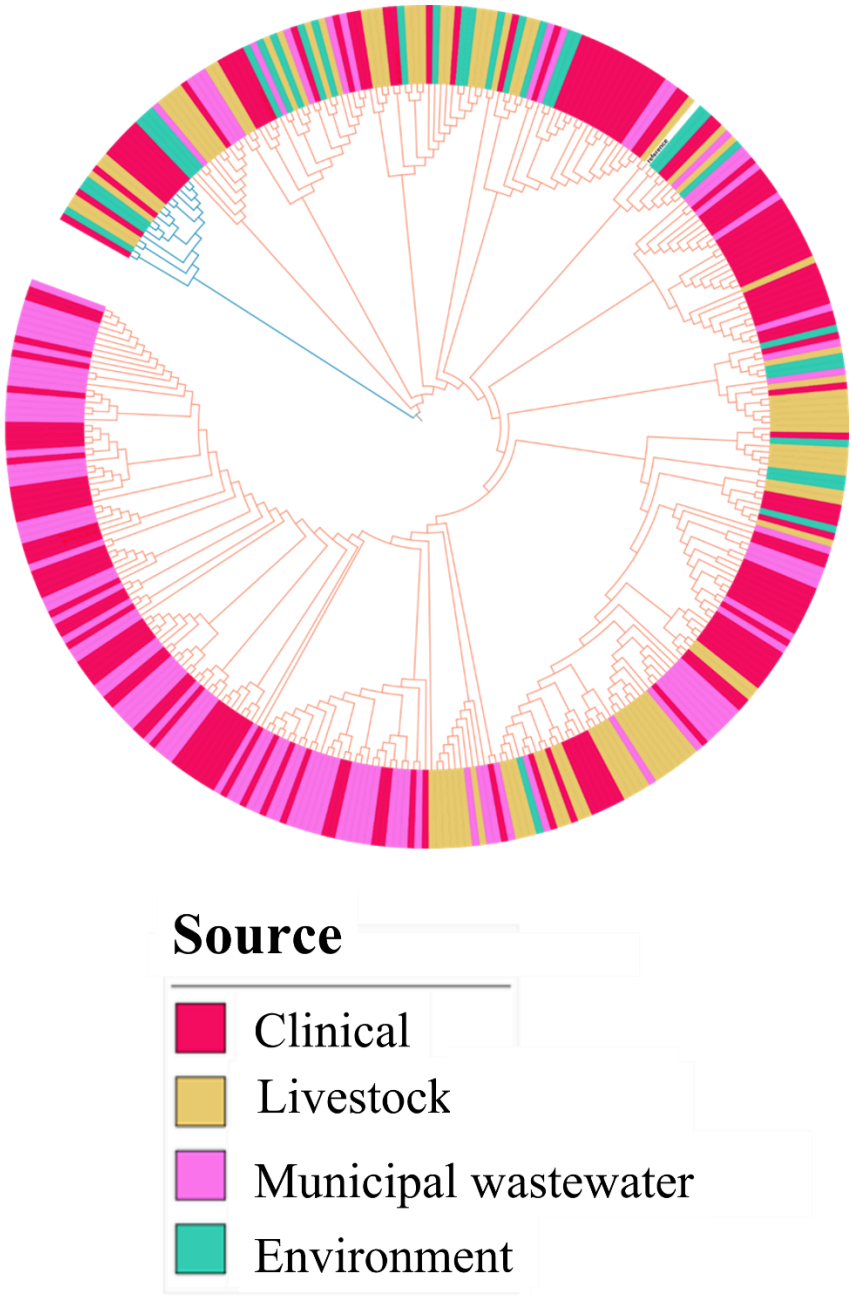

Figure 3.8 Maximum likelihood core-genome phylogenetic tree of 376 genomes *E. faecalis* genomes across the One-health continuum. *E. faecalis* ATCC 47077/OG1RF (CP002621.1) was used as a reference genome. Genomes were characterized based on their source of isolation into four groups: livestock, clinical, municipal wastewater and environmental.

3.5 Discussion

AMR is a serious concern for human and animal health and the global economy. One Health approaches to assess AMR recognize the role of multiple ecosystems in generating and spreading antimicrobial resistance genes (329). In One Health studies, *Enterococcus* species have been used as 'indicator bacteria' to monitor ARGs dissemination in ecosystems. In this study, we performed genomic characterization of *Enterococcus* species recovered in feces from weaners, finishers and sows. Furthermore, we evaluated the ARGs identified in *E. faecium* and *E. faecalis* genomes across livestock and poultry production systems and cumulatively across the overall One-health continuum.

E. hirae was predominantly identified in swine feces, followed by *E. faecium* and *E. faecalis*. In studies from the US and Canada, *E. hirae* was frequently recovered from livestock(217, 342) . In poultry, *E. faecium* has been isolated most frequently (343) and along with *E. faecium* and *E. faecalis* are often associated with human infections (217). In all identified enterococcal species, tetracycline resistance determinants *tetL and tetM* were frequently found on the mobile plasmid pM7M2 (NC_016009). This plasmid has been previously identified in *E. faecalis* isolated from dairy cattle feces and was shown to transfer into *Streptococcus mutans* UA159 through natural transformation (344). These findings show that three *Enterococcus* spp. (i.e., *E. faecium*, *E. faecalis* and *E. hirae*) can readily acquire ARGs in the gut micro-environment and possibly contribute to gene dissemination through plasmid-mediated ARG transfer.

We aimed to define the impact of differences in AMU (antimicrobial usage) across different livestock sectors on the occurrence of ARGs within enterococci. Across all livestock sectors, isolates from bovine sources were the least MDR, which may reflect the

extent of AMU in this livestock sector in Canada. According to the CIPARS 2019 report, most antimicrobials are administered to swine (<300 mg/PCU), followed by poultry (<200 mg/PCU) and cattle (<100 mg/PCU) (CIPARS, 2019). Regardless of the high MDR in poultry isolates, we did not find any isolates of poultry origin carrying ARGs conferring resistance to antimicrobials that were administered to poultry (Table 3.1). However, comparative genomics of enterococci identified that tetracycline and macrolide resistance genotypes were more prevalent in the beef production system compared to swine and poultry, a result that may reflect the greater use of these antimicrobials in beef cattle (345, 346).

Mobile genetic elements play a significant role in gene dissemination within and across ecosystems. In our study, all ARGs, except those that were intrinsic, were mapped to plasmids in almost 80% *E. faecium* and *E. faecalis* isolates. Resistance to aminoglycosides, tetracyclines, trimethoprim and MLS was identified across all ecosystems, with tetracycline and MLS being the most common. With these antimicrobials broadly used across sectors, the existence and persistence of resistant strains across the continuum is perhaps not surprising (347, 348). Their persistence may also be explained by the co-existence of these genes along with other ARGs, as well as others have found a strong association of tetracycline resistance ARGs (*tetL* and *tetM*) with other ARGs, including *ermB, ant(6)-la, aph(3')-IIIa*, *lnu(G), lsaE*, and *sat4* (349).. These ARGs were often found on MGEs that may facilitate their spread in different ecosystems. Continuous exposure to one antimicrobial class in a particular ecosystem can also select for ARGs conferring resistance to other antimicrobial classes (350-352).

Some antimicrobial resistance determinants were found in some sectors but not others. For example, *aac(6')-Ie/aph(2'')-Ia,* which is associated with high-level gentamicin resistance (HLGR), was only identified in *E. faecium* genomes from CL and MW. However, the association of this gene with MGEs may facilitate its spread to other human pathogens as it mapped to five different plasmids and was frequently associated with IS256 elements. Previously, *aac(6')-Ie/aph(2'')-Ia* was associated with IS256 on the *Tn*5281 composite transposon in a conjugative pBEM10 plasmid in *E.faecalis* (353), with Tn4001 on plasmid pSK1 in *Staphylococcus aureus* (354) and Tn4031 in *Staphylococcus epidermidis* (355). Glycopeptide-resistant genes *vanA* and *vanC* were identified in clinical and poultry isolates. The *vanA* operon was mapped to two plasmids in CL isolates, pV24-3 and pF856. Along with *vanA*-operon, other ARGs *ant(6)-Ia, aph(3')-IIIa*, *ermB* and *sat4* were also mapped to pF856. This particular plasmid was first reported in a hospitalized patient associated with a vancomycin-resistant *Enterococcus* outbreak in Ontario, Canada (356)
Our phylogenomic analysis revealed a similar topology of *gro-EL*-based (357) and core-genome-based trees, with *E. faecium* segregating into two main groups.

Our core-genome tree topology partitioned into two clades. In contrast, in a recent study by Sanderson et al. (358), clade B formed a paraphyletic clade rather than a monophyletic clade. Our findings also agree with previous studies (357, 358), as more ARGs and virulence genes were associated with clade A than clade B isolates. Furthermore, most of the genomes associated with CL isolates clustered in clade A. Phylogenetically, *E. faecalis* genomes did not cleanly partition into clades by source and instead formed multiple clades that originated from multiple sources.

In conclusion, our study suggests that some resistant strains are universally present in all ecosystems, irrespective of antimicrobial pressure. However, some ARGs are exclusive to particular ecosystems, reflecting antimicrobial usage within that sector. Moreover, we also found that co-selection and association of ARGs with different MGEs likely facilitate the spread of ARGs across the One Health Continuum. In addition, clinical *E. faecium* isolates formed a distinct cluster and were consistently mapped to a hospital associated clade'.

Chapter 4: Genomic Characterization of Carbapenem-Resistant Bacteria from Beef Cattle Feedlots.

Antibiotics vol. 12,6 960. 25 May 2023, 12, 960.
doi:10.3390/antibiotics12060960.

Author's Contributions

R.Z. and T.M., designed the study; S.-e.-Z.Z. and R.Z. analyzed sequence data; K.M. arranged for sample collection and bacterial isolation; S.A. performed short read sequencing; L.S., T. H. per-formed nanopore sequencing, M.S-E assembled nanopore reads, S.-e.-Z.Z. and S.A. performed CarbaNP testing, S.-e.-Z.Z. performed AST, genomic DNA extraction, generated figures and analyzed overall data/results and wrote the first draft of the manuscript; and T.M., A.Z., and R.Z. provided funding and supervision. All authors have read and agreed to the published version of the manuscript.

4.1 Abstract

Carbapenems are considered a last resort for the treatment of MDR bacterial infections in humans. In this study, we investigated the occurrence of carbapenem resistant bacteria in feedlots in Alberta, Canada (August 2016 to June 2019). The presumptive carbapenem resistant isolates (n=116) recovered after ertapenem enrichment were subjected to antimicrobial susceptibility testing against 12 different antibiotics including four carbapenems (ertapenem, meropenem, doripenem, and imipenem) and ceftazidime, chloramphenicol, gentamicin, levofloxacin, piperacillin, trimethoprim-sulfamethoxazole, tobramycin and tetracycline. Of these 72% of the isolates (n=84) showed resistance to ertapenem, while 27% of the isolates (n=31) were resistant to at least one other

carbapenem, with all except one isolate being resistant to at least two other drug classes. Of these 31 isolates, 90% were carbapenemase positive, while a subset of 36 ertapenem-only resistant isolates were carbapenemase negative. The positive isolates belonged to three genera; *Pseudomonas, Acinetobacter*, and *Stenotrophomonas* with the majority being *Pseudomonas aeruginosa* (n=20) as identified by 16S rRNA gene sequencing. Whole genome sequencing identified intrinsic carbapenem resistance genes including *blaOXA-50* and its variants (*P. aeruginosa*), *blaOXA-265* (*A. haemolyticus*), *blaOXA-648* (*A. lwoffii*), *blaOXA-278* (*A. junii*), and *blaL1* and *blaL2* (*S. maltophilia*). The acquired carbapenem resistance gene (*blaPST-2*) was identified in *P. saudiphocaensis* and *P. stutzeri*. In a comparative genomic analysis, clinical *P. aeruginosa*, clustered separately from those recovered from bovine feces. In conclusion, despite the use of selective enrichment methods, finding carbapenem resistant bacteria within feedlot environment was rare in Alberta.

4.2 Introduction

Carbapenems are β-lactam antibiotics that consist of a four-membered β-lactam ring fused with a secondary five-membered thiazolidine ring through the nitrogen and adjacent tetrahedral carbon atom. Unlike other β-lactams, carbapenems have two substitutions, at position one there is a substitution of sulfur for a carbon atom and at the fourth position of the thiazolidinic moiety, a carbon is substituted for a sulfone (359, 360). So far, four carbapenems, including ertapenem, meropenem, doripenem, and imipenem, have been approved for use in the US. These members differ in their side chains, influencing their antimicrobial activity. Carbapenems inhibit cell wall synthesis by preventing the formation of cross-linkages in peptidoglycan via binding to peptidoglycan binding protein (PBP),

thus leading to cell lysis and death (361). The ability of carbapenems to bind diverse PBPs with high affinity and their ability to resist extended-spectrum β-lactamases (ESBLs) and AmpC β-lactamases account for their broad-spectrum activity against both Gram-negative and Gram-positive bacteria (360, 362, 363). Ertapenem binds preferentially to PBPs 2 and 3 of *Escherichia coli* and has a low affinity for PBP 1a, 1b, 4, and 5. Imipenem binds with high affinity to PBP2, followed by PBP1a and 1b, but binds to PBP3 with low affinity. Meropenem possesses a high affinity for PBP 2, 3, and 4. Doripenem, similar to meropenem can bind to PBPs 2 and 3 of *P. aeruginosa* and also has affinity for PBP2 of *E. coli* (360). The use of carbapenem in human clinical settings increased due to rising resistance against penicillin, cephalosporins, fluoroquinolones and aminoglycosides (364). According to GLASS data on global antimicrobial consumption, approximately 9.8% (2.62 Defined Daily Dose/ 1000 individuals/ day) of all antimicrobial consumed globally in 2020 were belonged to "Other beta-lactam antibacterial (J01D)" class including carbapenems (365).

Bacteria may circumvent carbapenem hydrolysis through intrinsic or acquired resistance mechanisms including the production of β-lactamases, overexpression of efflux pumps and mutations that alter the expression or/and function of porin proteins and PBPs (366-368). The β-lactamases conferring resistance against carbapenem may belong to one of the Ambler classes including class A (e.g., KPC), class B (e.g., VIM, NDM), and class D (e.g., OXA-48) (369). The genes responsible for the production of β-lactamases are often associated with MGEs that can play a significant role in dissemination of carbapenem resistance in an ecosystem (370). Carbapenem resistance was first reported in opportunistic and environmental bacteria with intrinsic resistance (371). Since then, carbapenem-

resistant bacteria have been reported in the Asia-Pacific, India, Europe, and North and Latin America (372). Resistant isolates have been recovered from a variety of sources including human clinics (371), livestock, including dairy cattle (373, 374), beef cattle (375-377), and swine (378), sewage water (379), and wildlife (380, 381).

Consequently, carbapenem resistance in Gram-negative bacteria has become a global problem with carbapenem-resistant Enterobacteriaceae (CRE), *Pseudomonas aeruginosa*, and *Acinetobacter baumannii* listed by the World Health Organization as being on the "ESKAPE" list of pathogens for which the control and development of new antimicrobials is urgently needed (382). According to Public Health Canada "Findings from the 2022 Canadian Antimicrobial Resistance Surveillance System Report", the estimated incidence rates of healthcare-associated carbapenemase-producing Enterobacterales infections in 2020 was 0.0345 per 10,000 patients-day (383).

Considering the importance of carbapenem in human health, the objective of this study was to investigate if carbapenem-resistant bacteria could be recovered from beef cattle feedlots through carbapenem enrichment. Initially, our focus species was *E. coli,* as it serves as an indicator bacteria in an antimicrobial resistance surveillance program within the One-health context (134) . Later, recovered resistant isolates belonging to species other than *E. coli* such as *Pseudomonas spp.*, *Acinetobacter spp.*, *Ochrobactrum intermedium* (*Brucella intermedia*), and *Stenotrophomonas maltophilia* were also included in downstream characterization using phenotypic (AST and CarbaNP test) and genotypic (WGS) approaches. Furthermore, we investigated genomic relatedness among *Pseudomonas aeruginosa* isolates recovered in this study to previously published *P. aeruginosa* genomes from NCBI database through comparative analysis.

4.3 Methodology

4.3.1 Sampling, isolation and identification

The study includes bovine fecal and catch-basin water samples, collected from four feedlots (pens=301) located in Alberta, over a period of two years (August, 2016 to June, 2018). Briefly, 20 g of feces were collected from 20 fresh fecal pats and placed in a sterile plastic container and thoroughly mixed. Each pen housed approximately 180 feedlot cattle. These samples were transferred to the research facility and enrichment was done within 24 hours for sample collection. For enrichment, 0.5 g of the mixed fecal sample was inoculated into 4.5 ml of *E. coli* (EC) broth containing 2µg/ml of ertapenem (ETP, sigma Aldrich, ref: sml1238), followed by overnight incubation at 37°C in shaking incubator at 250 RPM. The enriched samples were then sub-cultured overnight on MacConkey agar supplemented with 2 µg/ml ertapenem. From each sample, a maximum of three colonies were selected and sub-cultured on nutrient agar supplemented with 0.5 µg/ml ertapenem (Dalynn Biologicals, Calgary, Canada). Water samples were collected from feedlot catch basins which collected run off water from the feedlot. Briefly, 1L of catch-basin water was collected into a polyethylene bottle attached to a telescopic pole at two different locations per site, which were combined to generate a composite sample. Water samples were transported on ice within 4 h of collection for processing. Composite catch basin water (10 mL) was filtered through a 0.45µm pore size filter of 47-mm diameter (S-Pak® EMD Millipore Corp, Billerica, MA) using a sterile vacuum-manifold filtration system (Pall Corporation, Port Washington, NY). The filter was then submerged in 4.5 mL EC broth-ertapenem (0.5mg/L) and incubated overnight at 37°C with shaking, followed by sub-culture onto MacConkey Agar supplemented with 0.5 µg/L ertapenem at 37°C (Dalynn

Biologicals, Calgary, Canada). A maximum of three colonies were selected from each sample (fecal samples= 301, water samples= catch basin water sample= 12) and sub-cultured on nutrient agar supplemented with 0.5 µg/ml ertapenem at 37°C for 24 h. For identification of recovered isolates (n=116), the 16S rRNA gene was amplified using universal bacterial 16S rRNA gene primers 27F (5′-AGAGTTTGATCCTGGCTCAG-3′) and 1492R (5′-GGTTACCTTGTTACGACTT-3′) followed by Sanger sequencing of the amplified PCR product (384). Species were identified using BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) search against the NCBI bacterial database.

4.3.2 Phenotypic Characterization

Isolates (n=116) were tested against 12 different antimicrobials including all four carbapenems (ertapenem, meropenem, doripenem and imipenem), ceftazidime, chloramphenicol, gentamicin, levofloxacin, piperacillin, trimethoprim-sulfamethoxazole, tobramycin and tetracycline, using the disk diffusion method according to the Clinical and Laboratory Standards Institute (CLSI) guideline M02-A12 and M100-S32. *Pseudomonas aeruginosa* ATCC 27853 was used as a reference quality control. Zones of inhibition were recorded using the BioMic V3 imaging system (Giles Scientific, Inc., Santa Barbara, CA, USA).

Isolates (n=116) were also tested for the production of carbapenemases using the chromogenic Carba NP test (RAPIDEC® CARBA NP kit, BioMérieux, St-Laurent, QC, Canada), with *Klebsiella pneumoniae* ATCC 700603 and *Klebsiella pneumoniae* OLC2685 used as negative and positive controls, respectively. A bacterial colony (10µl loop) was picked up from overnight-incubated Mueller-Hinton agar plates and mixed into API suspension medium. The bacterial suspension was then transferred to wells in a test

strips and incubated at 37°C. Test strips were read at 30 and 120 minutes. A 'positive' test corresponded to a color change from red to yellow-orange, while no-change in color was considered 'negative' for carbapenemase production.

4.3.3 Whole-genome sequencing, assembly and annotation

All Carba NP positive isolates (n=28) and a subset (n=14) of Carba NP negative isolates selected to be representative of identified species were subjected to whole genome sequencing using short and long reads sequencing technologies. High molecular weight genomic (HMW) DNA was extracted using a genomic DNA preparation kit with Genomic-tip 20/G (Cat: 13323; QIAGEN, USA) according to manufacturer's instructions. DNA quality and quantity were estimated using a Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific, Mississauga, ON, Canada) and a Qubit Fluorometer with PicoGreen (Q32850, Invitrogen, USA) respectively. The integrity of DNA was confirmed through agarose gel electrophoresis. Short-read sequencing was performed on the Illumina MiSeq platform. The genomic library was constructed using the Illumina NexteraXT DNA sample preparation kit (Illumina Inc., San Diego, CA, USA) followed by sequencing on the Illumina MiSeq platform using the MiSeq Reagent Kit V3, generating 2×300 base paired-end reads. Long-read sequencing was performed on PromethION platform from Oxford Nanopore Technologies (ONT). Sequencing libraries were prepared using LSK-109 genomic DNA preparation kit. One hundred nanograms of HMW genomic DNA were end-repaired using a NEBNext Ultra II End prep enzyme mix [Cat: E7646AA; New England Biolabs (NEB) Ltd. Whitby, ON, Canada] and FFPE DNA repair mix [Cat: M6630L; New England Biolabs (NEB) Ltd. Whitby, ON, Canada]. End-repaired DNA from each isolate was individually barcoded using a Nanopore barcoding kit EXP-NBD196 (Oxford

Nanopore Technologies (ONT), UK), followed by pooling and cleaning of barcoded DNA samples in a single tube for multiplexing. Pooled samples were cleaned using the Omega-bind NGS beads (M1378-01, Omegabiotek) at 0.5X volume following the manufacturer's instructions. Beads were dried for 1 min and DNA was eluted in 32 µL of deionized sterile H_2O. The DNA concentration was measured using the Invitrogen Qubit dsDNA BR assay (Q32854, Invitrogen). Sequencing adapters were then ligated to 30 µL of recovered DNA using Adapter Mix II Expansion kit EXP-AMII001 (ONT) along with the 5 µL T4 DNA ligase (NEB) [Cat: M0202M] and 10 µL quick ligation buffer [Cat: B6058S]. Adaptor-ligated DNA was bead-cleaned using a 0.4X volume of beads following manufacturer's instructions using the supplied polyethylene glycol (PEG) based wash buffer. The DNA from the last step (200-400 ng) was loaded onto an ONT PromethION sequencing flow cell as directed by the manufacturer. MinKNOW Core v. 3.1.20 and guppy v. 2.0.10 were used for flow cell signal processing and base calling during each run and reads were assembled de novo using Flye (https://github.com/fenderglass/Flye) (385).

Hybrid genome assemblies were generated using illumina short-reads and Flye-assembled contigs from ONT long-reads using the Unicycler assembly tool (179). Illumina paired-end short-read sequences were used to generate an assembly graph using SPAdes followed by bridge building using Flye-assembled contigs using Miniasm (https://github.com/lh3/miniasm) (386) and Racon (https://github.com/isovic/racon). Multiple rounds of short-read polishing was done using Pilon (https://github.com/broadinstitute/pilon). The contiguity and quality of each genome assembly was assessed using Quast (v. 5.2.0) by computing relevant metrics, including the number of contigs, total length (bp), and GC content (336). Contigs were then annotated

using Prokka v.v. 1.13.1 (337) and assembled contigs were screened for the presence of antimicrobial resistance and virulence genes using ABRicate v.v. 1.0.1 (https://github.com/tseemann/ABRICATE) against the NCBI Bacterial Antimicrobial Resistance Reference Gene Database (NCBI BioProject ID: PRJNA313047) and the VirulenceFinder database (PMID: 34850947) (338), respectively. To identify plasmids, Mob-recon tool v.v. 3.0.0 (https://github.com/phac-nml/mob-suite) was used (339).

4.3.4 Comparative genomic analysis

Whole genome comparative genomic analysis was conducted between all the *Pseudomonas aeruginosa* isolates (n=20) sequenced in this study to *P. aeruginosa* genomes (n=76) (Appendix 3; Table S1) originating from the North America, from the PathoSystems Resource Integration Center (PATRIC) (https://www.patricbrc.org; accessed on February 20, 2023). Although these same genomes were also available from NCBI database, the metadata (origin, source of isolation) in the PATRIC database was more detailed. These isolates originated from human clinical sources with the exception of a single isolate from cattle (CP013989).

The core-genome phylogenomic tree was constructed using the SNVphyl pipeline v. 1.2.3 (387). The phylogenetic tree was generated by aligning paired-end reads against the *P. aeruginosa* PAO1 reference genome (NC_002516.2) using SMALT (v. 0.7.5; https://sourceforge.net/projects/smalt/). The generated read pileups were then subjected to quality filtering (minimum mean mapping quality score of 30), coverage cut-offs (15× minimum depth of coverage), and a single nucleotide variant (SNV) abundance ratio filter of 0.75 to obtain a multiple sequence alignment of SNV-containing sites. The SNV alignment, with no density filtering was used to create a maximum likelihood phylogeny

using PhyML v.v. 3.0. The generated Newick file was visualized using the Interactive Tree Of Life (iTOL) v.6 (231).

The *blaOXA-50* gene and its variants identified in bovine (this study) and human (NCBI genomes) isolates were also compared to determine their genetic relatedness. For this, *blaOXA-50* coding sequencing (CDS) was computationally extracted from each genome and aligned using MAFFT v.v. 7.490 (388). The resultant alignment based tree was visualized using iTOL v. 6 (231).

4.4 Results

4.4.1 Recovery of carbapenem resistant isolates and species identification

A total of 116 presumptive carbapenem resistant isolates were recovered from bovine fecal and catch basin samples. Among these 8 were *E. coli*, 100 were *Pseudomonas* spp., 5 were *Acinetobacter* spp., 2 were *Ochrobactrum intermedium* (*Brucella intermedia*) and 1 was identified as *Stenotrophomonas maltophilia.*

4.4.2 Phenotypic characterization of carbapenem resistant bacteria

Figure 4.1 shows AMR profiles of 116 isolates tested using antibiotic discs. Of all the tested *E. coli* (n=8) isolates, one isolate was resistant to ertapenem and tetracycline, while all others exhibited an intermediate resistance to ertapenem. All *E. coli* isolates were found to be negative for carbapenemase production(Figure 4.1).

All *Acinetobacter* spp. isolates (n=5) were resistant to both ertapenem and meropenem and positive for carbapenemase production. One isolate of *Acinetobacter* spp. was found to also

be resistant to tetracycline and trimethoprim-sulfonamide. Tobramycin resistance was also identified in two *Acinetobacter* spp. isolates recovered from bovine feces (Figure 4.1).

For *Pseudomonas* spp, 7% of isolates were resistant to meropenem alone, 20% were resistant to both doripenem and meropenem, and 11% were resistant to all tested carbapenems. After carbapenems, *Pseudomonas spp.* were most commonly resistant to trimethoprim-sulfonamides (33%) followed by chloramphenicol (30%), tetracycline (23%), ceftazidime (8%), gentamicin (6%), tobramycin (6%), piperacillin (5%) and levofloxacin (5%) (Figure 4.1). A total of 56 *Pseudomonas* spp. were tested for the production of carbapenemase and 38% of these isolates were confirmed positive (Figure 4.1).

A single *O. intermedium* isolate was resistant to ertapenem and meropenem, with both *O. intermedium* isolates being negative for carbapenemase production. The single *S. maltophilia* isolate identified in this study was resistant to all tested antibiotics and was positive for carbapenemase production (Figure 4.1 and Table 4.1).

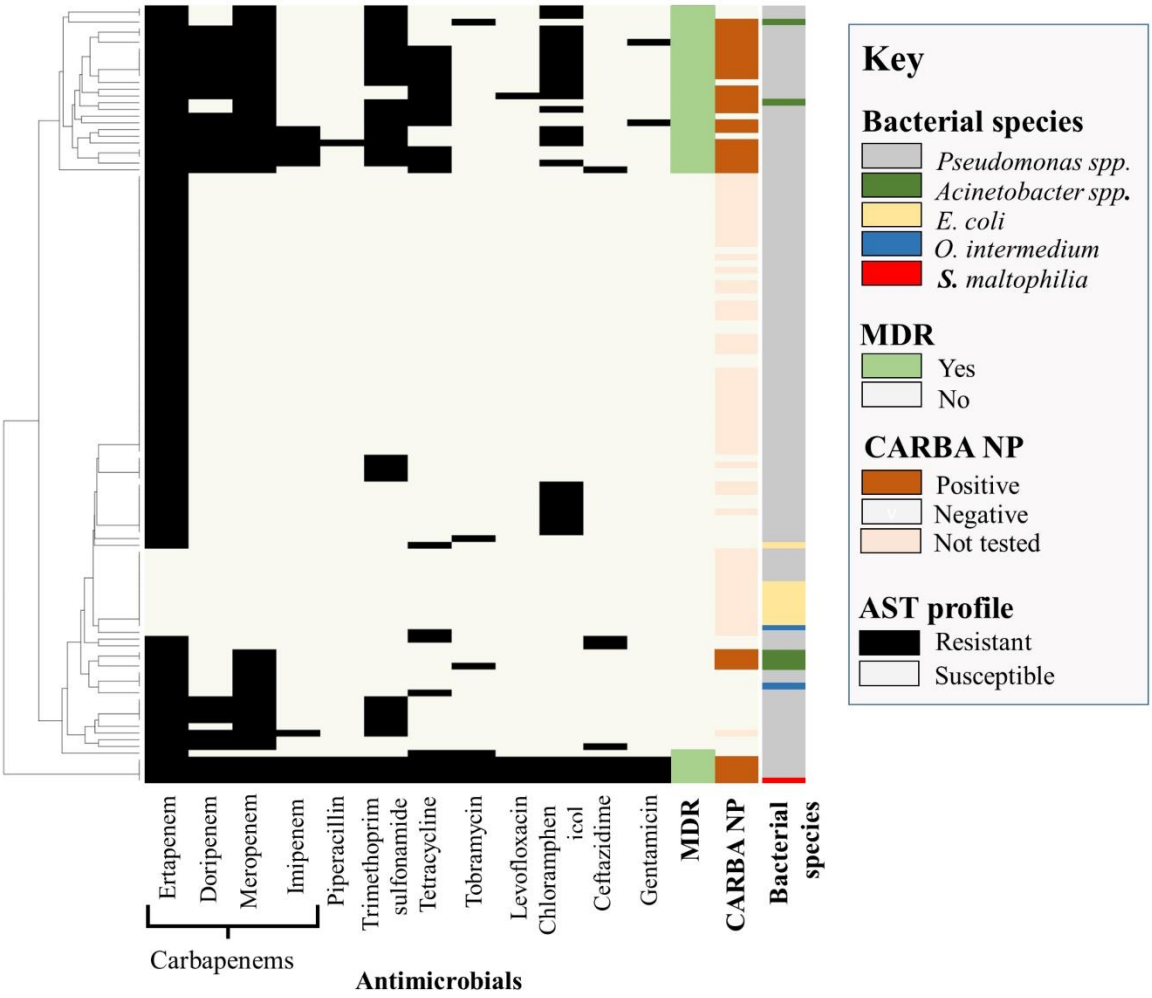

Figure 4.1 Antimicrobial susceptibility and carbapenemase production phenotypes of isolates recovered from beef cattle feedlots following carbapenem enrichment in carbapenem. Far right panel represents carbapenem bacterial species recovered from bovine and catch basin water samples. Next panel represents CarbaNP results where positive isolates were shown in brown and negative isolates were shown in grey. The heatmap showed phenotypic AST profiles of isolates where black blocks showed phenotypic resistance. Grey blocks represents susceptibility to particular antimicrobials. The tree showed association among phenotypic resistance profiles.

Table 4.1 AMR profiles of isolates recovered from bovine feces and catch basin water samples using WGS analysis.

Bacterial Species	**Antimicrobial resistance genes**	**Phenotype-based on Genotype**
E. coli (n=01)	*blaEC, blaCMY-2, aph(3'')-Ib, aph(6)-Id, sul2, tet(A), floR*	Aminoglycosides Chloramphenicol, Sulfisoxazole, Tetracycline
A. haemolyticus (n=3)	*blaOXA-265, aacA-ACI, blaPDC-197*	Carbapenem Aminoglycosides Cephalosporin
	blaOXA-265, aacA-ACI1	Carbapenem Aminoglycosides
A. lwoffii (n=1)	*blaOXA-648*	Carbapenem
A. junii (n=1)	*blaOXA -278*	Carbapenem
S. maltophilia (n=01)	*blaL1,blaL2, aph(6)-Smalt, aph(3')-IIc, oqxB9, oqxA10, floR2*	Carbapenem Aminoglycosides Phenicol Chloramphenicol
O. intermedium (n=01)	*floR, oqxB12, blaOCH-2*	Chloramphenicol Quinolone Cephalosporin
P. aeruginosa (n=20)	*blaOXA-50, blaPDC-197, aph(3')-IIb, catB7, fosA*	Carbapenem Cephalosporin Chloramphenicol Aminoglycoside Fosfomycin
	blaOXA-50, blaPDC-197, aph(3')-IIb, catB7, fosA, crpP	
	blaOXA-50, blaPDC-55, aph(3')-IIb, catB7, fosA	
	blaOXA-50, blaPDC-66, aph(3')-IIb, fosA, crpP, catB7	
	blaOXA-486, blaPDC-374, aph(3')-IIb, fosA, crpP, catB7	

	blaOXA-486, blaPDC-374, aph(3')-IIb, catB7, fosA	
	blaOXA-494, blaPDC-374, aph(3")-la, aph(6)-ld, catB7 fosA, crpP	
	blaOXA-902, blaPDC-133, aph(3')-IIb, catB7, fosA, crpP	
	blaOXA-902	Carbapenem
P. entomophila (n=01)	*blaPDC-33*	Cephalosporin
P. plecoglossicida (n=09)	No gene	-
P. mosselii (n=01)	No gene identified	-
P. putida (n=01)	No gene identified	-
P. saudiphocaensis (n=01)	*blaPST-2, aadA1*	Carbapenem, Aminoglycoside
P. stutzeri (n=01)	*blaPST-2*	Carbapenems

4.4.3 Genomic characterization

Illumina short read sequencing generated reads with an average sequencing coverage of 80% per isolate (Appendix 3; Table S2). Hybrid assembly using Illumina short and Flye-assembled long reads generated an average of 6 contigs per *P. aeruginosa* genome, with an average genome size of 6,581,091 bp and GC% content of 66.23%. The average genome sizes for *P. plecoglossicida* and *A. haemolyticus* were 5,836,945 bp (GC content= 62.24%) and 3,327,176 bp (GC content= 39.71%). The genome sizes of other *Pseudomonas* spp. isolates belonging to species, *P. mosselii*, *P. entomophila*, *P. putida*, *P. stutzeri*, and *P. saudiphocaensis* were 5,724,129 bp (GC content= 64.56%), 5,483,085 (GC content= 62.55%), 5,784,665 bp (GC content= 61.80%), 4,059,836 bp (GC content=63.28%), and 3,671,120 bp (GC content= 61.14%), respectively. Genomes of *Acinetobacter lwoffii* and *A. junii* had 3,648,566 bp (GC content= 40.38%) and 3,109,155 bp (43.13%), while *O. intermedium* and *S. maltophilia* genomes were 4,881,352 bp (GC content=57.53%) and 5,264,819 bp (GC content=66.60%) bp. The whole genome sequence data of the 42 bacterial isolates have been deposited in GenBank under BioProject PRJNA956966.

Sequencing of the ertapenem resistant *E. coli* isolate (n=1) revealed no antimicrobial resistance genes (ARG) associated with carbapenem resistance. However, identified ARGs associated with β-lactams (*blaEC*), cephalosporins (*blaCMY-2*), sulfonamide (*sul2*), aminoglycosides (*aph(3")-Ib*, *aph(6)-ld*), tetracycline (*tetA*), fluoroquinolones (*floR*) were identified (Table 4.1). All ARGs but *blaEC* were mapped to a conjugative plasmid (CP025245). The virulence genes associated with type II secretion system (T2SS), type III

secretion system (T3SS), pili and fimbriae synthesis, adhesion, iron import, curli biogenesis and enterobactin synthesis were also present in the *E. coli* genome. (Table 2).

A total 34 *Pseudomonas* spp. were sequenced including *P. aeruginosa* (n=20), *P. plecoglossicida* (n= 09), *P. entomophila* (n=1), *P. mosselii* (n=1), *P. putida* (n=1), *P. saudiphocaensis* (n=1) and *P. stutzeri* (n=1). All *P. aeruginosa* isolates, carried *blaOXA*-type carbapenemase (class D) gene [*blaOXA-50* (13/20, 65%), *blaOXA-486* (3/20, 15%), *blaOXA-494* (1/20, 5%), *blaOXA-902* (1/20, 5%), *blaOXA-648* (1/20, 5%)] (Table 4.1). Additionally, we identified ARGs associated with cephalosporin [*blaPDC-197* (9/20, 45%), *blaPDC-374* (3/20, 15%), *blaPDC-55* (3/20, 15%), *blaPDC-133*(1/20, 5%), *blaPDC-66* (1/20, 5%)], aminoglycoside [*aph(3')-Iib* (18/20, 90%), *aph(3'')-Ia* (1/20, 5%), *aph(6)-Id* (1/20, 5%)], chloramphenicol [*catB7* (18/20, 90%)], fluoroquinolone [*crpP* (4/20, 20%)] and fosfomycin [*fosA* (18/20, 90%)] resistance in *P. aeruginosa*. None of these ARGs were plasmid-associated. In *P. saudiphocaensis* and *P. stutzeri*, the carbapenem resistant gene *blaPST-2* was present, but we did not find any carbapenem resistant genes in carbapenemase negative *Pseudomonas* species (Table 4.1). In *P. aeruginosa*, an average of 225 virulence genes were identified per isolate, whereas other *Pseudomonas* spp. lacked virulence genes (Table 4.2). Virulence genes were associated with biofilm formation, secretion system, pili and fimbriae formation.

Five carbapenemase positive *Acinetobacter* spp. isolates were sequenced, with three isolates being identified as *A. haemolyticus* and the other two as *A. lwoffii* and *A. junii.* At least one carbapenem resistant gene was identified in all *Acinetobacter* spp. including; *blaOXA-265* (*A. haemolyticus*, n=3), *blaOXA-648* (*A. lwoffii*, n=1) and *blaOXA-278* (*A.*

junii, n=1). In *A. haemolyticus*, the aminoglycoside (*aacA-ACI*) and cephalosporin (*blaPDC-197*) resistance genes were also identified (Table 4.1).

In carbapenemase positive *S. maltophilia*, genes associated with carbapenem (*blaL1*, *blaL2*), aminoglycosides (*aph(6)-Smalt*, *aph(3')-IIc*), and chloramphenicol (*floR2*) were identified (Table 4.1). We also identified a multidrug efflux Resistance-nodulation-division (RND) transporter operon (*oqxB9, oqxA10*) responsible for quinolones resistance. This multidrug efflux RND transporter was mapped twice on *S. maltophilia* genome along with *aph(3')-IIc*. As expected, we did not identify any carbapenemase-associated ARGs in *O. intermedium*, as this isolate was negative for carbapenemase production. However, ARGs conferring resistance to chloramphenicol (*floR*), quinolones (*oqxB12*) and cephalosporin (*blaOCH-2*) were identified in this isolate (table 4.1).

Hybrid genome assemblies allowed complete circular genomes to be constructed for some of the isolates (Appendix 3; Table S1). Moreover, some ARGs including *blaOXA-265* in *A. haemolyticus,* and *blaPST-2* in *P. saudiphocaensis* were only identified after hybrid assemblies were generated.

Table 4.2 Virulence gene profiles of whole genome sequenced isolates recovered from bovine feces and catch-basin water samples.

Bacterial species	Virulence genes[14] (%)
P. aeruginosa (n=20)	**Biofilm and capsule synthesis:** *alg44* (21/24, 88%), *alg8* (22/24, 92%), *algA-G,I-L,P-R,U,W,X,Z* (22/24, 92%), *mucA-E,P* (21/24, 88%) **Pili and Fimbriae:** *chpA*-E (92%), *pilA* (9%), *pilB* (88%), *pilC* (34%), *pilE,F* (84%), *pilG-K,M-X,Y1,Y2* (92%), *fimT-V* (84%), *fleI/flag* (13%), *fleN* (100%), *fleP* (13%), *fleQ* (100%), *fleR* (88%), *fleS* (88%), *flgA* (84%), *flgB* (88%), *flgC* (100%), *flgD-F* (84%), *flgG-I* (24), *flgJ,K* (84%), *flgL* (13%), *flgM,N* (84%), *flhA,B,F* (100%), *fliA,C* (100%), *fliD* (13%), *fliE-L* (92%), *fliM,N* (100%), *fliO-R* (88%), *fliS* (13%), *motA-D,Y* (92%) **Biosynthesis of small ferric-ion-chelating molecules:** *pchA-I,R* (84%), *fptA* (84%), *fpvA* (21%), *mbtH-like* (100%), *pvcA-D* (84%), *pvdA* (84%), *pvdD,E* (21%), *pvdF-H* (88%), *pvdI,J* (21%), *pvdL,M* (100), *pvdM* (22), *pvdN-Q* (88%), *pvdS* (96%) **Phenazine biosynthesis:** *phzA1-G1,M,S* (80%), *phzH* (30%) **Rhamnolipids:** *rhlA,B,I* (84%), *rhlC* (75%) **Type VI secretion system:** *clpV1* (96%), *dotU1* (88%), *flha1* (80%), *hcp1* (96%), *hsiA1* (84%), *hsiB1/vipA* (100%), *hsiC1/vipB* (96%), *hsiE1,F1,H1* (84%), *hsiG1* (96%), *hsiJ1* (92%), *icmF1/tssM1* (84%), *tagQ* (88%), *tse1-3* (84%), *vgrG1a* (80%), *ppkA*(84%), *tagR* (100%) *tagS* (88%), *tagT* (88%), *pppA* (84%), *tagF/pppB* (84%)**Type III secretion system:** *pscB-L, N-U* (80%), *popB,D,N* (80%), *pcr1-4,D,H,R,V* (80%), *exsA*-E (80%), *exoS,T,Y* (80%), *ptxR* (84%) **Type I secretion system:** *aprA* (84%) **Type II secretion system (Xcp) and exo-proteins:** *xcpA/pilD* (88%), *lasA* (88%), *lasB* (84%), *plcH* (84%), *toxA* (21%), *lip1* (84%) **Quorum sensing:** *lasI* (88%) **Lipopolysaccharide core biosynthesis:** *waaA,C* (88%), *waaF* (96%), *waaG,P* (100%), *wzy* (17%), *wzz* (17%)
P. entomophila (n=1)	**Biofilm and capsule synthesis:** *algCBU, mucD* **Lipopolysaccharide core biosynthesis:** *waaF, wag* **Type VI secretion system:** *clpV1, hsiG1, hcp1, hsiC1/vipB, hsiB1/vipA, tagR* **Biosynthesis of small ferric-ion-chelating molecules:** *pvdH, pvdS, mbtH-like* **Pili and Fimbriae:** *motC, fleNQ, flhA, fliAIGMNPQ, flgCGHI, pilH*
P. mosselii (n=1)	**Biofilm and capsule synthesis**: *algA-D,U,W,I* (100%), *alg8* (100%), *mucD* (100%) **Biosynthesis of small ferric-ion-chelating molecules**: *mbtH-like* (100%), *pvdH,S,M* (100%)

[14] Bold indicates virulence genes associated with cellular component or metabolic activity of each specific bacterial species.

	Lipopolysaccharide core biosynthesis: *waaF,G,P* (100%) **Pili and Fimbriae:** *flgC,G-I* (100%), *fliA,F,G,I,M,-Q* (100%), *fleN,Q* (100%), *motA-C* (100%), *pilH* (100%) **Type VI secretion system:** *tagR, dotU1* (100%), *hsiB1/vipA* (100%), *hsiC1/vipB* (100%), *hcp1*(100%), *hsiG1* (100%), *clpV1* (100%)
P. putida (n=2)	**Biofilm and capsule synthesis**: *algA-D,U,I* (100%),*alg8* (100%), *mucD* (100%) **Biosynthesis of small ferric-ion-chelating molecules**: *mbtH-like* (100%), *pvdH,S* (100%) **Lipopolysaccharide core biosynthesis:** *waaF,G* (100%) **Pili and Fimbriae:** *flgC,G-I* (100%), *fliA,G,I,M,N,P,Q* (100%), *fleN,Q* (100%), *flhA* (100%), *motC,D* (100%), *pilH* (100%)
P. saudiphocaensis (n=1)	**Biofilm and capsule synthesis:** *algC,R,U* (100%) **Pili and Fimbriae:** *flgC,G,I* (100%), *flhA* (100%), *fliE,G,I,M-P* (100%), *pilG,H,U,T* (100%) **Type II secretion system (Xcp):** *xcpT,R* (100%) **Lipopolysaccharide core biosynthesis:** *waaF*
P. stutzeri (n=1)	**Biofilm and capsule synthesis**: *algA-C,R* (100%) **Pili and Fimbriae:** *flgG,I* (100%), *flhA* (100%), *fliA,E-G,I,M,N-R* (100%), *fleN,Q* (100%), *motA* (100%), *pilG,H,J,M,R,T,U* (100%) **Type II secretion system (Xcp):** *xcpT,R* (100%) **Lipopolysaccharide core biosynthesis:** *waaF,P* (100%)
P. plecoglossicida (n=9)	**Biofilm and capsule synthesis**: *algB,C,U* (100%), *algW* (67%), *mucD* (100%) **Pili and Fimbriae:** *flgC,G,H,I* (100%), *flhA* (100%), *fliA,G,I,M,N,P,Q* (100%), *fleN,Q* (100%), *motC* (100%), *pilH* (100%) **Lipopolysaccharide core biosynthesis:** *waaF* (67%), *waaG* (100%) **Type VI secretion system:** *clpV1* (100%), *hcp1* (100%), *hsiB1/vipA* (100%), *hsiC1/vipB* (100%), *hsiG1* (100%), *tagR* (100%) **Biosynthesis of small ferric-ion-chelating molecules:** *mbtH-like* (100%), *pvdSH* (100%)

E. coli (n=1)	**Type II secretion system:** *gspC-M* (100%) **Type III secretion system:** *espX2* (100%), *ompA* (100%), *espR1*(100%), *espR4* (100%), *espR3* (100%), *espL1* (100%), *espY3* (100%), *espY2* (100%), *espX1*(100%), *espY4* (100%), *espL4* (100%), *espX4* (100%), *espX5* (100%), **Curli biogenesis:** *csgB,D,F,G* (100%) **Iron import system:** *shuA,S,T,W,X* (100%), *chuUVW* (100%) **Pili and Fimbriae:** *fimA-H* (100%), *yagV/ecpE* (100%), *yagW/ecpD* (100%), *yagX/ecpC* (100%), *yagY/ecpB*, *yagZ/ecpA* (100%), *ykgK/ecpR* (100%) **Adhesion:** *fdeC* (100%) **Enterobactin:** *entA-E,F,S* (100%), *fepA-D,G* (100%)
A. haemolyticus (n=3)	No virulence gene
A. lwoffii (n=1)	
A. junii (n=1)	
S. maltophilia (n=1)	

4.4.4 Comparative genomic analysis

Core-genome based phylogenetic tree of *P. aeruginosa* isolates formed two clades where the majority of *P. aeruginosa* recovered from bovine sources clustered together in a paraphyletic clade (Figure 4.2). In human clinical isolates, collectively, we identified 20 different ARGs conferring resistance to aminoglycosides [*aac(3)-Ic, aac(3)-Id, aac(3)-VIa, aac(6')-Ib4, aac(6')-Ib-AKT, aac(6')-Ib-G, aac(6')-IIa, aac(6')-IIc, aac(6')-Il, aacA8, aadA1, aadA5, aadA6, ant(2'')-Ia, ant(4')-IIb, aph(3')-Ib, aph(3'')-Ib, aph(3')-IIb, aph(3')-VIa, aph(6)-Id*] while in bovine *P. aeruginosa* only 3 genes, *aph(3'')-la*, *aph(3')-IIb*, and *aph(6)-ld* were present (Appendix 3; Table S3). All clinical isolates carried *blaOXA-50* variants (Figure 4.3). In addition, metallo-carbapenemase gene *blaNDM-1* was found in two clinical genomes (2/75; 3%), but not in any of the bovine isolates. We also identified ARGs for cephalosporins (*blaCTX-M-30, blaGES-9*, *blaOXA-10, blaOXA-101*, *blaOXA-2, blaPDC-15, blaPDC-16, blaPDC-22*, *blaPDC-34, blaPDC-38, blaPDC-46, blaPDC-55, blaPDC-59*, *blaPDC-100, blaPDC-101*, *blaPDC-116, blaPDC-123*, *blaPDC-167, blaPDC-264, blaPDC-308, blaPDC-364*, *blaPDC-374*, *blaPDC-385*, *blaVEB-9*), chloramphenicol (*cmlA5, catB, catB7, catB, floR2*), colistin (*mcr-1*), quinolones (*crpP, qnrVC1*), fosfomycin (*fosA*), macrolides [*mcr(E), mph(E)*], sulfonamide (*sul1*), tetracycline[*tet(G)*] and trimethoprim (*dfrA1, dfrA10, dfrB5*) in human clinical *P. aeruginosa* genomes. Comparison of *blaOXA-50* gene variants and their genomic context (genes found in close proximity) did not reveal any significant genome variations across genomes from human and bovine sources (Figure 4.3).

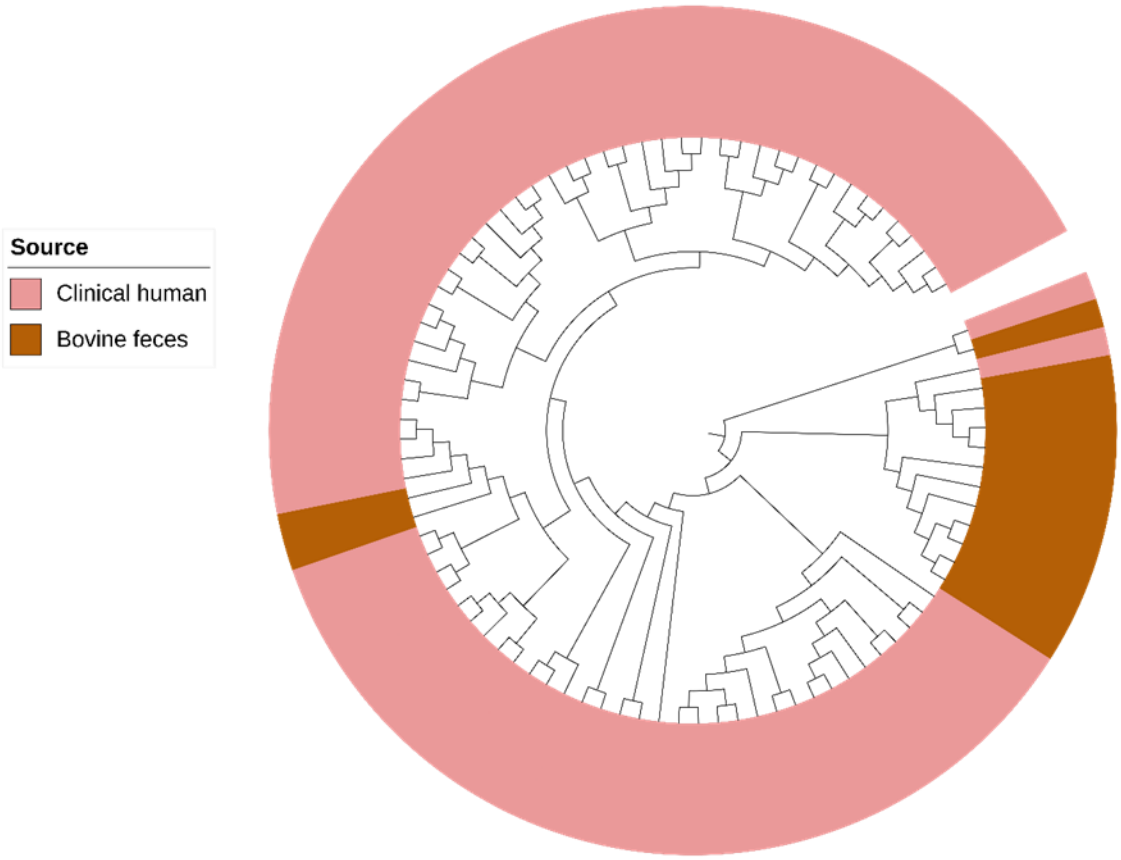

Figure 4.2 Maximum likelihood core-genome phylogenetic tree of *Pseudomonas aeruginosa* recovered from bovine and human clinical sources. The *Pseudomonas aeruginosa* POAO1 (NC_002516.2) was used as a reference genome.

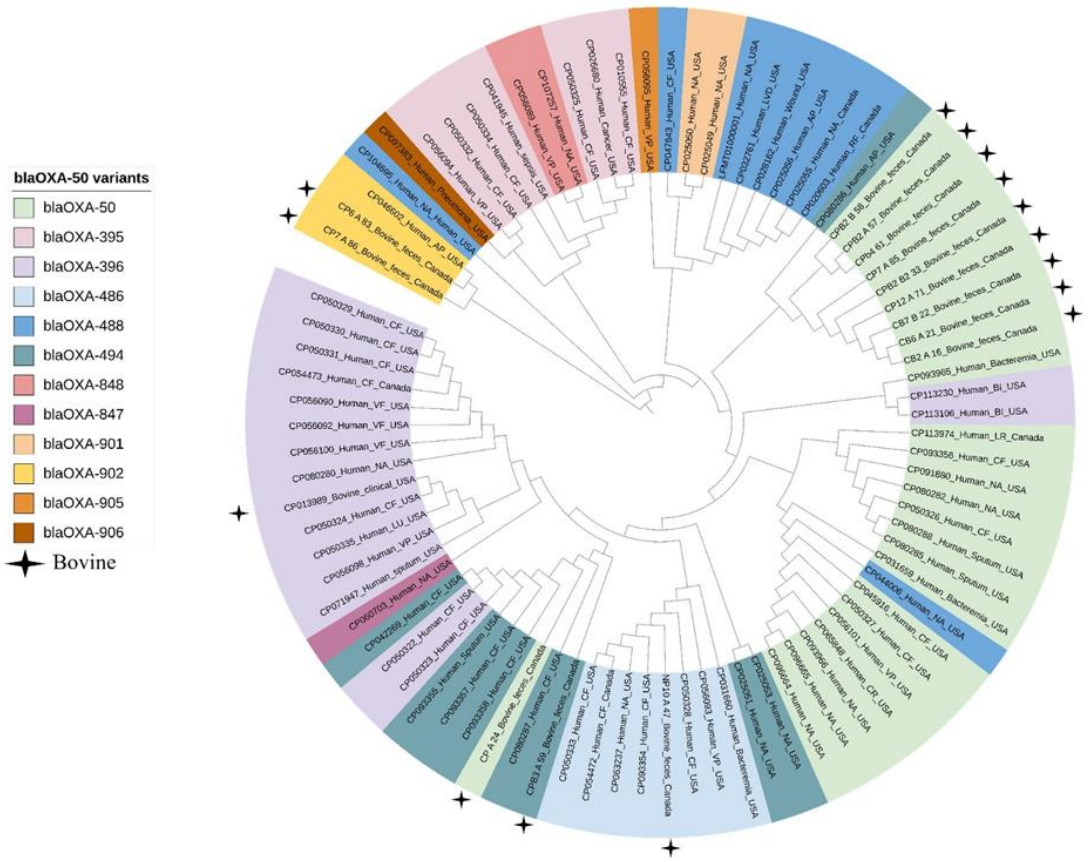

Figure 4.3 Phylogenetic tree of carbapenemase gene *blaOXA-50* variants identified in *Pseudomonas aeruginosa* isolates from human and bovine sources, constructed using MAFFT.

4.5 Discussion

Increasing carbapenem resistance has threatened the clinical utility of these drugs in human medicine, leading to the challenge of "extreme drug resistant" bacteria (389). Therefore, the Centers for Disease Control and Prevention (CDC) has declared carbapenem resistant Enterobacteriaceae as an urgent and serious threat to human health (390). In this study, we investigated if carbapenem resistant *E. coli* could be recovered using carbapenem enrichment from bovine feces or catch basin water samples collected in intensive feedlots. We had extremely low recovery of carbapenem resistant *E. coli* in bovine feces in investigated feedlots located Alberta. The majority of isolates that were presumed to be *E. coli* were subsequently identified as other bacterial species, primarily *Pseudomonas* species. This is not surprising as we used etrapenem for sample enrichment and *Pseudomonas* spp. are intrinsically resistant to this antimicrobial (391, 392).

In *Acinetobacter* spp., the carbapenem resistance genes identified were mostly intrinsic (393). For instance, *blaOXA-214, blaOXA-278 and blaOXA-648* in *A. haemolyticus, A. junii A. lwoffii* isolates, respectively are intrinsic ARGs in these species (394, 395). *A. lwoffii* is a clinically important *Acinetobacter* species, however, *A. haemolyticus* and *A. junii* are opportunistic only occasionally reported in human infections (396). The *blaL1 and blaL2* genes encoding for metallo-β-lactamase in *S. maltophilia* are also intrinsic and confer broad spectrum resistance against β-lactams including all carbapenems (397, 398). *S. maltophilia* is considered a “newly emerging pathogen of concern” and were frequently isolated from immunocompromised patients (398-400).

In *P. aeruginosa*, we identified four different genes variants belonging to the *blaOXA-50* class (*blaOXA-50, blaOXA-902, blaOXA-486* and *blaOXA-494*) encoding carbapenem-

hydrolyzing oxacillinase (CHDLs). All of these variants were also present in clinical isolates in addition to seven other variants *blaOXA-395*, *blaOXA-488*, *blaOXA-847*, *blaOXA-848* *blaOXA-901*, *blaOXA-905* and *blaOXA-906*. It has been reported that *blaOXA-50* naturally exists in all *P. aeruginosa* and does not appeared to be acquired based on the similar GC% content of the *blaOXA-50* gene to the overall *P. aeruginosa* genome (391). We did not identify known insertion sequences, repeat elements, relaxases, integrases, within the genomic proximity of *blaOXA-50*, suggesting that this gene may be less likely mobilized and is unlikely to be readily transferred to other bacterial species.

The majority of *P. aeruginosa* isolates from bovine sources clustered separately from those obtained from humans in the core-genome phylogenetic analysis. Similar observations had been reported in Brazil, where *P. aeruginosa* recovered from healthy bovine urine samples segregated from human clinical urinary tract isolates (401). These observations might be associated with a prolonged association and adaptation of these isolates within their respective environmental hosts. *P. aeruginosa* is known for its remarkable ability to adapt to diverse ecological niches from soil to various living hosts (402, 403). Genes associated with metabolism and pathogenesis constitute the core-genome of *P. aeruginosa*, whereas genes required to adapt to various niches constitute accessory genome. These genes were found in clusters in certain loci referred as 'regions of genomic plasticity' (404, 405). In this study, *P. aeruginosa* recovered from the beef production system were carrying ARGs conferring resistance to chloramphenicol (*catB7*), β-lactams (*blaOXA-50*), fosfomycin (*fosA*), aminoglycosides (*aph(3')-IIb*), cephalosporins (*blaPDC-55*, *blaPDC-374*). These ARGs were also found in the genomes of human clinical isolates. These ARGs were found on chromosomes in both bovine and human clinical isolates. In *P. aeruginosa,* the type IV

pilus uptake foreign DNA during transformation events (406, 407). The accessory genome can be distinguished from the core genome by its aberrant GC content, codon usage, and tetranucleotide usage (408). The GC content of *P. aeruginosa* (~66.3%) is mostly higher than the foreign DNA. Over time, the acquired DNA may lose the sequence compositional difference that distinguished it from the core genome of *P. aeruginosa* as it underwent the same pressure as the core genome (409).

We observed variable phenotypic susceptibility profiles among *P. aeruginosa* isolates that harbored similar AMR profiles. Some isolates were resistant to all carbapenems, some were resistant to both meropenem and doripenem and some were only resistant to meropenem. Other than β-lactamase production, alterations or lack of porin OprD, and overexpression of resistance-nodulation-division (RND) efflux pumps (MexAB-OprM, MexCD-OprJ MexEF-OprN and MexXY-OprM) are also associated with carbapenem resistance in *Pseudomonas species* (410, 411). Not only the combinations of these mechanisms confer reduced susceptibility to carbapenems, the overexpression of these efflux pumps results in β-lactam, tetracycline, trimethoprim, aminoglycosides and fluoroquinolone resistance (410, 412, 413). It is more likely that the variation in susceptibility profiles in our isolates is associated with the varied number of *MexAB-OprM* operon and *OprD* found in each isolate (414, 415). Moreover, isolates with multiple *MexAB-OprM* operons also possessed a single copy of *MexR* repressor gene. The presence of only one repressor gene in comparison to multiple efflux pumps as regulatory targets may disturb the molecular stoichiometry of the regulation affecting efflux pump expression, thus, resulting in varied phenotypic resistance profiles among *P. aeruginosa* isolates. The effect of efflux pump and respressor protein on phenotypic resistance profile

need to be explore in future. *P. stutzeri* (*Stutzerimonas stutzeri*) is an opportunistic pathogen that rarely causes infection in humans (416). In our study, we found the carbapenem resistant gene, *blaPST-2* in the *P. stutzeri* genome. The *blaPST-2* encodes for a subclass B1 metallo- β-lactamase and was first identified on the chromosome of *P. stutzeri* DSM 10701 (417). We did not find genes associated with mobility in proximity to this ARG, its genetic context. However, it is suggested that this species might have acquired this gene two reasons. First, this gene found in only 4 out of 19 *P. stutzeri* genomes found in NCBI database. Secondly, this gene is phylogenetically related to the previously characterized mobile subclass B1 metallo-β-lactamase families, including KHM, SIM, and IMP (418-420).

Pseudomonas saudiphocaensis has only recently been classified as a new species (421) and it has not been well characterized. There are only two studies where *P. saudiphocaensis* was recovered, first from air samples in the city environment of Makkah, Saudi Arabia in 2012 (421) and from a sheep dairy farm in New Zealand (422). To the best of our knowledge, this is first report of recovery of *P. saudiphocaensis* from bovine feces. With the hybrid genome assembly, we were able to construct a complete circular genome of this isolate. The genome of *P. saudiphocaensis* (3.6 Mbp) was small as compared to other *Pseudomonas* species which ranged from 5.5-6.7 Mbp. This isolates also carried *blaPST-2* gene, which has not been identified in other *P. saudiphocaensis* genomes (421, 422). The *blaPST-1* found in *P. stutzeri* and *P. saudiphocaensis* showed 95% amino acid similarity. Despite limited knowledge of *P. saudiphocaensis*, it is suggested that it may have acquired this gene.

In conclusion, the recovery of carbapenem resistant bacteria from beef production system in Alberta is rare. Majority of carbapenem resistant bacterial species including *P. aeruginosa, A. haemolyticus, A. junii, A. lwoffii* and *S. maltophilia* identified in this study carried intrinsic carbapenem resistant genes and were only recoverable following selective enrichment. *P. aeruginosa* found in this study were multidrug resistant and may reflect the fact that this species can readily acquire foreign genes via transformation through type IV pilus. Moreover, the phylogenetic analysis showed that bovine *P. aeruginosa* strains formed separate cluster from human clinical strains, indicating that they may have adapted to the cattle environment and prudent management in feedlot system has limited the spread of these isolates outside cattle production system.

Chapter 5: Conclusion and future perspective

5.1 Conclusion

Antimicrobial resistance is a cause for global concern due to its current and potential impact on global population health, the cost of health care, and a significant economic losses. It is estimated that by 2050 annual global GDP will fall by ≥1.1% due to AMR (423). In Canada, AMR added a $1.4 billion financial burden to the Canadian healthcare system in 2018. If the resistance rate increases to 40% in 2050, the cumulative cost of AMR to the healthcare system will reach 120 billion (424). AMR is a multifaceted problem as it encompasses human, animal, and environmental health. Despite restricted use of some antimicrobials including those that are of critical importance in human health, the antimicrobials used in humans and animals often overlap as the infections are caused by the same or similar pathogens. The prophylactic and metaphylactic administration of antimicrobials to livestock has been blamed for the emergence, propagation, dissemination, and persistence of AMR bacterial populations in the environment, potentially impacting both humans and livestock. Antimicrobial usage in livestock has come under increasing scrutiny in recent years following the emergence of AMR from antimicrobial usage in food-producing animals. AMR represents a global health problem, therefore it is important to conduct AMR surveillance studies to estimate the nature as well as the magnitude of these problems. It may also help in evaluating the effectiveness of efforts taken at the national and international levels to combat AMR in human and livestock ecosystems.

Commensal bacteria of the gastrointestinal tract of humans and livestock including enterococci and *Escherichia coli* are used as bioindicators in AMR surveillance programs.

These bacterial species are found universally in nature and can also be recovered from different environmental samples including soil, municipal wastewater, natural water sources, wildlife, marine, and freshwater (425). Although these bacterial species are often not the specific target of antimicrobials, the selective pressure imposed by antimicrobials can result in them gaining resistance and serving as a reservoir of AMR genes.

In recent years, the next-generation WGS has become an integral part of AMR surveillance programs. With the advancement of next-generation technologies, it is now possible to completely sequence bacterial genomes in a cost-effective and timely manner. The WGS generated from these platforms can be used to determine the genotypic resistance profiles of recovered isolates and the association of identified AMR with MGEs including plasmids, integrons, and integrated and conjugated elements. This genomic data along with phenotypic data can inform the phenotypic-genotypic association in investigated bacterial species. It is also possible to study the evolutionary relationship between isolates from different settings through phylogenetic analysis. In addition, the availability of WGS data in public databases allows us to perform independent reanalysis that contributes to understanding AMR patterns across different ecosystems.

The primary objective of this thesis was to contribute to the comprehension of the role played by indicator bacteria in the surveillance of antimicrobial resistance, as well as to gain a deeper understanding of the impact of antimicrobial usage in the human and livestock sectors on the development and spread of antimicrobial resistance. We did phenotypic and genotypic characterization of isolates recovered across a One-health continuum. The next-generation WGS data generated from short or/and long read sequencing platforms was used to determine the resistome, mobilome, and virulome. The

genomic relatedness among isolates recovered from different ecosystems, including; human clinical, livestock, natural water sources, and environment was also investigated.

Chapter 2: *Enterococcus hirae* is one member of the genus *Enterococci* that rarely causes infection in humans. We investigated the prevalence of *Enterococcus hirae* in a One-health continuum and characterized recovered isolates both phenotypically and genomically. *E. hirae* is an important *Enterococcus* species from a livestock perspective. It is important to investigate the AMR prevalence in this species to understand the role of this species in gene dissemination. From a human clinical perspective, although *E. hirae* is an opportunistic pathogen the rate of *E. hirae* infection is increasing. We performed comparative genomic analysis on *E. hirae* genomes recovered from different ecosystems and cross-species comparative genomic analysis among *E. hirae*, *E. faecium*, and *E. faecalis* to understand the genome structure of *E. hirae*. Some important findings of this work are highlighted below. We identified that *E. hirae* exists predominantly in the bovine feedlot environment whereas *E. faecium* and *E. faecalis* were more frequently isolated from human clinical settings. Among *E. hirae* isolates, phenotype and genotype resistance profiles reflect the type of antimicrobials used in that particular ecosystem. For instance, the majority of *E. hirae* isolates recovered from beef production systems were resistant to macrolides and tetracyclines which are commonly used antimicrobials in livestock for disease treatment and prevention (426). The genotype profiles also corroborated with the phenotypic profiles of *E. hirae* recovered from beef cattle feedlots, where macrolide and tetracycline antimicrobial resistance genes were prevalent. In the comparative genomic analysis of *E. hirae,* we identified that the pan-genome of this species is 'open' in nature, indicating that it can readily acquire genes from other bacterial species. The difference in

phenotype and genotype profiles in *E. hirae* recovered from humans and livestock suggested that the transfer of resistance genes across different ecosystems is likely limited. The cross-species comparative genomic analysis showed that *E. hirae* harbored genes that were beneficial for cattle's gut metabolism including those involved in vitamin B12 synthesis, whereas in *E. faecium* and *E. faecalis,* genes encoded for virulence factors and proteins involved in pathogenesis. Our study also identified that these three species are niche-specific as they have adapted to each specific niche, likely making the transfer of ARGs to human pathogens a rarity. From a One health perspective, it is important to consider *E. hirae* as an important enterococcal species in surveillance programs as this species can acquire ARGs that can be transferred to pathogenic bacterial species.

Chapter 3: *Enterococcus faecium* and *Enterococcus faecalis* are mainly opportunistic pathogens in humans and account for 14% of all nosocomial infections in the US in 2010. In this study, we performed a comparative genomic analysis of *E. faecium* and *E. faecalis* isolates recovered from the One-health sectors. This study aimed to provide an understanding of antimicrobial resistance in *E. faecium* and *E. faecalis* across livestock sectors and a One-health continuum. We included *E. faecium* and *E. faecalis* isolates recovered from livestock (swine, poultry, beef, and dairy cattle), human clinical isolates, environmental samples (natural water sources, river water, wild and domestic animals), and urban wastewater. Our study showed that tetracyclines and macrolide-resistant genotypes exist universally in *E. faecium* and *E. faecalis,* irrespective of differences in antimicrobial use across One-health sectors. These genes were often associated with mobile genetic elements along with other antimicrobial resistance genes conferring resistance to aminoglycosides, lincosamides, and streptogramins that may facilitate their

dissemination in ecosystems. The prevalence of MDR isolates in the livestock sector reflects the extent of antimicrobial use within livestock sectors in Canada. The beef production system had the least MDR isolates followed by poultry and swine. However, compared to poultry and swine isolates, tetracycline and macrolide-resistant isolates were found more in the beef production system, reflecting the extensive use of these antimicrobials in beef production (426). Vancomycin-resistant enterococci are a leading cause of nosocomial infections in humans (426). In our isolates from clinics and poultry, vancomycin resistance genes were identified. In the phylogenomic analysis of *E. faecium*, human clinical isolates were mapped to a hospital-associated clade. These findings showed that human clinical isolates likely evolved separately from livestock isolates.

Chapter 4: Carbapenems are reserved for the treatment of MDR infections in humans. Carbapenem-resistant bacterial species including *P. aeruginosa, Acinetobacter,* and *E. coli* are considered an urgent threat to human medicine (CDC, 2019). We investigated the prevalence of carbapenem-resistant bacteria in beef production systems after carbapenem enrichment procedures. We had an extremely low recovery of carbapenem-resistant bacteria from the beef production systems. Isolates that were recovered after enrichment were mostly intrinsically resistant to carbapenem. These include isolates that are recognized as human pathogens *Pseudomonas aeruginosa* and *Acinetobacter lwoffii* and isolates *that* are categorized as 'newly emerging pathogens' including *Acinetobacter haemolyticus, Acinetobacter junii,* and *Stenotrophomonas maltophilia.* MDR *P. aeruginosa* were also identified and carried genes against cephalosporin, chloramphenicol, aminoglycoside, and fosfomycin. MDR *P. aeruginosa* were categorized as a 'serious threat' to human health. To investigate the genomic relatedness among *P. aeruginosa* from

humans and cattle, we included genomes from public databases. *P. aeruginosa* from bovine sources formed a distinct cluster in the core-genome based phylogenomic tree. However, we identified that *P. aeruginosa* from human and bovine sources carried similar ARG profiles, indicating the widespread multidrug resistance in *P. aeruginosa* species. The acquired carbapenem resistance was found in two *Pseudomonas* species (*P. stutzeri* and *P. saudiphocaensis*). To our knowledge, this is the first report of the recovery of *P. saudiphocaensis* from bovine source, adding to the scarce existing literature (421, 422). Further comparative genomic analysis on these species was restricted by the limited number of genomes available in the public database (as of May 2023).

Overall, this thesis provides a snapshot of AMR patterns across a One-health continuum. As indicated by our analysis, the resistance to critically important antimicrobials is limited, indicating that AMR stewardship in the livestock industry has limited the spread of ARGs to the human ecosystem. Identification of ARGs carrying plasmids indicates the importance of these species in gene dissemination. This study contributed a massive amount of genomic data to public databases. Genomic databases are mostly dominated by human clinical isolates, with less representation from other sectors including livestock and the environment. For instance, the genome collection of *P. aeruginosa* consisted of human clinical isolates and a single bovine clinical isolate, with no representation from other livestock sectors. The submission of isolates from this study will fill these gaps in the genomic database. All genomic data generated in this thesis are publicly available through NCBI and can be utilized for further analysis. Overall, our study has added valuable information to the literature.

5.2 Future perspectives

The knowledge gained from this study can be applied in an extension of this study as well as in support of new studies. Below are some of the plausible future directions that could be explored:

AMR is a continuously evolving phenomenon; therefore, continuous monitoring of AMR is required. The information and genomic pipelines used in this study can be utilized in the future for similar surveillance programs to monitor AMR patterns in the One-health continuum.

Recent machine learning models have also shown promising outcomes in the investigations of AMR. In those models, the predictions of strain susceptibility or resistance are made based on genetic and AST data (427). The WGS-based AMR genotypic data and AST phenotypic data generated in this study can be utilized with the prediction models to further train the models in predicting resistant phenotypes and rank AMR genes with their importance.

MGEs play an important role in the dissemination of ARGs across bacterial species through HGT (428). Considering that a variety of MGEs like plasmids, and prophages were identified during this study, a future study could be designed to further investigate the role of these MGEs in the dissemination of ARGs across the One-health sectors identified in this study. In some isolates, ARGs were mapped on the MGEs. The MDR isolates carrying ARGs on MGEs could be studied further. Conjugative assays can be conducted to investigate the transferability of these identified MGEs. Furthermore, the knowledge gained from these experiments could be used to develop machine learning models to

determine the transfer potential of unknown MGEs across bacterial species. In artificial intelligence or machine learning, the prevailing knowledge base, such as genetic markers, is utilized to study complex relations or interactions. This involves using the information from trained models to identify similar or related patterns in unseen data. (427, 429, 430).

In this study, some bacterial species lack ARGs yet demonstrated a resistant phenotype. The majority of ARGs encode for enzymes that neutralize antimicrobial drugs through hydrolysis (hydrolases like beta-lactamases, cephalosporinases, carbapenemases, esterases, and epoxide hydrolases), transfer (transferases, such as acetyltransferases, phosphotransferases, nucleotidyltransferases, glycosyltransferases, ADP-ribosyltransferases, and S-transferases) and lysis (redox enzymes, including monooxygenases and lyases). A deep learning model can be developed using AST data and AMR genetic data generated from this study to decipher novel genes that have not yet been implicated in resistance to different families of antibiotics. For instance; in a study by Sunuwar et al. (2021) (431), AST phenotype data and AMR genetic data were used to predict the synergistic effect of AMR genes in resistance to different classes. A machine learning model was created with the ExtraTrees Classifier (ETC) algorithm and used to train binary matrices of genotypes (0 for absence and 1 for presence of an AMR gene) and relevant antibiotics' phenotypes (0 for susceptibility and 1 for resistance to an antibiotic) were created using the AMR genotypes and AST phenotypes data. In *Klebsiella*, phenotypic carbapenem resistance was associated with the presence of aminoglycosides [*aac(6)-Ib, aph(3)-Ia, aph(4)-Ia*] and sulfonamide (*sul1*) resistance genes were identified. This phenotype arose as a result of these genes encoding for proteins that had a similar function to carbapenemases (431).

From a One-health perspective, ESKAPE pathogens comprising *Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa,* and *Enterobacter* species are critical for human health and often exhibit MDR. In this study, we characterized and investigated MDR in *Enterococcus faecium* and *Pseudomonas aeruginosa* recovered from the One-health sectors. In the future, a similar study could be conducted including all ESKAPE pathogens to investigate the resistome, mobilome, and virulence genes in these pathogens from a One-health perspective.

The bacterial transcriptomic analysis includes the RNA-sequencing of expressed bacterial genes and has the potential to fill gaps in the AMR database. Transcriptomics links the genotype and phenotype, which is important in understanding AMR mechanisms. Bacteria exhibit diverse mechanisms of gene expression to maintain essential functions and modulate accessory functions in response to environmental signals. We have a limited understanding of the regulatory gene pathways involved in AMR. There are two major classes of regulatory RNA: RNA attenuators/riboswitches and small RNAs (sRNAs). RNA attenuators are sensory RNA and are part of the mRNA that they regulate. Transcriptomic analysis of *Enterococcus faecalis* harboring the tetracycline resistance determinant *tet(M),* on conjugative transposon Tn916 showed that the expression of *tet(M)* in the presence of tetracycline is regulated by an extension of a small transcript representing the upstream leader region in the resistance determinant (432). Regulatory sRNAs are 50–500 nucleotide long RNAs, transcribed independently or processed from messenger RNAs (mRNAs). They are involved in controlling complex physiological processes (433) and are also required for most pathogens to survive antimicrobial challenges (434). Similar to transcriptional factors, sRNAs can regulate antibiotic susceptibility through direct

interactions with mRNAs involved in drug import, efflux, and cell-wall synthesis. The regulatory sRNA contributes to intrinsic resistance and antimicrobial tolerance[15] in several bacterial species. Here are some examples: for *P. aeruginosa*, uptake of carbapenem is mediated by porin proteins encoded by *oprD* and *oprP*. A decrease in the expression of porin proteins are associated with the carbapenem resistant phenotype. Novel interactions between the 5′-untranslated region (5′-UTR) of *oprD* and two sRNAs (Sr0161 and ErsA) were identified. These sRNAs negatively regulate oprD expression, and overexpression of *Sr0161* increased resistance to meropenem in *P. aeruginosa* (435). In *E. faecium*, the sRNA_0160 was found to be associated with daptomycin resistance and responses to daptomycin (436). The role of sRNA in regulating antimicrobial resistance genes could be a future area of investigation.

[15] Antimicrobial tolerance refers to the ability of microorganisms to survive exposure to antimicrobial agents at concentrations that would normally be lethal to the majority of the population. Unlike AMR, antimicrobial tolerance is characterized by a temporary or reversible state in which microorganisms can persist in the presence of the drug without acquiring permanent resistance mechanisms.

References

1. **Gould IM, Bal AM.** 2013. New antibiotic agents in the pipeline and how they can help overcome microbial resistance. *Virulence* **4:**185-191.
2. **WHO.** https://www.who.int/news/item/29-04-2019-new-report-calls-for-urgent-action-to-avert-antimicrobial-resistance-crisis. https://www.who.int/news/item/29-04-2019-new-report-calls-for-urgent-action-to-avert-antimicrobial-resistance-crisis. Accessed
3. **Antimicrobial Resistance C.** 2022. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet (London, England)* **399:**629-655.
4. **Murray CJ, Ezzati M, Flaxman AD, Lim S, Lozano R, Michaud C, Naghavi M, Salomon JA, Shibuya K, Vos T, Wikler D, Lopez AD.** 2012. GBD 2010: design, definitions, and metrics. *Lancet* **380:**2063-2066.
5. **Aminov RI.** 2010. A brief history of the antibiotic era: lessons learned and challenges for the future. *Front Microbiol* **1:**134.
6. **Rice LB.** 2012. Mechanisms of resistance and clinical relevance of resistance to β-lactams, glycopeptides, and fluoroquinolones. *Mayo Clin Proc* **87:**198-208.
7. **Randall CP, Mariner KR, Chopra I, O'Neill AJ.** 2013. The target of daptomycin is absent from Escherichia coli and other gram-negative pathogens. *Antimicrob Agents Chemother* **57:**637-639.
8. **Blanco P, Hernando-Amado S, Reales-Calderon JA, Corona F, Lira F, Alcalde-Rico M, Bernardini A, Sanchez MB, Martinez JL.** 2016. Bacterial Multidrug Efflux Pumps: Much More Than Antibiotic Resistance Determinants. *Microorganisms* **4**.
9. **Cox G, Wright GD.** 2013. Intrinsic antibiotic resistance: Mechanisms, origins, challenges and solutions. *International Journal of Medical Microbiology* **303:**287-292.
10. **Capita R, Alonso-Calleja C.** 2013. Antibiotic-Resistant Bacteria: A Challenge for the Food Industry. *Critical Reviews in Food Science and Nutrition* **53:**11 - 48.
11. **Anonymous.** 2013. The antibiotic alarm. *Nature* **495:**141.
12. **Johnson AP.** 2015. Surveillance of antibiotic resistance. *Philosophical Transactions of the Royal Society B: Biological Sciences* **370:**20140080.
13. **Partridge SR, Kwong SM, Firth N, Jensen SO.** 2018. Mobile Genetic Elements Associated with Antimicrobial Resistance. *Clin Microbiol Rev* **31**.
14. **McInnes RS, McCallum GE, Lamberte LE, van Schaik W.** 2020. Horizontal transfer of antibiotic resistance genes in the human gut microbiome. *Curr Opin Microbiol* **53:**35-43.
15. **Pal C, Bengtsson-Palme J, Kristiansson E, Larsson DJ.** 2015. Co-occurrence of resistance genes to antibiotics, biocides and metals reveals novel insights into their co-selection potential. *BMC genomics* **16:**1-14.
16. **Spellberg B, Hansen GR, Kar A, Cordova CD, Price LB, Johnson JR.** 2016. Antibiotic resistance in humans and animals. *NAM Perspectives*.
17. **Curiao T, Marchi E, Grandgirard D, León-Sampedro R, Viti C, Leib SL, Baquero F, Oggioni MR, Martinez JL, Coque TM.** 2016. Multiple adaptive routes of Salmonella enterica Typhimurium to biocide and antibiotic exposure. *BMC genomics* **17:**1-16.
18. **Hall LM, Holtum JA, Powles SB.** 2018. Mechanisms responsible for cross resistance and multiple resistance, p 243-262, *Herbicide Resistance in Plants*. CRC Press.
19. **Vogwill T, MacLean RC.** 2015. The genetic basis of the fitness costs of antimicrobial resistance: a meta-analysis approach. *Evolutionary Applications* **8:**284-295.
20. **Koraimann G, Wagner MA.** 2014. Social behavior and decision making in bacterial conjugation. *Frontiers in Cellular and Infection Microbiology* **4**.
21. **Canchaya C, Fournous G, Brüssow H.** 2004. The impact of prophages on bacterial chromosomes. *Molecular Microbiology* **53:**9-18.
22. **Bragg JG, Wagner A.** 2009. Protein material costs: single atoms can make an evolutionary difference. *Trends in Genetics* **25:**5-8.
23. **WHO.** 14 October 2022 2022. One health joint plan of action (2022–2026): working together for the health of humans, animals, plants and the environment. https://www.who.int/publications/i/item/9789240059139. Accessed 06/02.

24. **Conrad C, Stanford K, McAllister T, Thomas J, Reuter T.** 2016. Shiga toxin-producing Escherichia coli and current trends in diagnostics. *Animal Frontiers* **6:**37-43.
25. **Harbarth S, Balkhy HH, Goossens H, Jarlier V, Kluytmans J, Laxminarayan R, Saam M, Van Belkum A, Pittet D.** 2015. *Antimicrobial resistance: one world, one fight!* Springer.
26. **Morrill HJ, Caffrey AR, Jump RL, Dosa D, LaPlante KL.** 2016. Antimicrobial stewardship in long-term care facilities: a call to action. *Journal of the American Medical Directors Association* **17:**183. e181-183. e116.
27. **Anonymous.** October 4, 2022 2022. Outpatient Antibiotic Prescriptions — United States, 2021, *on Centers for Disease Control and Prevention, National Center for Emerging and Zoonotic Infectious Diseases (NCEZID), Division of Healthcare Quality Promotion (DHQP)*. https://www.cdc.gov/antibiotic-use/data/report-2021.html. Accessed October 12.
28. **Michael CA, Dominey-Howes D, Labbate M.** 2014. The antimicrobial resistance crisis: causes, consequences, and management. *Front Public Health* **2:**145.
29. **Bartlett JG, Gilbert DN, Spellberg B.** 2013. Seven ways to preserve the miracle of antibiotics. *Clin Infect Dis* **56:**1445-1450.
30. **Nickell JS, White BJ.** 2010. Metaphylactic Antimicrobial Therapy for Bovine Respiratory Disease in Stocker and Feedlot Cattle. *Veterinary Clinics of North America: Food Animal Practice* **26:**285-301.
31. **Van Boeckel TP, Brower C, Gilbert M, Grenfell BT, Levin SA, Robinson TP, Teillant A, Laxminarayan R.** 2015. Global trends in antimicrobial use in food animals. *Proceedings of the National Academy of Sciences* **112:**5649-5654.
32. **Bruno A, Mackay C.** 2012. Antimicrobial resistance and the activities of the Codex Alimentarius Commission. *Revue Scientifique et Technique (International Office of Epizootics)* **31:**317-323.
33. **Berendsen BJ, Wegh RS, Memelink J, Zuidema T, Stolker LA.** 2015. The analysis of animal faeces as a tool to monitor antibiotic usage. *Talanta* **132:**258-268.
34. **Salama N, Abou-Raya S, Shalaby A, Emam W, Mehaya F.** 2011. Incidence of tetracycline residues in chicken meat and liver retailed to consumers. *Food Additives and Contaminants: Part B* **4:**88-93.
35. **Maurya GK, Pachauri S.** 2022. Chapter 16 - Fungi: The indicators of pollution, p 277-296. *In* Bandh SA, Shafi S (ed), *Freshwater Mycology* doi:https://doi.org/10.1016/B978-0-323-91232-7.00012-X. Elsevier.
36. **Singh B, Christina E.** 2021. Chapter 11 - Bacterial metabolites: an unexplored quarry, p 205-234. *In* Kumar A, Singh J, Samuel J (ed), *Volatiles and Metabolites of Microbes* doi:https://doi.org/10.1016/B978-0-12-824523-1.00006-7. Academic Press.
37. **Arunima A, Das JK, Suar M.** 2019. Chapter 10 - Gut Microbes in Liver Diseases: Dietary Intervention for Promoting Hepatic Health, p 117-131. *In* Watson RR, Preedy VR (ed), *Dietary Interventions in Gastrointestinal Diseases* doi:https://doi.org/10.1016/B978-0-12-814468-8.00010-7. Academic Press.
38. **Barton MD.** 2000. Antibiotic use in animal feed and its impact on human healt. *Nutr Res Rev* **13:**279-299.
39. **Authority EFS.** 2008. Report from the Task Force on Zoonoses Data Collection including guidance for harmonized monitoring and reporting of antimicrobial resistance in commensal Escherichia coli and Enterococcus spp. from food animals. *EFSA journal* **6:**141r.
40. **Aspevall O, Dohnhammar U, Nilsson O, Pringle MJPHAoS, Institute NV.** 2019. Swedres-Svarm 2018. Consumption of Antibiotics and Occurrence of Antibiotic Resistance in Sweden.
41. **Chazel M, Jouy E, Meunier D, Haenni M, Gay E, Calavas D, Madec J-YJBé.** 2009. Résapath: réseau d'épidémiosurveillance de l'antibiorésistance des bactéries pathogènes animales. **34:**7-8.
42. **Directorate VMJNH, Addlestone: Veterinary Medicines Directorate.** 2019. UK one health report. Joint report on antibiotic use and antibiotic resistance, 2013–2017.
43. **De Greeff S, Schoffelen A, Verduin C.** 2020. NethMap 2020: Consumption of antimicrobial agents and antimicrobial resistance among medically important bacteria in the Netherlandsin 2019/MARAN 2020: Monitoring of Antimicrobial Resistance and Antibiotic Usage in Animals in the Netherlands in 2019.
44. **Korsgaard HB, Ellis-Iversen J, Hendriksen RS, Høg BB, Ronco T, Attauabi M, Boel J, Dalby T, Hammerum AM, Hansen F.** 2020. DANMAP 2019-Use of antimicrobial agents and occurrence of antimicrobial resistance in bacteria from food animals, food and humans in Denmark.

45. **Agunos A, Gow SP, Deckert AE, Kuiper G, Léger DFJP.** 2021. Informing Stewardship Measures in Canadian Food Animal Species through Integrated Reporting of Antimicrobial Use and Antimicrobial Resistance Surveillance Data—Part I, Methodology Development. **10:**1492.
46. **Leimbach A, Hacker J, Dobrindt U.** 2013. E. coli as an All-Rounder: The Thin Line Between Commensalism and Pathogenicity, p 3-32. *In* Dobrindt U, Hacker JH, Svanborg C (ed), *Between Pathogenicity and Commensalism* doi:10.1007/82_2012_303. Springer Berlin Heidelberg, Berlin, Heidelberg.
47. **Martinson JNV, Walk ST.** 2020. Escherichia coli Residency in the Gut of Healthy Human Adults. *EcoSal Plus* **9**.
48. **Jang J, Hur HG, Sadowsky MJ, Byappanahalli MN, Yan T, Ishii S.** 2017. Environmental Escherichia coli: ecology and public health implications-a review. *J Appl Microbiol* **123:**570-581.
49. **Ramos S, Silva V, Dapkevicius MdLE, Caniça M, Tejedor-Junco MT, Igrejas G, Poeta P.** 2020. Escherichia coli as Commensal and Pathogenic Bacteria Among Food-Producing Animals: Health Implications of Extended Spectrum β-lactamase (ESBL) Production. *Animals : an open access journal from MDPI* **10:**2239.
50. **Tivendale KA, Logue CM, Kariyawasam S, Jordan D, Hussein A, Li G, Wannemuehler Y, Nolan LK.** 2010. Avian-pathogenic Escherichia coli strains are similar to neonatal meningitis E. coli strains and are able to cause meningitis in the rat model of human disease. *Infect Immun* **78:**3412-3419.
51. **Lai J, Zhu Y, Tang L, Lin X.** 2021. Epidemiology and antimicrobial susceptibility of invasive Escherichia coli infection in neonates from 2012 to 2019 in Xiamen, China. *BMC Infect Dis* **21:**295.
52. **La Combe B, Clermont O, Messika J, Eveillard M, Kouatchet A, Lasocki S, Corvec S, Lakhal K, Billard-Pomares T, Fernandes R, Armand-Lefevre L, Bourdon S, Reignier J, Fihman V, de Prost N, Bador J, Goret J, Wallet F, Denamur E, Ricard JD.** 2019. Pneumonia-Specific Escherichia coli with Distinct Phylogenetic and Virulence Profiles, France, 2012-2014. *Emerg Infect Dis* **25:**710-718.
53. **Sun M, Gao X, Zhao K, Ma J, Yao H, Pan Z.** 2021. Insight Into the Virulence Related Secretion Systems, Fimbriae, and Toxins in O2:K1 Escherichia coli Isolated From Bovine Mastitis. **8**.
54. **Tantoso E, Eisenhaber B, Kirsch M, Shitov V, Zhao Z, Eisenhaber F.** 2022. To kill or to be killed: pangenome analysis of Escherichia coli strains reveals a tailocin specific for pandemic ST131. *BMC Biol* **20:**146.
55. **van Elsas JD, Semenov AV, Costa R, Trevors JT.** 2011. Survival of Escherichia coli in the environment: fundamental and public health aspects. *Isme j* **5:**173-183.
56. **Tetz VV.** 2005. The pangenome concept: a unifying view of genetic information. *Medical science monitor: international medical journal of experimental and clinical research* **11:**HY24-29.
57. **Tonkin-Hill G, Gladstone RA, Pöntinen AK, Arredondo-Alonso S, Bentley SD, Corander J.** 2023. Robust analysis of prokaryotic pangenome gene gain and loss rates with Panstripe. *Genome Research*.
58. **Leimbach A, Hacker J, Dobrindt U.** 2013. E. coli as an all-rounder: the thin line between commensalism and pathogenicity. *Curr Top Microbiol Immunol* **358:**3-32.
59. **Rasko DA, Rosovitz MJ, Myers GS, Mongodin EF, Fricke WF, Gajer P, Crabtree J, Sebaihia M, Thomson NR, Chaudhuri R, Henderson IR, Sperandio V, Ravel J.** 2008. The pangenome structure of Escherichia coli: comparative genomic analysis of E. coli commensal and pathogenic isolates. *J Bacteriol* **190:**6881-6893.
60. **Lukjancenko O, Wassenaar TM, Ussery DW.** 2010. Comparison of 61 sequenced Escherichia coli genomes. *Microbial ecology* **60:**708-720.
61. **Kaas RS, Friis C, Ussery DW, Aarestrup FM.** 2012. Estimating variation within the genes and inferring the phylogeny of 186 sequenced diverse Escherichia coli genomes. *BMC genomics* **13:**577-577.
62. **Ding W, Baumdicker F, Neher RAJNar.** 2018. panX: pan-genome analysis and exploration. **46:**e5-e5.
63. **Beghain J, Bridier-Nahmias A, Le Nagard H, Denamur E, Clermont O.** 2018. ClermonTyping: an easy-to-use and accurate in silico method for Escherichia genus strain phylotyping. *Microb Genom* **4**.

64. **Clermont O, Christenson JK, Denamur E, Gordon DM.** 2013. The Clermont Escherichia coli phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. *Environ Microbiol Rep* **5:**58-65.
65. **Tenaillon O, Skurnik D, Picard B, Denamur E.** 2010. The population genetics of commensal Escherichia coli. *Nat Rev Microbiol* **8:**207-217.
66. **Tlaskalová-Hogenová H, Stepánková R, Hudcovic T, Tucková L, Cukrowska B, Lodinová-Zádníková R, Kozáková H, Rossmann P, Bártová J, Sokol D, Funda DP, Borovská D, Reháková Z, Sinkora J, Hofman J, Drastich P, Kokesová A.** 2004. Commensal bacteria (normal microflora), mucosal immunity and chronic inflammatory and autoimmune diseases. *Immunol Lett* **93:**97-108.
67. **Chao L, Levin BR.** 1981. Structured habitats and the evolution of anticompetitor toxins in bacteria. *Proc Natl Acad Sci U S A* **78:**6324-6328.
68. **Wirth T, Falush D, Lan R, Colles F, Mensa P, Wieler LH, Karch H, Reeves PR, Maiden MC, Ochman H, Achtman M.** 2006. Sex and virulence in Escherichia coli: an evolutionary perspective. *Mol Microbiol* **60:**1136-1151.
69. **Nowrouzian F, Adlerberth I, Wold AE.** 2001. P fimbriae, capsule and aerobactin characterize colonic resident Escherichia coli. *Epidemiol Infect* **126:**11-18.
70. **Wildschutte H, Wolfe DM, Tamewitz A, Lawrence JG.** 2004. Protozoan predation, diversifying selection, and the evolution of antigenic diversity in Salmonella. *Proc Natl Acad Sci U S A* **101:**10644-10649.
71. **Escobar-Páramo P, Le Menac'h A, Le Gall T, Amorin C, Gouriou S, Picard B, Skurnik D, Denamur E.** 2006. Identification of forces shaping the commensal Escherichia coli genetic structure by comparing animal and human isolates. *Environ Microbiol* **8:**1975-1984.
72. **Teuber MJC, CMLS MLS.** 1999. Spread of antibiotic resistance with food-borne pathogens. **56:**755-763.
73. **Datta NJBmJ.** 1969. Drug resistance and R factors in the bowel bacteria of London patients before and after admission to hospital. **2:**407-411.
74. **Guinee P, Ugueto N, Van Leeuwen NJAM.** 1970. Escherichia coli with resistance factors in vegetarians, babies, and nonvegetarians. **20:**531-535.
75. **Linton K, Lee PA, Richmond M, Gillespie W, Rowland A, Baker VNJE, Infection.** 1972. Antibiotic resistance and transmissible R-factors in the intestinal coliform flora of healthy adults and children in an urban and a rural community. **70:**99-104.
76. **Nijsten R, London N, Van den Bogaard A, Stobberingh EJJoAC.** 1996. In-vitro transfer of antibiotic resistance between faecal Escherichia coli strains isolated from pig farmers and pigs. **37:**1141-1154.
77. **Dunn SJ, Connor C, McNally A.** 2019. The evolution and transmission of multi-drug resistant Escherichia coli and Klebsiella pneumoniae: the complexity of clones and plasmids. *Current Opinion in Microbiology* **51:**51-56.
78. **Valentin L, Sharp H, Hille K, Seibt U, Fischer J, Pfeifer Y, Michael GB, Nickel S, Schmiedel J, Falgenhauer L, Friese A, Bauerfeind R, Roesler U, Imirzalioglu C, Chakraborty T, Helmuth R, Valenza G, Werner G, Schwarz S, Guerra B, Appel B, Kreienbrock L, Käsbohrer A.** 2014. Subgrouping of ESBL-producing Escherichia coli from animal and human sources: an approach to quantify the distribution of ESBL types between different reservoirs. *Int J Med Microbiol* **304:**805-816.
79. **World Health O, Resistance WHOAGoISoA.** 2017. *Critically important antimicrobials for human medicine: ranking of antimicrobial agents for risk management of antimicrobial resistance due to non-human use*, 5th rev. ed. World Health Organization, Geneva.
80. **Organization WH.** 2017. *Global priority list of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics. Geneva: WHO; 2017*.
81. **Wilke MS, Lovering AL, Strynadka NCJ.** 2005. β-Lactam antibiotic resistance: a current structural perspective. *Current Opinion in Microbiology* **8:**525-533.
82. **Logan LK, Weinstein RA.** 2017. The Epidemiology of Carbapenem-Resistant Enterobacteriaceae: The Impact and Evolution of a Global Menace. *The Journal of infectious diseases* **215:**S28-S36.
83. **Patel G, Bonomo R.** 2013. "Stormy waters ahead": global emergence of carbapenemases. *Frontiers in Microbiology* **4**.

84. **Walsh TR.** 2010. Emerging carbapenemases: a global perspective. *Int J Antimicrob Agents* **36 Suppl 3:**S8-14.
85. **Gupta V.** 2007. An update on newer β-lactamases. *Indian Journal of Medical Research* **126:**417-427.
86. **Dhillon RH-P, Clark J.** 2012. ESBLs: a clear and present danger? *Critical care research and practice* **2012**.
87. **Silveira MC, Azevedo da Silva R, Faria da Mota F, Catanho M, Jardim R, R Guimarães AC, de Miranda AB.** 2018. Systematic Identification and Classification of β-Lactamases Based on Sequence Similarity Criteria: β-Lactamase Annotation. *Evolutionary bioinformatics online* **14:**1176934318797351-1176934318797351.
88. **Bebrone C.** 2007. Metallo-beta-lactamases (classification, activity, genetic organization, structure, zinc coordination) and their superfamily. *Biochem Pharmacol* **74:**1686-1701.
89. **Smith HZ, Kendall B.** 2019. Carbapenem Resistant Enterobacteriacea (CRE), *StatPearls*. StatPearls Publishing

StatPearls Publishing LLC., Treasure Island (FL).
90. **Walther-Rasmussen J, Høiby N.** 2007. Class A carbapenemases. *Journal of Antimicrobial Chemotherapy* **60:**470-482.
91. **Nordmann P, Cuzon G, Naas T.** 2009. The real threat of Klebsiella pneumoniae carbapenemase-producing bacteria. *The Lancet infectious diseases* **9:**228-236.
92. **Chen LF, Anderson DJ, Paterson DL.** 2012. Overview of the epidemiology and the threat of Klebsiella pneumoniae carbapenemases (KPC) resistance. *Infection and drug resistance* **5:**133.
93. **Del Franco M, Paone L, Novati R, Giacomazzi CG, Bagattini M, Galotto C, Montanera PG, Triassi M, Zarrilli R.** 2015. Molecular epidemiology of carbapenem resistant Enterobacteriaceae in Valle d'Aosta region, Italy, shows the emergence of KPC-2 producing Klebsiella pneumoniae clonal complex 101 (ST101 and ST1789). *BMC microbiology* **15:**260.
94. **Guerreiro LF.** 2018. *Cefalosporinas e carbapenemos no tratamento de infeções provocadas por bactérias Gram-negativas*.
95. **Nordmann P, Poirel L.** 2002. Emerging carbapenemases in Gram-negative aerobes. *Clinical Microbiology and Infection* **8:**321-331.
96. **Meletis G.** 2016. Carbapenem resistance: overview of the problem and future perspectives. *Therapeutic advances in infectious disease* **3:**15-21.
97. **Bedenić B, Sardelić S.** 2018. Carbapenemases, *Growing and Handling of Bacterial Cultures*. IntechOpen.
98. **Hoelle J, Johnson JR, Johnston BD, Kinkle B, Boczek L, Ryu H, Hayes S.** 2019. Survey of US wastewater for carbapenem-resistant Enterobacteriaceae. *J Water Health* **17:**219-226.
99. **Byappanahalli MN, Nevers MB, Korajkic A, Staley ZR, Harwood VJ.** 2012. Enterococci in the environment. *Microbiol Mol Biol Rev* **76:**685-706.
100. **Anonymous.** 2014. *In* Gilmore MS, Clewell DB, Ike Y, Shankar N (ed), *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*. Massachusetts Eye and Ear Infirmary, Boston.
101. **Lebreton F, Willems RJL, Gilmore MS.** 2014. Enterococcus Diversity, Origins in Nature, and Gut Colonization. *In* Gilmore MS, Clewell DB, Ike Y, Shankar N (ed), *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*. Massachusetts Eye and Ear Infirmary, Boston.
102. **Hsueh PR, Teng LJ, Chen YC, Yang PC, Ho SW, Luh KT.** 2000. Recurrent bacteremic peritonitis caused by Enterococcus cecorum in a patient with liver cirrhosis. *J Clin Microbiol* **38:**2450-2452.
103. **Yan T, Goto DK, Feng F.** 2011. *Concentration Dynamics of Fecal Indicators in Hawaiian Coastal and Inland Sand, Soil, and Water during Rainfall Events*. Water Environment Research Foundation Alexandria, VA.
104. **Byappanahalli M, Fujioka R.** 2004. Indigenous soil bacteria and low moisture may limit but allow faecal bacteria to multiply and become a minor population in tropical soils. *Water Science and Technology* **50:**27-32.

105. **Harwood VJ, Whitlock J, Withington V.** 2000. Classification of antibiotic resistance patterns of indicator bacteria by discriminant analysis: use in predicting the source of fecal contamination in subtropical waters. *Applied and Environmental Microbiology* **66:**3698-3704.
106. **Layton B, Walters S, Lam L, Boehm A.** 2010. Enterococcus species distribution among human and animal hosts using multiplex PCR. *Journal of applied microbiology* **109:**539-547.
107. **Fiore E, Van Tyne D, Gilmore MS.** 2019. Pathogenicity of Enterococci. *Microbiology spectrum* **7:**10.1128/microbiolspec.GPP1123-0053-2018.
108. **Facklam Rr, Carvalho MdG, Teixeira LM.** History, Taxonomy, Biochemical Characteristics, and Antibiotic Susceptibility Testing of Enterococci, p. *In* (ed),
109. **Carrero-Colón M, Wickham GS, Turco RF.** 2010. Taxonomy, Phylogeny, and Physiology of Fecal Indicator Bacteria, p 23-38, *The Fecal Bacteria* doi:https://doi.org/10.1128/9781555816865.ch2.
110. **van Schaik W, Top J, Riley DR, Boekhorst J, Vrijenhoek JEP, Schapendonk CME, Hendrickx APA, Nijman IJ, Bonten MJM, Tettelin H, Willems RJL.** 2010. Pyrosequencing-based comparative genome analysis of the nosocomial pathogen Enterococcus faecium and identification of a large transferable pathogenicity island. *BMC genomics* **11:**239-239.
111. **Miller WR, Munita JM, Arias CA.** 2014. Mechanisms of antibiotic resistance in enterococci. *Expert review of anti-infective therapy* **12:**1221-1236.
112. **Herrera-Hidalgo L, de Alarcón A, López-Cortes LE, Luque-Márquez R, López-Cortes LF, Gutiérrez-Valencia A, Gil-Navarro MV.** 2020. Enterococcus faecalis Endocarditis and Outpatient Treatment: A Systematic Review of Current Alternatives. *Antibiotics (Basel)* **9**.
113. **Mikalsen T, Pedersen T, Willems R, Coque TM, Werner G, Sadowy E, van Schaik W, Jensen LB, Sundsfjord A, Hegstad K.** 2015. Investigating the mobilome in clinically important lineages of Enterococcus faecium and Enterococcus faecalis. *BMC Genomics* **16:**282.
114. **Butler MS, Hansford KA, Blaskovich MAT, Halai R, Cooper MA.** 2014. Glycopeptide antibiotics: Back to the future. *The Journal of Antibiotics* **67:**631-644.
115. **Bonten MJ, Willems R, Weinstein RAJTLid.** 2001. Vancomycin-resistant enterococci: why are they here, and where do they come from? **1:**314-325.
116. **Kobayashi N, Mahbub Alam MD, Nishimoto Y, Urasawa S, Uehara N, Watanabe N.** 2001. Distribution of aminoglycoside resistance genes in recent clinical isolates of Enterococcus faecalis, Enterococcus faecium and Enterococcus avium. *Epidemiology and Infection* **126:**197-204.
117. **Silby MW, Winstanley C, Godfrey SA, Levy SB, Jackson RW.** 2011. Pseudomonas genomes: diverse and adaptable. *FEMS microbiology reviews* **35:**652-680.
118. **Silby MW, Winstanley C, Godfrey SAC, Levy SB, Jackson RW.** 2011. Pseudomonas genomes: diverse and adaptable. *FEMS Microbiology Reviews* **35:**652-680.
119. **Jeukens J, Freschi L, Kukavica-Ibrulj I, Emond-Rheault JG, Tucker NP, Levesque RC.** 2019. Genomics of antibiotic-resistance prediction in Pseudomonas aeruginosa. *Ann N Y Acad Sci* **1435:**5-17.
120. **Mowat E, Paterson S, Fothergill JL, Wright EA, Ledson MJ, Walshaw MJ, Brockhurst MA, Winstanley C.** 2011. Pseudomonas aeruginosa population diversity and turnover in cystic fibrosis chronic infections. *Am J Respir Crit Care Med* **183:**1674-1679.
121. **OneHealthTrust. ResistanceMap Ar.** 2023. https://resistancemap.onehealthtrust.org/AntibioticResistance.phpOneHealthTrust. . Accessed
122. **Treepong P, Kos VN, Guyeux C, Blanc DS, Bertrand X, Valot B, Hocquet D.** 2018. Global emergence of the widespread Pseudomonas aeruginosa ST235 clone. *Clin Microbiol Infect* **24:**258-266.
123. **Cabot G, López-Causapé C, Ocampo-Sosa AA, Sommer LM, Domínguez M, Zamorano L, Juan C, Tubau F, Rodríguez C, Moyà B, Peña C, Martínez-Martínez L, Plesiat P, Oliver A.** 2016. Deciphering the Resistome of the Widespread Pseudomonas aeruginosa Sequence Type 175 International High-Risk Clone through Whole-Genome Sequencing. *Antimicrob Agents Chemother* **60:**7415-7423.
124. **Horcajada JP, Montero M, Oliver A, Sorlí L, Luque S, Gómez-Zorrilla S, Benito N, Grau S.** 2019. Epidemiology and Treatment of Multidrug-Resistant and Extensively Drug-Resistant Pseudomonas aeruginosa Infections. *Clin Microbiol Rev* **32**.
125. **van Belkum A, Tassios PT, Dijkshoorn L, Haeggman S, Cookson B, Fry NK, Fussing V, Green J, Feil E, Gerner-Smidt P, Brisse S, Struelens M.** 2007. Guidelines for the validation and

application of typing methods for use in bacterial epidemiology. *Clin Microbiol Infect* **13 Suppl 3:**1-46.

126. **Guo X, Xia R, Han N, Xu H.** 2011. Genetic diversity analyses of class 1 integrons and their associated antimicrobial resistance genes in Enterobacteriaceae strains recovered from aquatic habitats in China. *Lett Appl Microbiol* **52:**667-675.
127. **Bouzouita I, Draoui H, Mahdhi S, Essalah L, Saidi LS.** 2021. Evaluation of PCR pncA-restriction fragment length polymorphism and PCR amplification of genomic regions of difference for the identification of M. bovis strains in lymph nodes cultures. *African Health Sciences* **21:**985-989.
128. **Charretier Y, Schrenzel J.** 2016. Mass spectrometry methods for predicting antibiotic resistance. *PROTEOMICS – Clinical Applications* **10:**964-981.
129. **Parizad EG, Parizad EG, Valizadeh A.** 2016. The Application of Pulsed Field Gel Electrophoresis in Clinical Studies. *J Clin Diagn Res* **10:**De01-04.
130. **Quainoo S, Coolen JPM, van Hijum S, Huynen MA, Melchers WJG, van Schaik W, Wertheim HFL.** 2017. Whole-Genome Sequencing of Bacterial Pathogens: the Future of Nosocomial Outbreak Analysis. *Clin Microbiol Rev* **30:**1015-1063.
131. **Sanger F.** 1988. Sequences, sequences, and sequences. *Annu Rev Biochem* **57:**1-28.
132. **Organization WH.** 2020. GLASS whole-genome sequencing for surveillance of antimicrobial resistance.
133. **Stubberfield E, AbuOun M, Sayers E, O'Connor HM, Card RM, Anjum MF.** 2019. Use of whole genome sequencing of commensal Escherichia coli in pigs for antimicrobial resistance surveillance, United Kingdom, 2018. *Eurosurveillance* **24:**1900136.
134. **AMR NGHRUoGSo.** 2020. Whole-genome sequencing as part of national and international surveillance programmes for antimicrobial resistance: a roadmap. *BMJ global health* **5:**e002244.
135. **Goodwin S, McPherson JD, McCombie WR.** 2016. Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet* **17:**333-351.
136. **Kircher M, Kelso J.** 2010. High-throughput DNA sequencing--concepts and limitations. *Bioessays* **32:**524-536.
137. **Mardis ER.** 2008. Next-generation DNA sequencing methods. *Annu Rev Genomics Hum Genet* **9:**387-402.
138. **Green RE, Krause J, Ptak SE, Briggs AW, Ronan MT, Simons JF, Du L, Egholm M, Rothberg JM, Paunovic M.** 2006. Analysis of one million base pairs of Neanderthal DNA. *Nature* **444:**330-336.
139. **Valouev A, Ichikawa J, Tonthat T, Stuart J, Ranade S, Peckham H, Zeng K, Malek JA, Costa G, McKernan K.** 2008. A high-resolution, nucleosome position map of C. elegans reveals a lack of universal sequence-dictated positioning. *Genome research* **18:**1051-1063.
140. **Liu L, Li Y, Li S, Hu N, He Y, Pong R, Lin D, Lu L, Law M.** 2012. Comparison of next-generation sequencing systems. *Journal of Biomedicine and biotechnology* **2012**.
141. **Ravi RK, Walton K, Khosroheidari M.** 2018. MiSeq: A Next Generation Sequencing Platform for Genomic Analysis, p 223-232. *In* DiStefano JK (ed), *Disease Gene Identification: Methods and Protocols* doi:10.1007/978-1-4939-7471-9_12. Springer New York, New York, NY.
142. **Schadt EE, Turner S, Kasarskis A.** 2010. A window into third-generation sequencing. *Human Molecular Genetics* **19:**R227-R240.
143. **Langmead B, Trapnell C, Pop M, Salzberg S.** 2009. *Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol10: R25*.
144. **Li H, Durbin R.** 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25:**1754-1760.
145. **Lin HN, Hsu WL.** 2017. Kart: a divide-and-conquer algorithm for NGS read alignment. *Bioinformatics* **33:**2281-2287.
146. **Huang L, Popic V, Batzoglou S.** 2013. Short read alignment with populations of genomes. *Bioinformatics* **29:**i361-i370.
147. **Edwards DJ, Holt KE.** 2013. Beginner's guide to comparative bacterial genome analysis using next-generation sequence data. *Microbial Informatics and Experimentation* **3:**2.
148. **Hernandez D, François P, Farinelli L, Osterås M, Schrenzel J.** 2008. De novo bacterial genome sequencing: millions of very short reads assembled on a desktop computer. *Genome Res* **18:**802-809.

149. **Simpson JT, Durbin R.** 2012. Efficient de novo assembly of large genomes using compressed data structures. *Genome Res* **22:**549-556.
150. **Li H.** 2012. Exploring single-sample SNP and INDEL calling with whole-genome de novo assembly. *Bioinformatics* **28:**1838-1844.
151. **Nagarajan N, Pop M.** 2013. Sequence assembly demystified. *Nat Rev Genet* **14:**157-167.
152. **Chu TC, Lu CH, Liu T, Lee GC, Li WH, Shih AC.** 2013. Assembler for de novo assembly of large genomes. *Proc Natl Acad Sci U S A* **110:**E3417-3424.
153. **Medvedev P, Pop M.** 2021. What do Eulerian and Hamiltonian cycles have to do with genome assembly? *PLoS computational biology* **17:**e1008928-e1008928.
154. **Zerbino DR, Birney E.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* **18:**821-829.
155. **Anonymous.** 2012. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology* **19:**455-477.
156. **Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, Tang J, Wu G, Zhang H, Shi Y, Liu Y, Yu C, Wang B, Lu Y, Han C, Cheung DW, Yiu SM, Peng S, Xiaoqian Z, Liu G, Liao X, Li Y, Yang H, Wang J, Lam TW, Wang J.** 2015. Erratum: SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience* **4:**30.
157. **Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJM, Birol I.** 2009. ABySS: a parallel assembler for short read sequence data. *Genome research* **19:**1117-1123.
158. **Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A.** 2020. Using SPAdes de novo assembler. *Current protocols in bioinformatics* **70:**e102.
159. **Benjamini Y, Speed TP.** 2012. Summarizing and correcting the GC content bias in high-throughput sequencing. *Nucleic Acids Res* **40:**e72.
160. **Lan JH, Yin Y, Reed EF, Moua K, Thomas K, Zhang Q.** 2015. Impact of three Illumina library construction methods on GC bias and HLA genotype calling. *Hum Immunol* **76:**166-175.
161. **Kumar KR, Cowley MJ, Davis RL.** 2019. Next-Generation Sequencing and Emerging Technologies. *Semin Thromb Hemost* **45:**661-673.
162. **Kono N, Arakawa K.** 2019. Nanopore sequencing: Review of potential applications in functional genomics. *Development, Growth & Differentiation* **61:**316-326.
163. **David M, Dursi LJ, Yao D, Boutros PC, Simpson JT.** 2016. Nanocall: an open source basecaller for Oxford Nanopore sequencing data. *Bioinformatics* **33:**49-55.
164. **Boža V, Brejová B, Vinař T.** 2017. DeepNano: Deep recurrent neural networks for base calling in MinION nanopore reads. *PLOS ONE* **12:**e0178751.
165. **Stoiber M, Brown JJB.** 2017. BasecRAWller: Streaming nanopore basecalling directly from raw signal.133058.
166. **Sovic I, Sikic M, Wilm A, Fenlon SN, Chen S, Nagarajan N.** 2015. Fast and sensitive mapping of error-prone nanopore sequencing reads with GraphMap. doi:10.1101/020719 %J bioRxiv**:**020719.
167. **Li H.** 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34:**3094-3100.
168. **Nurk S, Walenz BP, Rhie A, Vollger MR, Logsdon GA, Grothe R, Miga KH, Eichler EE, Phillippy AM, Koren S.** 2020. HiCanu: accurate assembly of segmental duplications, satellites, and allelic variants from high-fidelity long reads. doi:10.1101/2020.03.14.992248 %J bioRxiv**:**2020.2003.2014.992248.
169. **Kolmogorov M, Yuan J, Lin Y, Pevzner PA.** 2019. Assembly of long, error-prone reads using repeat graphs. *Nature Biotechnology* **37:**540-546.
170. **Li H.** 2016. Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. *Bioinformatics* **32:**2103-2110.
171. **Liu H, Wu S, Li A, Ruan JJG.** 2020. SMARTdenovo: a de novo assembler using long noisy reads.
172. **Vaser R, Šikić M.** 2021. Time- and memory-efficient genome assembly with Raven. *Nature Computational Science* **1:**332-336.
173. **van Dijk EL, Jaszczyszyn Y, Naquin D, Thermes C.** 2018. The Third Revolution in Sequencing Technology. *Trends in Genetics* **34:**666-681.
174. **Branton D, Deamer DW, Marziali A, Bayley H, Benner SA, Butler T, Di Ventra M, Garaj S, Hibbs A, Huang X, Jovanovich SB, Krstic PS, Lindsay S, Ling XS, Mastrangelo CH, Meller A, Oliver JS, Pershin YV, Ramsey JM, Riehn R, Soni GV, Tabard-Cossa V, Wanunu M,**

Wiggin M, Schloss JA. 2008. The potential and challenges of nanopore sequencing. *Nature biotechnology* **26:**1146-1153.

175. **Utturkar SM, Klingeman DM, Land ML, Schadt CW, Doktycz MJ, Pelletier DA, Brown SD.** 2014. Evaluation and validation of de novo and hybrid assembly techniques to derive high-quality genome sequences. *Bioinformatics* **30:**2709-2716.
176. **Koren S, Schatz MC, Walenz BP, Martin J, Howard JT, Ganapathy G, Wang Z, Rasko DA, McCombie WR, Jarvis ED.** 2012. Hybrid error correction and de novo assembly of single-molecule sequencing reads. *Nature biotechnology* **30:**693-700.
177. **Zimin AV, Marçais G, Puiu D, Roberts M, Salzberg SL, Yorke JA.** 2013. The MaSuRCA genome assembler. *Bioinformatics* **29:**2669-2677.
178. **Antipov D, Korobeynikov A, McLean JS, Pevzner PA.** 2016. hybridSPAdes: an algorithm for hybrid assembly of short and long reads. *Bioinformatics* **32:**1009-1015.
179. **Wick RR, Judd LM, Gorrie CL, Holt KE.** 2017. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* **13:**e1005595.
180. **Ye C, Hill CM, Wu S, Ruan J, Ma ZS.** 2016. DBG2OLC: Efficient Assembly of Large Genomes Using Long Erroneous Reads of the Third Generation Sequencing Technologies. *Sci Rep* **6:**31900.
181. **Di Genova A, Buena-Atienza E, Ossowski S, Sagot M-F.** 2021. Efficient hybrid de novo assembly of human genomes with WENGAN. *Nature Biotechnology* **39:**422-430.
182. **Ijaq J, Chandrasekharan M, Poddar R, Bethi N, Sundararajan VSJFiG.** 2015. Annotation and curation of uncharacterized proteins- challenges. **6**.
183. **Seemann T.** 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30:**2068-2069.
184. **Ejigu GF, Jung J.** 2020. Review on the Computational Genome Annotation of Sequences Obtained by Next-Generation Sequencing. *Biology* **9:**295.
185. **Su M, Satola SW, Read TD.** 2019. Genome-Based Prediction of Bacterial Antibiotic Resistance. **57:**e01405-01418.
186. **Yee R, Simner PJ.** 2022. Next-Generation Sequencing Approaches to Predicting Antimicrobial Susceptibility Testing Results. *Clin Lab Med* **42:**557-572.
187. **Mariem Saidani LM, Amira Chaouechi, Ines Tabib, Estelle Saras, Alya Soudani, Monia Daaloul-Jedidi, Aymen Mamlouk, Faten Ben Chehida, Chahid Chakroun, Jean-Yves Madec, and Marisa Haenni.** 2019. High Genetic Diversity of Enterobacteriaceae Clones and Plasmids Disseminating Resistance to Extended-Spectrum Cephalosporins and Colistin in Healthy Chicken in Tunisia. **25:**1507-1513.
188. **Naseer U, Sundsfjord AJMdr.** 2011. The CTX-M conundrum: dissemination of plasmids and Escherichia coli clones. **17:**83-97.
189. **Nicolas-Chanoine M-H, Bertrand X, Madec J-YJCmr.** 2014. Escherichia coli ST131, an intriguing clonal group. **27:**543-574.
190. **Hasman H, Aarestrup FM.** 2005. Relationship between copper, glycopeptide, and macrolide resistance among Enterococcus faecium strains isolated from pigs in Denmark between 1997 and 2003. *Antimicrob Agents Chemother* **49:**454-456.
191. **Wong MH-y, Liu L, Yan M, Chan EW-c, Chen S.** 2015. Dissemination of IncI2 Plasmids That Harbor the bla_{CTX-M} Element among Clinical Salmonella Isolates. **59:**5026-5028.
192. **Seiler C, Berendonk TU.** 2012. Heavy metal driven co-selection of antibiotic resistance in soil and water bodies impacted by agriculture and aquaculture. *Front Microbiol* **3:**399.
193. **Pal C, Bengtsson-Palme J, Kristiansson E, Larsson DGJJBG.** 2015. Co-occurrence of resistance genes to antibiotics, biocides and metals reveals novel insights into their co-selection potential. **16**.
194. **Villagómez-Estrada S, Pérez JF, Darwich L, Vidal A, van Kuijk S, Melo-Durán D, Solà-Oriol D.** 2020. Effects of copper and zinc sources and inclusion levels of copper on weanling pig performance and intestinal microbiota. *J Anim Sci* **98**.
195. **Johnning A, Karami N, Tång Hallbäck E, Müller V, Nyberg L, Buongermino Pereira M, Stewart C, Ambjörnsson T, Westerlund F, Adlerberth I, Kristiansson E.** 2018. The resistomes of six carbapenem-resistant pathogens - a critical genotype-phenotype analysis. *Microbial genomics* **4:**e000233.
196. **Trongjit S, Chuanchuen R.** 2021. Whole genome sequencing and characteristics of Escherichia coli with co-existence of ESBL and mcr genes from pigs. *PloS one* **16:**e0260011-e0260011.

197. **Wang Y, Zhang R, Li J, Wu Z, Yin W, Schwarz S, Tyrrell JM, Zheng Y, Wang S, Shen Z, Liu Z, Liu J, Lei L, Li M, Zhang Q, Wu C, Zhang Q, Wu Y, Walsh TR, Shen J.** 2017. Comprehensive resistome analysis reveals the prevalence of NDM and MCR-1 in Chinese poultry production. *Nature Microbiology* **2:**16260.
198. **Liu YY, Wang Y, Walsh TR, Yi LX, Zhang R, Spencer J, Doi Y, Tian G, Dong B, Huang X, Yu LF, Gu D, Ren H, Chen X, Lv L, He D, Zhou H, Liang Z, Liu JH, Shen J.** 2016. Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. *Lancet Infect Dis* **16:**161-168.
199. **Hasman H, Hammerum AM, Hansen F, Hendriksen RS, Olesen B, Agersø Y, Zankari E, Leekitcharoenphon P, Stegger M, Kaas RS, Cavaco LM, Hansen DS, Aarestrup FM, Skov RL.** 2015. Detection of mcr-1 encoding plasmid-mediated colistin-resistant Escherichia coli isolates from human bloodstream infection and imported chicken meat, Denmark 2015. *Euro Surveill* **20**.
200. **Dantas Palmeira J, Haenni M, Madec JY, Ferreira HMN.** 2021. First Global Report of Plasmid-Mediated mcr-1 and Extended-Spectrum Beta-Lactamase-Producing Escherichia coli from Sheep in Portugal. *Antibiotics (Basel)* **10**.
201. **Hasman H, Hammerum AM, Hansen F, Hendriksen RS, Olesen B, Agersø Y, Zankari E, Leekitcharoenphon P, Stegger M, Kaas RS, Cavaco LM, Hansen DS, Aarestrup FM, Skov RL.** 2015. Detection of mcr-1 encoding plasmid-mediated colistin-resistant Escherichia coli isolates from human bloodstream infection and imported chicken meat, Denmark 2015. **20:**30085.
202. **Hammerl JA, Borowiak M, Schmoger S, Shamoun D, Grobbel M, Malorny B, Tenhagen B-A, Käsbohrer A.** 2018. mcr-5 and a novel mcr-5.2 variant in Escherichia coli isolates from food and food-producing animals, Germany, 2010 to 2017. *Journal of Antimicrobial Chemotherapy* **73:**1433-1435.
203. **Doumith M, Godbole G, Ashton P, Larkin L, Dallman T, Day M, Day M, Muller-Pebody B, Ellington MJ, de Pinna E, Johnson AP, Hopkins KL, Woodford N.** 2016. Detection of the plasmid-mediated mcr-1 gene conferring colistin resistance in human and food isolates of Salmonella enterica and Escherichia coli in England and Wales. *J Antimicrob Chemother* **71:**2300-2305.
204. **Ricaurte JC, Boucher HW, Turett GS, Moellering RC, Labombardi VJ, Kislak JW.** 2001. Chloramphenicol treatment for vancomycin-resistant Enterococcus faecium bacteremia. *Clin Microbiol Infect* **7:**17-21.
205. **Jiang HX, Zeng ZL, Chen ZL, Liu JJ, Fung KF.** 2006. Pharmacokinetics of florfenicol in pigs following intravenous, intramuscular or oral administration and the effects of feed intake on oral dosing. *J Vet Pharmacol Ther* **29:**153-156.
206. **Jung Y-H, Cha M-H, Woo G-J, Chi Y-M.** 2021. Characterization of oxazolidinone and phenicol resistance genes in non-clinical enterococcal isolates from Korea. *Journal of Global Antimicrobial Resistance* **24:**363-369.
207. **Cai J, Schwarz S, Chi D, Wang Z, Zhang R, Wang Y.** 2019. Faecal carriage of optrA-positive enterococci in asymptomatic healthy humans in Hangzhou, China. *Clin Microbiol Infect* **25:**630.e631-630.e636.
208. **Kang Z-Z, Lei C-W, Kong L-H, Wang Y-L, Ye X-L, Ma B-H, Wang X-C, Li C, Zhang Y, Wang H-N.** 2019. Detection of transferable oxazolidinone resistance determinants in Enterococcus faecalis and Enterococcus faecium of swine origin in Sichuan Province, China. *Journal of Global Antimicrobial Resistance* **19:**333-337.
209. **Hiltunen T, Virta M, Laine A-L.** 2017. Antibiotic resistance in the wild: an eco-evolutionary perspective. *Philosophical transactions of the Royal Society of London Series B, Biological sciences* **372:**20160039.
210. **Hernando-Amado S, Coque TM, Baquero F, Martínez JL.** 2019. Defining and combating antibiotic resistance from One Health and Global Health perspectives. *Nature Microbiology* **4:**1432-1442.
211. **Murray BE.** 1990. The life and times of the Enterococcus. *Clin Microbiol Rev* **3:**46-65.
212. **Moreno MF, Sarantinopoulos P, Tsakalidou E, De Vuyst L.** 2006. The role and application of enterococci in food and health. *International journal of food microbiology* **106:**1-24.
213. **Barnes AMT, Frank KL, Dunny GM.** 2021. Enterococcal Endocarditis: Hiding in Plain Sight. *Front Cell Infect Microbiol* **11:**722482.

214. **Pinkes ME, White C, Wong CS.** 2019. Native-valve Enterococcus hirae endocarditis: a case report and review of the literature. *BMC infectious diseases* **19:**891-891.
215. **Brayer S, Linn A, Holt S, Ellery K, Mitchell S, Williams J.** 2019. Enterococcus hirae Bacteremia in an Infant: Case Report and Review of the Literature. *J Pediatric Infect Dis Soc* **8:**571-573.
216. **Winther M, Dalager-Pedersen M, Tarpgaard IH, Nielsen HL.** 2020. First Danish case of infective endocarditis caused by Enterococcus hirae. *BMJ Case Rep* **13**.
217. **Zaheer R, Cook SR, Barbieri R, Goji N, Cameron A, Petkau A, Polo RO, Tymensen L, Stamm C, Song J, Hannon S, Jones T, Church D, Booker CW, Amoako K, Van Domselaar G, Read RR, McAllister TA.** 2020. Surveillance of Enterococcus spp. reveals distinct species and antimicrobial resistance diversity across a One-Health continuum. *Sci Rep* **10:**3937.
218. **Heval Can B, Aydın D, Sema Ü, Yeliz Tanrıverdi Ç, Esra T.** 2020. Enterococcus hirae as a cause of bacteremic urinary tract infection: first case report from Turkey. *The Journal of Infection in Developing Countries* **14**.
219. **Nakamura T, Ishikawa K, Matsuo T, Kawai F, Uehara Y, Mori N.** 2021. Enterococcus hirae bacteremia associated with acute pyelonephritis in a patient with alcoholic cirrhosis: a case report and literature review. *BMC Infectious Diseases* **21:**999.
220. **Pãosinho A, Azevedo T, Alves JV, Costa IA, Carvalho G, Peres SR, Baptista T, Borges F, Mansinho K.** 2016. Acute Pyelonephritis with Bacteremia Caused by Enterococcus hirae: A Rare Infection in Humans. *Case reports in infectious diseases* **2016:**4698462.
221. **Chan T-S, Wu M-S, Suk F-M, Chen C-N, Chen Y-F, Hou Y-H, Lien G-S.** 2012. Enterococcus hirae-related acute pyelonephritis and cholangitis with bacteremia: An unusual infection in humans. *The Kaohsiung Journal of Medical Sciences* **28:**111-114.
222. **Pinkes M, White C, Wong C.** 2019. Native-valve Enterococcus hirae endocarditis: a case report and review of the literature. *BMC infectious diseases* **19:**891.
223. **Talarmin J-P, Pineau S, Guillouzouic A, Boutoille D, Giraudeau C, Reynaud A, Lepelletier D, Corvec S.** 2011. Relapse of Enterococcus hirae Prosthetic Valve Endocarditis. *Journal of clinical microbiology* **49:**1182-1184.
224. **Pinkes ME, White C, Wong CS.** Native-valve Enterococcus hirae endocarditis: a case report and review of the literature. *BMC infectious diseases*. **19**(1)**:**891. doi:10.1186/s12879-019-4532-z.
225. **Tan C-K, Lai C-C, Wang J-Y, Lin S-H, Liao C-H, Huang Y-T, Wang C-Y, Lin H-I, Hsueh P-R.** 2010. Bacteremia caused by non-faecalis and non-faecium enterococcus species at a Medical center in Taiwan, 2000 to 2008. *Journal of Infection* **61:**34-43.
226. **Bourafa N, Loucif L, Boutefnouchet N, Rolain JM.** 2015. Enterococcus hirae, an unusual pathogen in humans causing urinary tract infection in a patient with benign prostatic hyperplasia: first case report in Algeria. *New Microbes New Infect* **8:**7-9.
227. **Wingett SW, Andrews S.** 2018. FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Research* **7**.
228. **Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA.** 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* **19:**455-477.
229. **Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV.** 2012. Identification of acquired antimicrobial resistance genes. *Journal of antimicrobial chemotherapy* **67:**2640-2644.
230. **Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS.** 2016. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic acids research* **44:**W16-W21.
231. **Letunic I, Bork P.** 2021. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research* **49:**W293-W296.
232. **Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J.** 2015. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics (Oxford, England)* **31:**3691-3693.
233. **Hadfield J, Croucher NJ, Goater RJ, Abudahab K, Aanensen DM, Harris SR.** 2017. Phandango: an interactive viewer for bacterial population genomics. *Bioinformatics* **34:**292-293.
234. **Marinier E, Zaheer R, Berry C, Weedmark KA, Domaratzki M, Mabon P, Knox NC, Reimer AR, Graham MR, Chui L, Patterson-Fortin L, Zhang J, Pagotto F, Farber J, Mahony J, Seyer K, Bekal S, Tremblay C, Isaac-Renton J, Prystajecky N, Chen J, Slade P, Van Domselaar G.**

2017. Neptune: a bioinformatics tool for rapid discovery of genomic variation in bacterial populations. *Nucleic acids research* **45:**e159-e159.

235. **Giraffa G.** 2014. Enterococcus, p 674-679. *In* Batt CA, Tortorello ML (ed), *Encyclopedia of Food Microbiology (Second Edition)* doi:https://doi.org/10.1016/B978-0-12-384730-0.00098-7. Academic Press, Oxford.
236. **De Graef EM, Devriese LA, Vancanneyt M, Baele M, Collins MD, Lefebvre K, Swings J, Haesebrouck F.** 2003. Description of Enterococcus canis sp. nov. from dogs and reclassification of Enterococcus porcinus Teixeira et al. 2001 as a junior synonym of Enterococcus villorum Vancanneyt et al 2001. *Int J Syst Evol Microbiol* **53:**1069-1074.
237. **de Vaux A, Laguerre G, Diviès C, Prévost H.** 1998. Enterococcus asini sp. nov. isolated from the caecum of donkeys (Equus asinus). *Int J Syst Bacteriol* **48 Pt 2:**383-387.
238. **Devriese LA, Ceyssens K, Rodrigues UM, Collins MD.** 1990. Enterococcus columbae, a species from pigeon intestines. *FEMS Microbiol Lett* **59:**247-251.
239. **Farrow JA, Collins MD.** 1985. Enterococcus hirae, a new species that includes amino acid assay strain NCDO 1258 and strains causing growth depression in young chickens. *International Journal of Systematic and Evolutionary Microbiology* **35:**73-75.
240. **Fogarty LR, Haack SK, Wolcott MJ, Whitman RL.** 2003. Abundance and characteristics of the recreational water quality indicator bacteria Escherichia coli and enterococci in gull faeces. *J Appl Microbiol* **94:**865-878.
241. **Andrew PW, Mitchell TJ.** The biology of Streptococci and Enterococci, p. *In* (ed), Journals Customer Services Blackwell Scientific Publications Ltd,
242. **Law-Brown J, Meyers PR.** 2003. Enterococcus phoeniculicola sp. nov., a novel member of the enterococci isolated from the uropygial gland of the Red-billed Woodhoopoe, Phoeniculus purpureus. *Int J Syst Evol Microbiol* **53:**683-685.
243. **Muniesa M, Jofre J, Lucena F.** 1999. Occurrence and numbers of bacteriophages and bacterial indicators in faeces of yellow-legged seagull (Larus cachinnans). *Lett Appl Microbiol* **29:**421-423.
244. **Layton BA, Walters SP, Lam LH, Boehm AB.** 2010. Enterococcus species distribution among human and animal hosts using multiplex PCR. *J Appl Microbiol* **109:**539-547.
245. **Maria da Gloria SC, Shewmaker PL, Steigerwalt AG, Morey RE, Sampson A, Joyce K, Barrett TJ, Teixeira LM, Facklam RR.** 2006. Enterococcus caccae sp. nov., isolated from human stools. *International journal of systematic and evolutionary microbiology* **56:**1505-1508.
246. **Byappanahalli MN, Nevers MB, Korajkic A, Staley ZR, Harwood VJ.** 2012. Enterococci in the environment. *Microbiology and molecular biology reviews : MMBR* **76:**685-706.
247. **Svec P, Devriese LA, Sedlácek I, Baele M, Vancanneyt M, Haesebrouck F, Swings J, Doskar J.** 2001. Enterococcus haemoperoxidus sp. nov. and Enterococcus moraviensis sp. nov., isolated from water. *Int J Syst Evol Microbiol* **51:**1567-1574.
248. **Švec P, Vancanneyt M, Sedláček I, Naser SM, Snauwaert C, Lefebvre K, Hoste B, Swings J.** 2006. Enterococcus silesiacus sp. nov. and Enterococcus termitis sp. nov. *Int J Syst Evol Microbiol* **56:**577-581.
249. **Švec P, Vancanneyt M, Sedláček I, Naser SM, Snauwaert C, Lefebvre K, Hoste B, Swings J.** 2006. Enterococcus silesiacus sp. nov. and Enterococcus termitis sp. nov. *International journal of systematic and evolutionary microbiology* **56:**577-581.
250. **Klein G.** 2003. Taxonomy, ecology and antibiotic resistance of enterococci from food and the gastro-intestinal tract. *Int J Food Microbiol* **88:**123-131.
251. **Müller T, Ulrich A, Ott EM, Müller M.** 2001. Identification of plant-associated enterococci. *Journal of Applied Microbiology* **91:**268-278.
252. **Caretti C, Lubello C.** 2003. Wastewater disinfection with PAA and UV combined treatment: a pilot plant study. *Water research* **37:**2365-2371.
253. **Heim S, Lleo M, Bonato B, Guzman CA, Canepari P.** 2002. The viable but nonculturable state and starvation are different stress responses of Enterococcus faecalis, as determined by proteome analysis. *J Bacteriol* **184:**6739-6745.
254. **Anderson KL, Whitlock JE, Harwood VJ.** 2005. Persistence and differential survival of fecal indicator bacteria in subtropical waters and sediments. *Appl Environ Microbiol* **71:**3041-3048.
255. **Francino MP.** 2016. Antibiotics and the Human Gut Microbiome: Dysbioses and Accumulation of Resistances. *Frontiers in Microbiology* **6**.

256. **Cameron A, McAllister TA.** 2016. Antimicrobial usage and resistance in beef production. *Journal of Animal Science and Biotechnology* **7:**1-22.
257. **Zaheer R, Cook S, Klima C, Stanford K, Alexander T, Topp E, Read R, Mcallister T.** 2013. Effect of subtherapeutic vs. therapeutic administration of macrolides on antimicrobial resistance in Mannheimia haemolytica and enterococci isolated from beef cattle. *Frontiers in Microbiology* **4**.
258. **Jackson CR, Lombard JE, Dargatz DA, Fedorka-Cray PJ.** 2011. Prevalence, species distribution and antimicrobial resistance of enterococci isolated from US dairy cattle. *Letters in applied microbiology* **52:**41-48.
259. **Beukers AG, Zaheer R, Cook SR, Stanford K, Chaves AV, Ward MP, McAllister TA.** 2015. Effect of in-feed administration and withdrawal of tylosin phosphate on antibiotic resistance in enterococci isolated from feedlot steers. *Frontiers in Microbiology* **6**.
260. **Gilad J, Borer A, Riesenberg K, Peled N, Shnaider A, Schlaeffer F.** 1998. Enterococcus hirae septicemia in a patient with end-stage renal disease undergoing hemodialysis. *European Journal of Clinical Microbiology and Infectious Diseases* **17:**576-577.
261. **Poyart C, Lambert T, Morand P, Abassade P, Quesne G, Baudouy Y, Trieu-Cuot P.** 2002. Native valve endocarditis due to Enterococcus hirae. *Journal of Clinical Microbiology* **40:**2689-2690.
262. **Talarmin J, Pineau S, Guillouzouic A, Boutoille D, Giraudeau C, Reynaud A, Lepelletier D, Corvec S.** 2011. Relapse of Enterococcus hirae prosthetic valve endocarditis. *Journal of Clinical Microbiology* **49:**1182-1184.
263. **Canalejo E, Ballesteros R, Cabezudo J, García-Arata M, Moreno J.** 2008. Bacteraemic spondylodiscitis caused by Enterococcus hirae. *European Journal of Clinical Microbiology & Infectious Diseases* **27:**613-615.
264. **Dicpinigaitis PV, De Aguirre M, Divito J.** 2015. Enterococcus hirae bacteremia associated with acute pancreatitis and septic shock. *Case reports in infectious diseases* **2015**.
265. **Hurd HS, Malladi S.** 2008. A stochastic assessment of the public health risks of the use of macrolide antibiotics in food animals. *Risk Analysis: An International Journal* **28:**695-710.
266. **Vikram A, Rovira P, Agga GE, Arthur TM, Bosilevac JM, Wheeler TL, Morley PS, Belk KE, Schmidt JW.** 2017. Impact of "raised without antibiotics" beef cattle production practices on occurrences of antimicrobial resistance. *Applied and environmental microbiology* **83:**e01682-01617.
267. **Zaheer R, Cook S, Klima C, Stanford K, Alexander T, Topp E, Read R, Mcallister T.** 2013. Effect of subtherapeutic vs. therapeutic administration of macrolides on antimicrobial resistance in Mannheimia haemolytica and enterococci isolated from beef cattle. *Frontiers in microbiology* **4:**133.
268. **Zaheer R, Lakin SM, Polo RO, Cook SR, Larney FJ, Morley PS, Booker CW, Hannon SJ, Van Domselaar G, Read RR.** 2019. Comparative diversity of microbiomes and Resistomes in beef feedlots, downstream environments and urban sewage influent. *BMC microbiology* **19:**1-17.
269. **Rovira P, McAllister T, Lakin SM, Cook SR, Doster E, Noyes NR, Weinroth MD, Yang X, Parker JK, Boucher C.** 2019. Characterization of the microbial resistome in conventional and "raised without antibiotics" beef and dairy production systems. *Frontiers in microbiology* **10:**1980.
270. **Safferling M, Griffith H, Jin J, Sharp J, De Jesus M, Ng C, Krulwich TA, Wang DN.** 2003. TetL tetracycline efflux protein from Bacillus subtilis is a dimer in the membrane and in detergent solution. *Biochemistry* **42:**13969-13976.
271. **Dönhöfer A, Franckenberg S, Wickles S, Berninghausen O, Beckmann R, Wilson DN.** 2012. Structural basis for TetM-mediated tetracycline resistance. *Proc Natl Acad Sci U S A* **109:**16900-16905.
272. **Barile S, Devirgiliis C, Perozzi G.** 2012. Molecular characterization of a novel mosaic tet(S/M) gene encoding tetracycline resistance in foodborne strains of Streptococcus bovis. *Microbiology (Reading)* **158:**2353-2362.
273. **Crespo MD, Olson JW, Altermann E, Siletzky RM, Kathariou S.** 2012. Chromosomal tet(O)-harboring regions in Campylobacter coli isolates from turkeys and swine. *Appl Environ Microbiol* **78:**8488-8491.
274. **Kazimierczak KA, Rincon MT, Patterson AJ, Martin JC, Young P, Flint HJ, Scott KP.** 2008. A new tetracycline efflux gene, tet(40), is located in tandem with tet(O/32/O) in a human gut

firmicute bacterium and in metagenomic library clones. *Antimicrob Agents Chemother* **52:**4001-4009.

275. **Opal SM, Pop-Vicas A.** 2015. 18 - Molecular Mechanisms of Antibiotic Resistance in Bacteria, p 235-251.e233. *In* Bennett JE, Dolin R, Blaser MJ (ed), *Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases (Eighth Edition)* doi:https://doi.org/10.1016/B978-1-4557-4801-3.00018-7. W.B. Saunders, Philadelphia.
276. **Okitsu N, Kaieda S, Yano H, Nakano R, Hosaka Y, Okamoto R, Kobayashi T, Inoue M.** 2005. Characterization of ermB gene transposition by Tn1545 and Tn917 in macrolide-resistant Streptococcus pneumoniae isolates. *Journal of clinical microbiology* **43:**168-173.
277. **Portillo A, Ruiz-Larrea F, Zarazaga M, Alonso A, Martinez JL, Torres C.** 2000. Macrolide Resistance Genes inEnterococcus spp. *Antimicrobial Agents and Chemotherapy* **44:**967-971.
278. **Chajęcka-Wierzchowska W, Zadernowska A, Łaniewska-Trokenheim Ł.** 2016. Diversity of Antibiotic Resistance Genes in Enterococcus Strains Isolated from Ready-to-Eat Meat Products. *Journal of Food Science* **81:**M2799-M2807.
279. **Kim YB, Seo KW, Jeon HY, Lim S-K, Sung HW, Lee YJ.** 2019. Molecular characterization of erythromycin and tetracycline-resistant Enterococcus faecalis isolated from retail chicken meats. *Poultry Science* **98:**977-983.
280. **Anonymous.** 2021. Detection of Linezolid-Resistant Enterococcus faecalis and Enterococcus faecium Isolates from the Layer Operation System in Korea. *Microbial Drug Resistance* **27:**1443-1449.
281. **TREMBLAY C-L, LETELLIER A, QUESSY S, BOULIANNE M, DAIGNAULT D, ARCHAMBAULT M.** 2011. Multiple-Antibiotic Resistance of Enterococcus faecalis and Enterococcus faecium from Cecal Contents in Broiler Chicken and Turkey Flocks Slaughtered in Canada and Plasmid Colocalization of tetO and ermB Genes. *Journal of Food Protection* **74:**1639-1648.
282. **Demirgül F, Tuncer Y.** 2017. Detection of Antibiotic Resistance and Resistance Genes in Enterococci Isolated from Sucuk, a Traditional Turkish Dry-Fermented Sausage. *Korean journal for food science of animal resources* **37:**670-681.
283. **Aarestrup FM.** 2000. Characterization of glycopeptide-resistant Enterococcus faecium (GRE) from broilers and pigs in Denmark: genetic evidence that persistence of GRE in pig herds is associated with coselection by resistance to macrolides. *Journal of Clinical Microbiology* **38:**2774-2777.
284. **Anonymous.** 2013. Phenotypic and Molecular Antibiotic Resistance Profile of Enterococcus faecalis and Enterococcus faecium Isolated from Different Traditional Fermented Foods. *Foodborne Pathogens and Disease* **10:**143-149.
285. **Chen L, Ruan Z, Wu Q, Zhou T, Li C.** 2010. Relationship between ermB gene and transposons Tn1545 and Tn917 in Enterococcus faecalis. *Chinese Journal of Nosocomiology* **20:**763-766.
286. **Wang Y, Xiong Y, Wang Z, Zheng J, Xu G, Deng Q, Wen Z, Yu Z.** 2021. Comparison of solithromycin with erythromycin in Enterococcus faecalis and Enterococcus faecium from China: antibacterial activity, clonality, resistance mechanism, and inhibition of biofilm formation. *The Journal of Antibiotics* **74:**143-151.
287. **Schmitz F-J, Sadurski R, Kray A, Boos M, Geisel R, Köhrer K, Verhoef J, Fluit AC.** 2000. Prevalence of macrolide-resistance genes in Staphylococcus aureus and Enterococcus faecium isolates from 24 European university hospitals. *Journal of Antimicrobial Chemotherapy* **45:**891-894.
288. **Beukers AG, Zaheer R, Goji N, Amoako KK, Chaves AV, Ward MP, McAllister TA.** 2017. Comparative genomics of Enterococcus spp. isolated from bovine feces. *BMC Microbiology* **17:**52.
289. **Palmer KL, Kos VN, Gilmore MS.** 2010. Horizontal gene transfer and the genomics of enterococcal antibiotic resistance. *Curr Opin Microbiol* **13:**632-639.
290. **Flannagan SE, Clewell DB.** 2002. Identification and characterization of genes encoding sex pheromone cAM373 activity in Enterococcus faecalis and Staphylococcus aureus. *Molecular Microbiology* **44:**803-817.
291. **Johnson CN, Sheriff EK, Duerkop BA, Chatterjee A, Margolin W.** 2021. Let Me Upgrade You: Impact of Mobile Genetic Elements on Enterococcal Adaptation and Evolution. *Journal of Bacteriology* **203:**e00177-00121.

292. **Costa Y, Galimand M, Leclercq R, Duval J, Courvalin P.** 1993. Characterization of the chromosomal aac(6')-Ii gene specific for Enterococcus faecium. *Antimicrobial Agents and Chemotherapy* **37:**1896-1903.
293. **Hashem YA, Amin HM, Essam TM, Yassin AS, Aziz RK.** 2017. Biofilm formation in enterococci: genotype-phenotype correlations and inhibition by vancomycin. *Sci Rep* **7:**5733.
294. **Vernikos GS.** 2020. A Review of Pangenome Tools and Recent Studies, p 89-112. *In* Tettelin H, Medini D (ed), *The Pangenome: Diversity, Dynamics and Evolution of Genomes* doi:10.1007/978-3-030-38281-0_4. Springer

295. **Ding W, Baumdicker F, Neher RA.** 2018. panX: pan-genome analysis and exploration. *Nucleic Acids Res* **46:**e5.
296. **Costa SS, Guimarães LC, Silva A, Soares SC, Baraúna RA.** 2020. First Steps in the Analysis of Prokaryotic Pan-Genomes. *Bioinformatics and Biology Insights* **14:**1177932220938064.
297. **Silhavy TJ, Kahne D, Walker S.** 2010. The bacterial cell envelope. *Cold Spring Harbor perspectives in biology* **2:**a000414.
298. **Koch AL.** 2003. Bacterial wall as target for attack: past, present, and future research. *Clinical microbiology reviews* **16:**673-687.
299. **Parisien A, Allain B, Zhang J, Mandeville R, Lan CQ.** 2008. Novel alternatives to antibiotics: bacteriophages, bacterial cell wall hydrolases, and antimicrobial peptides. *J Appl Microbiol* **104:**1-13.
300. **Jeckelmann J-M, Erni B.** 2020. Transporters of glucose and other carbohydrates in bacteria. *Pflügers Archiv-European Journal of Physiology***:**1-25.
301. **Özcengiz G, Öğülür İ.** 2015. Biochemistry, genetics and regulation of bacilysin biosynthesis and its significance more than an antibiotic. *New Biotechnology* **32:**612-619.
302. **Babasaki K, Takao T, Shimonishi Y, Kurahashi K.** 1985. Subtilosin A, a new antibiotic peptide produced by Bacillus subtilis 168: isolation, structural analysis, and biogenesis. *J Biochem* **98:**585-603.
303. **Algburi A, Zehm S, Netrebov V, Bren AB, Chistyakov V, Chikindas ML.** 2017. Subtilosin Prevents Biofilm Formation by Inhibiting Bacterial Quorum Sensing. *Probiotics Antimicrob Proteins* **9:**81-90.
304. **Sarasa SB, Mahendran R, Muthusamy G, Thankappan B, Selta DRF, Angayarkanni J.** 2020. A Brief Review on the Non-protein Amino Acid, Gamma-amino Butyric Acid (GABA): Its Production and Role in Microbes. *Current Microbiology* **77:**534-544.
305. **Mamuad LL, Lee S-S.** 2021. The Role of Glutamic Acid-producing Microorganisms in Rumen Microbial Ecosystems. *Journal of Life Science* **31:**520-526.
306. **Ben Braïek O, Smaoui S.** 2019. Enterococci: Between Emerging Pathogens and Potential Probiotics. *Biomed Res Int* **2019:**5938210.
307. **Adnan M, Patel M, Hadi S.** 2017. Functional and health promoting inherent attributes of Enterococcus hirae F2 as a novel probiotic isolated from the digestive tract of the freshwater fish Catla catla. *PeerJ* **5:**e3085-e3085.
308. **Arokiyaraj S, Islam VIH, Bharanidharan R, Raveendar S, Lee J, Kim DH, Oh YK, Kim E-K, Kim KH.** 2014. Antibacterial, anti-inflammatory and probiotic potential of Enterococcus hirae isolated from the rumen of Bos primigenius. *World Journal of Microbiology and Biotechnology* **30:**2111-2118.
309. **Daillère R, Vétizou M, Waldschmitt N, Yamazaki T, Isnard C, Poirier-Colame V, Duong CP, Flament C, Lepage P, Roberti MP.** 2016. Enterococcus hirae and Barnesiella intestinihominis facilitate cyclophosphamide-induced therapeutic immunomodulatory effects. *Immunity* **45:**931-943.
310. **Ali L, Goraya MU, Arafat Y, Ajmal M, Chen JL, Yu D.** 2017. Molecular Mechanism of Quorum-Sensing in Enterococcus faecalis: Its Role in Virulence and Therapeutic Approaches. *Int J Mol Sci* **18**.
311. **Zhou X, Willems RJL, Friedrich AW, Rossen JWA, Bathoorn E.** 2020. Enterococcus faecium: from microbiological insights to practical recommendations for infection control and diagnostics. *Antimicrob Resist Infect Control* **9:**130.

312. **Michaux C, Sanguinetti M, Reffuveille F, Auffray Y, Posteraro B, Gilmore MS, Hartke A, Giard J-C.** 2011. SlyA is a transcriptional regulator involved in the virulence of Enterococcus faecalis. *Infection and immunity* **79:**2638-2645.
313. **Mäkinen PL, Clewell DB, An F, Mäkinen KK.** 1989. Purification and Substrate Specificity of a Strongly Hydrophobic Extracellular Metalloendopeptidase ("Gelatinase") from Streptococcus faecalis (Strain 0G1-10). *Journal of Biological Chemistry* **264:**3325-3334.
314. **Yang J, Tan G, Zhang T, White RH, Lu J, Ding H.** 2015. Deletion of the Proposed Iron Chaperones IscA/SufA Results in Accumulation of a Red Intermediate Cysteine Desulfurase IscS in Escherichia coli. *The Journal of biological chemistry* **290:**14226-14234.
315. **Zhang H, Niesel DW, Peterson JW, Klimpel GR.** 1998. Lipoprotein release by bacteria: potential factor in bacterial pathogenesis. *Infection and immunity* **66:**5196-5201.
316. **Hutchings MI, Palmer T, Harrington DJ, Sutcliffe IC.** 2009. Lipoprotein biogenesis in Gram-positive bacteria: knowing when to hold 'em, knowing when to fold 'em. *Trends in microbiology* **17:**13-21.
317. **Fulde M, Willenborg J, Huber C, Hitzmann A, Willms D, Seitz M, Eisenreich W, Valentin-Weigand P, Goethe R.** 2014. The arginine-ornithine antiporter ArcD contributes to biological fitness of Streptococcus suis. *Frontiers in Cellular and Infection Microbiology* **4**.
318. **Bilek HC, Deveci A, Ünal S, Tanrıverdi Çaycı Y, Tanyel E.** 2020. Enterococcus hirae as a cause of bacteremic urinary tract infection: first case report from Turkey. *J Infect Dev Ctries* **14:**1780-1482.
319. **Pinkes ME, White C, Wong CS.** 2019. Native-valve Enterococcus hirae endocarditis: a case report and review of the literature. *BMC Infect Dis* **19:**891.
320. **Dicpinigaitis PV, De Aguirre M, Divito J.** 2015. Enterococcus hirae Bacteremia Associated with Acute Pancreatitis and Septic Shock. *Case Rep Infect Dis* **2015:**123852.
321. **Ebeling CG, Romito BT.** 2019. Aortic valve endocarditis from Enterococcus hirae infection. *Proc (Bayl Univ Med Cent)* **32:**249-250.
322. **Nicklas JL, Moisan P, Stone MR, Gookin JL.** 2010. In situ molecular diagnosis and histopathological characterization of enteroadherent Enterococcus hirae infection in pre-weaning-age kittens. *J Clin Microbiol* **48:**2814-2820.
323. **Poweleit N, Czudnochowski N, Nakagawa R, Trinidad DD, Murphy KC, Sassetti CM, Rosenberg OS.** 2019. The structure of the endogenous ESX-3 secretion system. *Elife* **8**.
324. **Borst LB, Suyemoto MM, Scholl EH, Fuller FJ, Barnes HJ.** 2015. Comparative genomic analysis identifies divergent genomic features of pathogenic Enterococcus cecorum including a type IC CRISPR-Cas system, a capsule locus, an epa-like locus, and putative host tissue binding proteins. *PLoS One* **10:**e0121294.
325. **Hatosy SM, Martiny AC.** 2015. The ocean as a global reservoir of antibiotic resistance genes. *Applied and environmental microbiology* **81:**7593-7599.
326. **Stauff DL, Bagaley D, Torres VJ, Joyce R, Anderson KL, Kuechenmeister L, Dunman PM, Skaar EP.** 2008. Staphylococcus aureus HrtA is an ATPase required for protection against heme toxicity and prevention of a transcriptional heme stress response. *Journal of bacteriology* **190:**3588-3596.
327. **Rutherford ST, Bassler BL.** 2012. Bacterial quorum sensing: its role in virulence and possibilities for its control. *Cold Spring Harbor perspectives in medicine* **2:**a012427.
328. **Baquero F.** 2021. Threats of antibiotic resistance: an obliged reappraisal. *Int Microbiol* **24:**499-506.
329. **White A, Hughes JM.** 2019. Critical Importance of a One Health Approach to Antimicrobial Resistance. *Ecohealth* **16:**404-409.
330. **Lechner I, Freivogel C, Stärk K, Visschers V.** 2020. Exposure Pathways to Antimicrobial Resistance at the Human-Animal Interface—A Qualitative Comparison of Swiss Expert and Consumer Opinions. *Frontiers in Public Health* **8:**345.
331. **García-Solache M, Rice LB.** 2019. The Enterococcus: a Model of Adaptability to Its Environment. *Clinical microbiology reviews* **32:**e00058-00018.
332. **Poulin-Laprade D, Brouard JS, Gagnon N, Turcotte A, Langlois A, Matte JJ, Carrillo CD, Zaheer R, McAllister T, Topp E, Talbot G.** 2021. Resistance determinants and their genetic context in enterobacteria from a longitudinal study of pigs reared under various husbandry conditions. *Appl Environ Microbiol* **87**.

333. **Rehman MA, Yin X, Zaheer R, Goji N, Amoako KK, McAllister T, Pritchard J, Topp E, Diarra MS.** 2018. Genotypes and Phenotypes of Enterococci Isolated From Broiler Chickens. *Frontiers in Sustainable Food Systems* **2**.
334. **Bolger AM, Lohse M, Usadel B.** 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30:**2114-2120.
335. **Anton Bankevich SN, Dmitry Antipov, Alexey A. Gurevich, Mikhail Dvorkin, Alexander S. Kulikov, Valery M. Lesin, Sergey I. Nikolenko, Son Pham, Andrey D. Prjibelski, Alexey V. Pyshkin, Alexander V. Sirotkin, Nikolay Vyahhi, Glenn Tesler, Max A. Alekseyev, and Pavel A. Pevzner.** 2012. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology* **19:**455-477.
336. **Gurevich A, Saveliev V, Vyahhi N, Tesler G.** 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* **29:**1072-1075.
337. **Seemann T.** 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30:**2068-2069.
338. **Liu B, Zheng D, Zhou S, Chen L, Yang J.** 2022. VFDB 2022: a general classification scheme for bacterial virulence factors. *Nucleic Acids Res* **50:**D912-d917.
339. **Robertson J, Nash JHE.** 2018. MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. *Microbial Genomics* **4**.
340. **Hung W-W.** 2019. *Using groEL as the target for identification of Enterococcus faecium clades and 7 clinically relevant Enterococcus species*. Elsevier.
341. **Jolley K, Bray J, Maiden M.** 2018. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications [version 1; peer review: 2 approved]. *Wellcome Open Research* **3**.
342. **Jackson CR, Lombard JE, Dargatz DA, Fedorka-Cray PJ.** 2011. Prevalence, species distribution and antimicrobial resistance of enterococci isolated from US dairy cattle. *Lett Appl Microbiol* **52:**41-48.
343. **Diarra MS, Rempel H, Champagne J, Masson L, Pritchard J, Topp E.** 2010. Distribution of antimicrobial resistance and virulence genes in Enterococcus spp. and characterization of isolates from broiler chickens. *Appl Environ Microbiol* **76:**8033-8043.
344. **Li X, Alvarez V, Harper WJ, Wang HH.** 2011. Persistent, toxin-antitoxin system-independent, tetracycline resistance-encoding plasmid from a dairy Enterococcus faecium isolate. *Appl Environ Microbiol* **77:**7096-7103.
345. **Carson CA, Reid-Smith R, Irwin RJ, Martin WS, McEwen SA.** 2008. Antimicrobial use on 24 beef farms in Ontario. *Canadian journal of veterinary research = Revue canadienne de recherche veterinaire* **72:**109-118.
346. **CARSS.** 2023-01-23 2022. Canadian Antimicrobial Resistance Surveillance System (CARSS) Report 2022. https://www.canada.ca/en/public-health/services/publications/drugs-health-products/canadian-antimicrobial-resistance-surveillance-system-report-2022.html#a4.1. Accessed
347. **Bartoloni A, Pallecchi L, Rodríguez H, Fernandez C, Mantella A, Bartalesi F, Strohmeyer M, Kristiansson C, Gotuzzo E, Paradisi F, Rossolini GM.** 2009. Antibiotic resistance in a very remote Amazonas community. *Int J Antimicrob Agents* **33:**125-129.
348. **Nejidat A, Diaz-Reck D, Gelfand I, Zaady E.** 2021. Persistence and spread of tetracycline resistance genes and microbial community variations in the soil of animal corrals in a semi-arid planted forest. *FEMS Microbiol Ecol* **97**.
349. **Stefańska I, Kwiecień E, Kizerwetter-Świda M, Chrobak-Chmiel D, Rzewuska M.** 2022. Tetracycline, Macrolide and Lincosamide Resistance in Streptococcus canis Strains from Companion Animals and Its Genetic Determinants. *Antibiotics (Basel)* **11**.
350. **Agga GE, Kasumba J, Loughrin JH, Conte ED.** 2020. Anaerobic Digestion of Tetracycline Spiked Livestock Manure and Poultry Litter Increased the Abundances of Antibiotic and Heavy Metal Resistance Genes. *Front Microbiol* **11:**614424.
351. **Amachawadi RG, Scott HM, Alvarado CA, Mainini TR, Vinasco J, Drouillard JS, Nagaraja TG.** 2013. Occurrence of the transferable copper resistance gene tcrB among fecal enterococci of U.S. feedlot cattle fed copper-supplemented diets. *Appl Environ Microbiol* **79:**4369-4375.
352. **Yazdankhah S, Rudi K, Bernhoft A.** 2014. Zinc and copper in animal feed - development of resistance and co-resistance to antimicrobial agents in bacteria of animal origin. *Microbial ecology in health and disease* **25:**10.3402/mehd.v3425.25862.

353. **Hodel-Christian SL, Murray BE.** 1991. Characterization of the gentamicin resistance transposon Tn5281 from Enterococcus faecalis and comparison to staphylococcal transposons Tn4001 and Tn4031. *Antimicrob Agents Chemother* **35:**1147-1152.
354. **Mlynarczyk-Bonikowska B, Kowalewski C, Krolak-Ulinska A, Marusza W.** 2022. Molecular Mechanisms of Drug Resistance in Staphylococcus aureus. *Int J Mol Sci* **23**.
355. **Thomas WD, Jr., Archer GL.** 1989. Mobility of gentamicin resistance genes from staphylococci isolated in the United States: identification of Tn4031, a gentamicin resistance transposon from Staphylococcus epidermidis. *Antimicrob Agents Chemother* **33:**1335-1341.
356. **Szakacs TA, Kalan L, McConnell MJ, Eshaghi A, Shahinas D, McGeer A, Wright GD, Low DE, Patel SN.** 2014. Outbreak of vancomycin-susceptible Enterococcus faecium containing the wild-type vanA gene. *J Clin Microbiol* **52:**1682-1686.
357. **Palmer KL, Godfrey P, Griggs A, Kos VN, Zucker J, Desjardins C, Cerqueira G, Gevers D, Walker S, Wortman J, Feldgarden M, Haas B, Birren B, Gilmore MS.** 2012. Comparative genomics of enterococci: variation in Enterococcus faecalis, clade structure in E. faecium, and defining characteristics of E. gallinarum and E. casseliflavus. *mBio* **3:**e00318-00311.
358. **Sanderson H, Gray KL, Manuele A, Maguire F, Khan A, Liu C, Rudrappa CN, Nash JHE, Robertson J, Bessonov K, Oloni M, Alcock BP, Raphenya AR, McAllister TA, Peacock SJ, Raven KE, Gouliouris T, McArthur AG, Brinkman FSL, Fink RC, Zaheer R, Beiko RG.** 2022. Exploring the mobilome and resistome of Enterococcus faecium in a One Health context across two continents. *bioRxiv* doi:10.1101/2022.04.11.487771**:**2022.2004.2011.487771.
359. **Bonfiglio G, Russo G, Nicoletti G.** 2002. Recent developments in carbapenems. *Expert Opin Investig Drugs* **11:**529-544.
360. **Nicolau DP.** 2008. Carbapenems: a potent class of antibiotics. *Expert Opin Pharmacother* **9:**23-37.
361. **Zhanel GG, Wiebe R, Dilay L, Thomson K, Rubinstein E, Hoban DJ, Noreddin AM, Karlowsky JA.** 2007. Comparative review of the carbapenems. *Drugs* **67:**1027-1052.
362. **El-Gamal MI, Brahim I, Hisham N, Aladdin R, Mohammed H, Bahaaeldin A.** 2017. Recent updates of carbapenem antibiotics. *European Journal of Medicinal Chemistry* **131:**185-195.
363. **Zhanel GG, Lawrence CK, Adam H, Schweizer F, Zelenitsky S, Zhanel M, Lagacé-Wiens PRS, Walkty A, Denisuik A, Golden A, Gin AS, Hoban DJ, Lynch JP, Karlowsky JA.** 2018. Imipenem–Relebactam and Meropenem–Vaborbactam: Two Novel Carbapenem-β-Lactamase Inhibitor Combinations. *Drugs* **78:**65-98.
364. **O'Donnell JN, Miglis CM, Lee JY, Tuvell M, Lertharakul T, Scheetz MH.** 2016. Carbapenem susceptibility breakpoints, clinical implications with the moving target. *Expert Review of Anti-infective Therapy* **14:**389-401.
365. **Anonymous.** 2021. Global Antimicrobial Resistance and Use Surveillance System (GLASS), *on World Health Organization*. https://worldhealthorg.shinyapps.io/glass-dashboard/_w_685891e2/#!/amc. Accessed 11-June.
366. **Bush K, Bradford PA.** 2016. β-Lactams and β-Lactamase Inhibitors: An Overview. *Cold Spring Harbor Perspectives in Medicine* **6**.
367. **Schweizer HP.** 2003. Efflux as a mechanism of resistance to antimicrobials in Pseudomonas aeruginosa and related bacteria: unanswered questions. *Genet Mol Res* **2:**48-62.
368. **Meletis G, Exindari M, Vavatsi N, Sofianou D, Diza E.** 2012. Mechanisms responsible for the emergence of carbapenem resistance in Pseudomonas aeruginosa. *Hippokratia* **16:**303-307.
369. **Ambler RP, Coulson AFW, Frère JM, Ghuysen JM, Joris B, Forsman M, Levesque RC, Tiraby G, Waley SG.** 1991. A standard numbering scheme for the class A β-lactamases. *Biochemical Journal* **276:**269-270.
370. **Sugawara Y, Akeda Y, Sakamoto N, Takeuchi D, Motooka D, Nakamura S, Hagiya H, Yamamoto N, Nishi I, Yoshida H, Okada K, Zin KN, Aye MM, Tomono K, Hamada S.** 2017. Genetic characterization of blaNDM-harboring plasmids in carbapenem-resistant Escherichia coli from Myanmar. *PLoS One* **12:**e0184720.
371. **Franco MRG, Caiaffa-Filho HH, Burattini MN, Rossi F.** 2010. Metallo-beta-lactamases among imipenem-resistant Pseudomonas aeruginosa in a brazilian university hospital. *Clinics* **65:**825-829.
372. **Elshamy AA, Aboshanab KM.** 2020. A review on bacterial resistance to carbapenems: epidemiology, detection and treatment options. *Future Sci OA* **6:**Fso438.

373. **Webb HE, Bugarel M, den Bakker HC, Nightingale KK, Granier SA, Scott HM, Loneragan GH.** 2016. Carbapenem-Resistant Bacteria Recovered from Faeces of Dairy Cattle in the High Plains Region of the USA. *PLoS One* **11:**e0147363.
374. **Tello M, Oporto B, Lavín JL, Ocejo M, Hurtado A.** 2021. Characterization of a carbapenem-resistant Escherichia coli from dairy cattle harbouring blaNDM-1 in an IncC plasmid. *Journal of Antimicrobial Chemotherapy* **77:**843-845.
375. **Klotz P, Göttig S, Leidner U, Semmler T, Scheufen S, Ewers C.** 2017. Carbapenem-resistance and pathogenicity of bovine Acinetobacter indicus-like isolates. *PLOS ONE* **12:**e0171986.
376. **Anonymous.** 2022. Multilocus Sequence Typing of Carbapenem-Resistant Acinetobacter baumannii Isolates Harboring blaOXA-23 and blaIMP in Cattle from Punjab, Pakistan. *Microbial Drug Resistance* **28:**997-1002.
377. **Tshitshi L, Manganyi MC, Montso PK, Mbewe M, Ateba CN.** 2020. Extended Spectrum Beta-Lactamase-Resistant Determinants among Carbapenem-Resistant Enterobacteriaceae from Beef Cattle in the North West Province, South Africa: A Critical Assessment of Their Possible Public Health Implications. *Antibiotics* **9:**820.
378. **Fischer J, Rodríguez I, Schmoger S, Friese A, Roesler U, Helmuth R, Guerra B.** 2012. Escherichia coli producing VIM-1 carbapenemase isolated on a pig farm. *Journal of Antimicrobial Chemotherapy* **67:**1793-1795.
379. **Savin M, Bierbaum G, Mutters NT, Schmithausen RM, Kreyenschmidt J, García-Meniño I, Schmoger S, Käsbohrer A, Hammerl JA.** 2022. Genetic Characterization of Carbapenem-Resistant Klebsiella spp. from Municipal and Slaughterhouse Wastewater. *Antibiotics* **11:**435.
380. **Ahlstrom CA, Frick A, Pongratz C, Spink K, Xavier C, Bonnedahl J, Ramey AM.** 2021. Genomic comparison of carbapenem-resistant Enterobacteriaceae from humans and gulls in Alaska. *Journal of Global Antimicrobial Resistance* **25:**23-25.
381. **Darwich L, Vidal A, Seminati C, Albamonte A, Casado A, López F, Molina-López RA, Migura-Garcia L.** 2019. High prevalence and diversity of extended-spectrum β-lactamase and emergence of OXA-48 producing Enterobacterales in wildlife in Catalonia. *PLOS ONE* **14:**e0210686.
382. **De Oliveira DMP, Forde BM, Kidd TJ, Harris PNA, Schembri MA, Beatson SA, Paterson DL, Walker MJ.** 2020. Antimicrobial Resistance in ESKAPE Pathogens. *Clin Microbiol Rev* **33**.
383. **Canada PH.** 2022-11-23. Antimicrobial Resistance in Canada: Findings from the 2022 Canadian Antimicrobial Resistance Surveillance System Report. https://health-infobase.canada.ca/carss/blog.html?utm_source=International_Stakeholder&utm_medium=Email&utm_campaign=launch_2022_CARSS_report_November_2022_ENG. Accessed 11-June.
384. **Zaheer R, Cook S, Klima C, Stanford K, Alexander T, Topp E, Read R, Mcallister T.** 2013. Effect of subtherapeutic vs. therapeutic administration of macrolides on antimicrobial resistance in Mannheimia haemolytica and enterococci isolated from beef cattle. *Frontiers in Microbiology* **4**.
385. **Freire B, Ladra S, Paramá JR.** 2022. Memory-Efficient Assembly Using Flye. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* **19:**3564-3577.
386. **Li H.** 2016. Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. *Bioinformatics* **32:**2103-2110.
387. **Petkau A, Mabon P, Sieffert C, Knox NC, Cabral J, Iskander M, Iskander M, Weedmark K, Zaheer R, Katz LS, Nadon C, Reimer A, Taboada E, Beiko RG, Hsiao W, Brinkman F, Graham M, Van Domselaar G.** SNVPhyl: a single nucleotide variant phylogenomics pipeline for microbial genomic epidemiology. *Microbial genomics*. **3**(6)**:**e000116. doi:10.1099/mgen.0.000116.
388. **Katoh K, Rozewicki J, Yamada KD.** 2019. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform* **20:**1160-1166.
389. **Codjoe FS, Donkor ES.** 2018. Carbapenem Resistance: A Review. *Medical Sciences* **6:**1.
390. **Prevention CfDCa.** 2019. National Center for Emerging and Zoonotic Infectious Diseases. Antibiotic, Antimicrobial Resistance. https://www.cdc.gov/drugresistance/biggest-threats.html. Accessed 15 April 2022.
391. **Girlich D, Naas T, Nordmann P.** 2004. Biochemical characterization of the naturally occurring oxacillinase OXA-50 of Pseudomonas aeruginosa. *Antimicrob Agents Chemother* **48:**2043-2048.
392. **Wexler HM.** 2004. In vitro activity of ertapenem: review of recent studies. *J Antimicrob Chemother* **53 Suppl 2:**ii11-21.

393. **Tian J, Zhang G, Ju Y, Tang N, Li J, Jia R, Feng J.** 2018. Five novel carbapenem-hydrolysing OXA-type β-lactamase groups are intrinsic in Acinetobacter spp. *Journal of Antimicrobial Chemotherapy* **73:**3279-3284.
394. **Figueiredo S, Bonnin R, Poirel L, Duranteau J, Nordmann P.** 2012. Identification of the naturally occurring genes encoding carbapenem-hydrolysing oxacillinases from Acinetobacter haemolyticus, Acinetobacter johnsonii, and Acinetobacter calcoaceticus. *Clinical microbiology and infection* **18:**907-913.
395. **Figueiredo S, Poirel L, Seifert H, Mugnier P, Benhamou D, Nordmann P.** 2010. OXA-134, a Naturally Occurring Carbapenem-Hydrolyzing Class D β-Lactamase from Acinetobacter lwoffii. *Antimicrobial Agents and Chemotherapy* **54:**5372-5375.
396. **Wong D, Nielsen TB, Bonomo RA, Pantapalangkoor P, Luna B, Spellberg B.** 2017. Clinical and Pathophysiological Overview of Acinetobacter Infections: a Century of Challenges. *Clinical Microbiology Reviews* **30:**409-447.
397. **Yang Z, Liu W, Cui Q, Niu W, Li H, Zhao X, Wei X, Wang X, Huang S, Dong D, Lu S, Bai C, Li Y, Huang L, Yuan J.** 2014. Prevalence and Detection of Stenotrophomonas maltophilia Carrying Metallo-β-lactamase blaL1 in Beijing, China. *Frontiers in Microbiology* **5**.
398. **Gideskog M, Welander J, Melhus Å.** 2020. Cluster of S. maltophilia among patients with respiratory tract infections at an intensive care unit. *Infection Prevention in Practice* **2:**100097.
399. **Geller M, Nunes CP, Oliveira L, Nigri R.** 2018. S. maltophilia pneumonia: A case report. *Respir Med Case Rep* **24:**44-45.
400. **Brooke JS.** 2014. New strategies against Stenotrophomonas maltophilia: a serious worldwide intrinsically drug-resistant opportunistic pathogen. *Expert Rev Anti Infect Ther* **12:**1-4.
401. **Giannattasio-Ferraz S, Ene A, Gomes VJ, Queiroz CO, Maskeri L, Oliveira AP, Putonti C, Barbosa-Stancioli EF.** 2022. Escherichia coli and Pseudomonas aeruginosa Isolated From Urine of Healthy Bovine Have Potential as Emerging Human and Bovine Pathogens. *Frontiers in Microbiology* **13**.
402. **Gellatly SL, Hancock RE.** 2013. Pseudomonas aeruginosa: new insights into pathogenesis and host defenses. *Pathog Dis* **67:**159-173.
403. **Moradali MF, Ghods S, Rehm BHA.** 2017. Pseudomonas aeruginosa Lifestyle: A Paradigm for Adaptation, Survival, and Persistence. *Frontiers in Cellular and Infection Microbiology* **7**.
404. **Wolfgang MC, Kulasekara BR, Liang X, Boyd D, Wu K, Yang Q, Miyada CG, Lory S.** 2003. Conservation of genome content and virulence determinants among clinical and environmental isolates of Pseudomonas aeruginosa. *Proceedings of the National Academy of Sciences* **100:**8484-8489.
405. **Mathee K, Narasimhan G, Valdes C, Qiu X, Matewish JM, Koehrsen M, Rokas A, Yandava CN, Engels R, Zeng E, Olavarietta R, Doud M, Smith RS, Montgomery P, White JR, Godfrey PA, Kodira C, Birren B, Galagan JE, Lory S.** 2008. Dynamics of Pseudomonas aeruginosa genome evolution. *Proceedings of the National Academy of Sciences* **105:**3100-3105.
406. **Strom MS, Lory S.** 1993. STRUCTURE-FUNCTION AND BIOGENESIS OF THE TYPE IV PILI. *Annual Review of Microbiology* **47:**565-596.
407. **Graupner S, Frey V, Hashemi R, Lorenz MG, Brandes G, Wackernagel W.** 2000. Type IV Pilus Genes pilA andpilC of Pseudomonas stutzeri Are Required for Natural Genetic Transformation, and pilA Can Be Replaced by Corresponding Genes from Nontransformable Species. *Journal of Bacteriology* **182:**2184-2190.
408. **Zhou H-Q, Ning L-W, Zhang H-X, Guo F-B.** 2014. Analysis of the Relationship between Genomic GC Content and Patterns of Base Usage, Codon Usage and Amino Acid Usage in Prokaryotes: Similar GC Content Adopts Similar Compositional Frequencies Regardless of the Phylogenetic Lineages. *PLOS ONE* **9:**e107319.
409. **Kung VL, Ozer EA, Hauser AR.** 2010. The accessory genome of Pseudomonas aeruginosa. *Microbiol Mol Biol Rev* **74:**621-641.
410. **Puzari M, Chetia P.** 2017. RND efflux pump mediated antibiotic resistance in Gram-negative bacteria Escherichia coli and Pseudomonas aeruginosa: a major issue worldwide. *World J Microbiol Biotechnol* **33:**24.

411. **Savinova TA, Samchenko AA, Bocharova YA, Mayansky NA, Chebotar IV.** 2021. Computer Program for Detection and Analyzing the Porin-Mediated Antibiotic Resistance of Bacteria. *Sovrem Tekhnologii Med* **13:**15-22.
412. **Li XZ, Nikaido H, Poole K.** 1995. Role of mexA-mexB-oprM in antibiotic efflux in Pseudomonas aeruginosa. *Antimicrob Agents Chemother* **39:**1948-1953.
413. **Köhler T, Kok M, Michea-Hamzehpour M, Plesiat P, Gotoh N, Nishino T, Curty LK, Pechere JC.** 1996. Multidrug efflux in intrinsic resistance to trimethoprim and sulfamethoxazole in Pseudomonas aeruginosa. *Antimicrobial Agents and Chemotherapy* **40:**2288-2290.
414. **Rojo-Bezares B, Estepa V, Cebollada R, de Toro M, Somalo S, Seral C, Castillo FJ, Torres C, Sáenz Y.** 2014. Carbapenem-resistant Pseudomonas aeruginosa strains from a Spanish hospital: characterization of metallo-beta-lactamases, porin OprD and integrons. *Int J Med Microbiol* **304:**405-414.
415. **López-García A, del Carmen Rocha-Gracia R, Bello-López E, Juárez-Zelocualtecalt C, Sáenz Y, Castañeda-Lucio M, López-Pliego L, González-Vázquez MC, Torres C, Ayala-Nuñez T.** 2018. Characterization of antimicrobial resistance mechanisms in carbapenem-resistant Pseudomonas aeruginosa carrying IMP variants recovered from a Mexican hospital. *Infection and drug resistance* **11:**1523.
416. **Noble RC, Overman SB.** 1994. Pseudomonas stutzeri infection a review of hospital isolates and a review of the literature. *Diagnostic Microbiology and Infectious Disease* **19:**51-56.
417. **Berglund F, Marathe NP, Österlund T, Bengtsson-Palme J, Kotsakis S, Flach C-F, Larsson DGJ, Kristiansson E.** 2017. Identification of 76 novel B1 metallo-β-lactamases through large-scale screening of genomic and metagenomic data. *Microbiome* **5:**134.
418. **Fallah F, Borhan RS, Hashemi A.** 2013. Detection of bla(IMP) and bla(VIM) metallo-β-lactamases genes among Pseudomonas aeruginosa strains. *International journal of burns and trauma* **3 2:**122-124.
419. **Senda K, Arakawa Y, Nakashima K, Ito H, Ichiyama S, Shimokata K, Kato N, Ohta M.** 1996. Multifocal outbreaks of metallo-beta-lactamase-producing Pseudomonas aeruginosa resistant to broad-spectrum beta-lactams, including carbapenems. *Antimicrob Agents Chemother* **40:**349-353.
420. **Sekiguchi J-i, Morita K, Kitao T, Watanabe N, Okazaki M, Miyoshi-Akiyama T, Kanamori M, Kirikae T.** 2008. KHM-1, a Novel Plasmid-Mediated Metallo-β-Lactamase from a Citrobacter freundii Clinical Isolate. *Antimicrobial Agents and Chemotherapy* **52:**4194-4197.
421. **Azhar EI, Papadioti A, Bibi F, Ashshi AM, Raoult D, Angelakis E.** 2017. 'Pseudomonas saudiphocaensis' sp. nov., a new bacterial species isolated from currency notes collected during the Hajj pilgrimage in 2012 at Makkah, Saudi Arabia. *New Microbes New Infect* **15:**131-133.
422. **Gupta TB, Risson AN, Jáuregui R, Maclean PH, Brightwell G.** 2021. Draft Genome Sequence of Pseudomonas saudiphocaensis Strain AGROB56, Isolated from a Dairy Farm in New Zealand. *Microbiology Resource Announcements* **10**.
423. **Ahmad M, Khan AU.** 2019. Global economic impact of antibiotic resistance: A review. *J Glob Antimicrob Resist* **19:**313-316.
424. **Bright-Ponte SJ.** 2020. Antimicrobial use data collection in animal agriculture. *Zoonoses Public Health* **67 Suppl 1:**1-5.
425. **Unno T, Staley C, Brown CM, Han D, Sadowsky MJ, Hur HG.** 2018. Fecal pollution: new trends and challenges in microbial source tracking using next-generation sequencing. *Environ Microbiol* **20:**3132-3140.
426. **Canada PHAo.** 2022-04-04. Canadian Antimicrobial Resistance Surveillance System Report 2021. https://www.canada.ca/en/public-health/services/publications/drugs-health-products/canadian-antimicrobial-resistance-surveillance-system-report-2021.html#a52. Accessed 11-June.
427. **Tang R, Luo R, Tang S, Song H, Chen X.** 2022. Machine learning in predicting antimicrobial resistance: a systematic review and meta-analysis. *Int J Antimicrob Agents* **60:**106684.
428. **Clewell DB, Weaver KE, Dunny GM, Coque TM, Francia MV, Hayes F.** 2014. Extrachromosomal and Mobile Elements in Enterococci: Transmission, Maintenance, and Epidemiology. *In* Gilmore MS, Clewell DB, Ike Y, Shankar N (ed), *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*. Massachusetts Eye and Ear Infirmary, Boston.

429. **Larrañaga P, Calvo B, Santana R, Bielza C, Galdiano J, Inza I, Lozano JA, Armañanzas R, Santafé G, Pérez A, Robles V.** 2006. Machine learning in bioinformatics. *Brief Bioinform* **7:**86-112.
430. **Maguire F, Rehman MA, Carrillo C, Diarra MS, Beiko RG.** 2019. Identification of Primary Antimicrobial Resistance Drivers in Agricultural Nontyphoidal Salmonella enterica Serovars by Using Machine Learning. *mSystems* **4**.
431. **Sunuwar J, Azad RK.** 2021. A machine learning framework to predict antibiotic resistance traits and yet unknown genes underlying resistance to specific antibiotics in bacterial strains. *Brief Bioinform* **22**.
432. **Su YA, He P, Clewell DB.** 1992. Characterization of the tet(M) determinant of Tn916: evidence for regulation by transcription attenuation. *Antimicrob Agents Chemother* **36:**769-778.
433. **Hör J, Gorski SA, Vogel J.** 2018. Bacterial RNA Biology on a Genome Scale. *Mol Cell* **70:**785-799.
434. **Molina-Santiago C, Daddaoua A, Gómez-Lozano M, Udaondo Z, Molin S, Ramos JL.** 2015. Differential transcriptional response to antibiotics by Pseudomonas putida DOT-T1E. *Environ Microbiol* **17:**3251-3262.
435. **Zhang YF, Han K, Chandler CE, Tjaden B, Ernst RK, Lory S.** 2017. Probing the sRNA regulatory landscape of P. aeruginosa: post-transcriptional control of determinants of pathogenicity and antibiotic susceptibility. *Mol Microbiol* **106:**919-937.
436. **Sinel C, Augagneur Y, Sassi M, Bronsard J, Cacaci M, Guérin F, Sanguinetti M, Meignen P, Cattoir V, Felden B.** 2017. Small RNAs in vancomycin-resistant Enterococcus faecium involved in daptomycin response and resistance. *Sci Rep* **7:**11067.

Appendices

Appendix 1- Supplementary for Chapter 2: Genomic characterization of *Enterococcus hirae* from beef cattle feedlots and associated environmental continuum

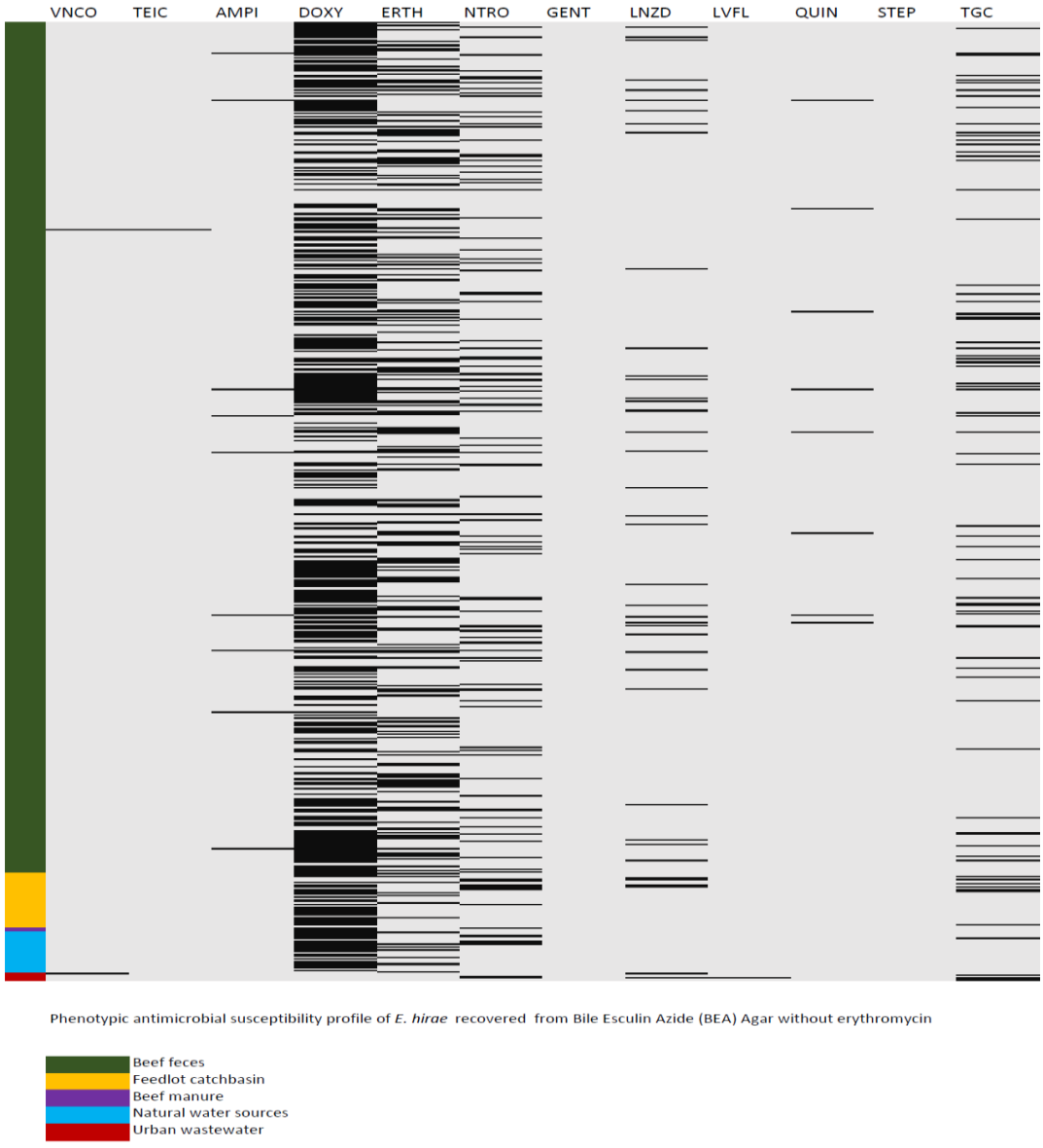

Figure S1. Heatmap of phenotypic resistance profile of *E. hirae* isolates (n=666), recovered from beef production system (n=632), natural water sources (n=28), and urban waste water (n= 6).

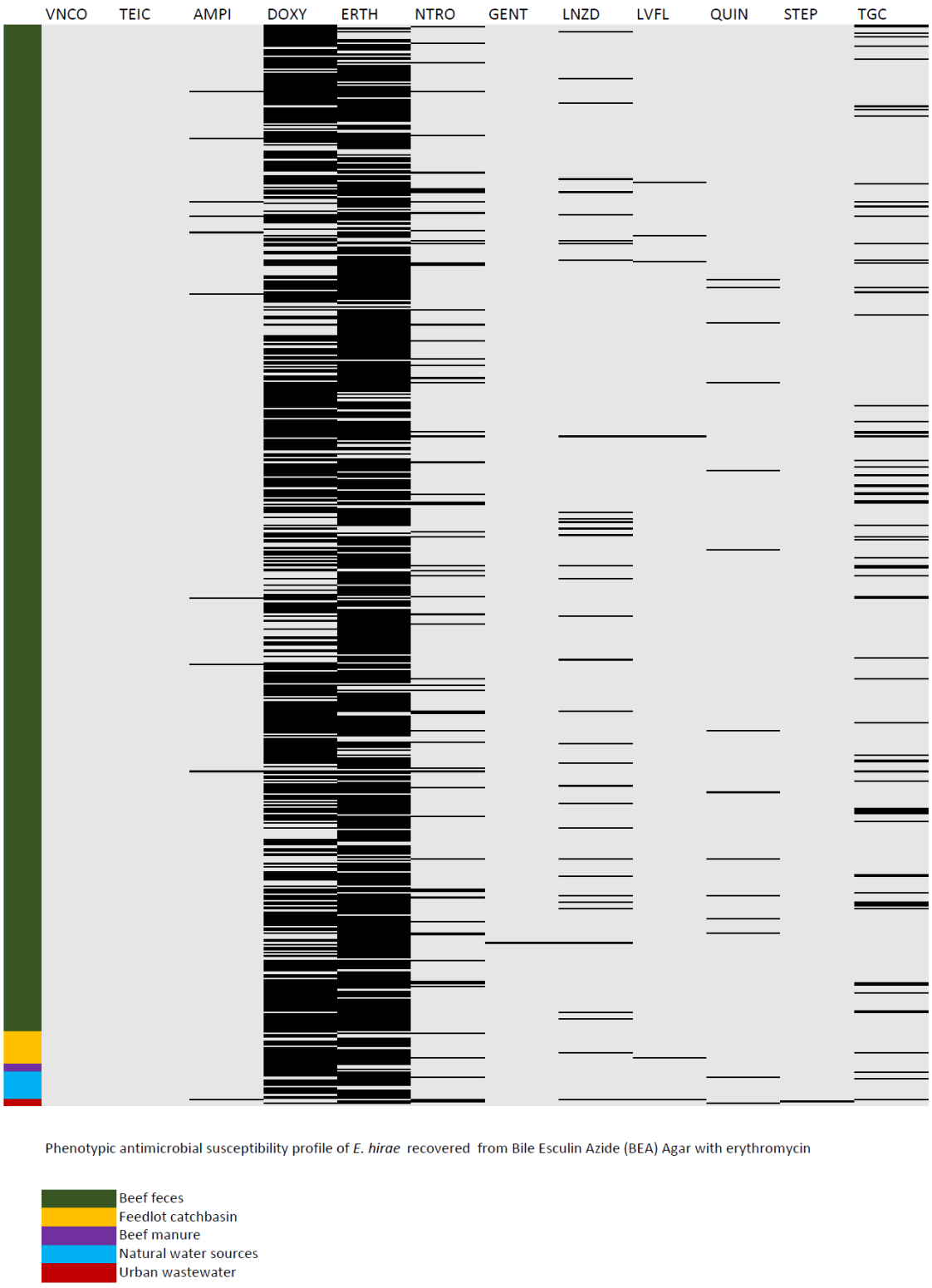

Figure S2. Heatmap of phenotypic resistance profile of *E. hirae* isolates (n=652), recovered from beef production system (n=632), natural water sources (n=17), and urban waste water (n= 3).

Table S1 Human clinical isolates retrieved from public database

Strain	Biosample	Source	SRA	Collection date	Geographical location
FDAARGOS_234	SAMN04875571	Human abscess	SRS1665862	.9/14/2014	USA:DC
708	SAMN15223906	Human	NA	2014	France: Jouy-en-Josas
13344	SAMN15223909	Human blood	NA	2012	France: Caen

Table S2 Resistant phenotypes identified in *E. hirae*

S.No.	Phenotype	Bovine feces	Catch basin	Bovine Manure	Natural water sources	Urban wastewater	Total
1	DOX+ERTH+NTRO+LNZD+LVFL+TGC	1	0	0	0	0	**1**
2	AMPI+DOX+ERTH+QUIN+TGC	1	0	0	0	0	**1**
3	AMPI+DOX+ERTH+NTRO+TGC	1	0	0	0	0	**1**
4	AMPI+NTRO+LNZD+LVFL+TGC	0	0	0	0	1	**1**
5	DOX+ERTH+LNZD+QUIN+TGC	1	0	0	0	0	**1**
6	DOX+ERTH+NTRO+LNZD+TGC	2	0	0	0	0	**2**
7	AMPI+DOX+ERTH+NTRO	3	0	0	0	0	**3**
8	AMPI+DOX+ERTH+TGC	2	0	0	0	0	**2**
9	AMPI+DOX+LNZD+QUIN	1	0	0	0	0	**1**
10	AMPI+ERTH+NTRO+TGC	1	0	0	0	0	**1**
11	DOX+ERTH+NTRO+LNZD	4	0	0	0	0	**4**
12	DOX+ERTH+NTRO+LVFL	0	1	0	0	0	**1**
13	DOX+ERTH+NTRO+QUIN	1	0	0	0	0	**1**
14	DOX+ERTH+NTRO+TGC	12	2	0	0	0	**14**
15	DOX+ERTH+LNZD+QUIN	1	0	0	0	0	**1**
16	DOX+ERTH+LNZD+TGC	7	1	0	0	0	**8**
17	DOX+ERTH+QUIN+TGC	1	0	0	0	0	**1**
18	DOX+ERTH+NTRO+QUIN	3	0	0	0	0	**3**
19	DOX+ERTH+NTRO+LVFL	0	1	0	0	0	**1**
20	DOX+NTRO+LNZD+TGC	3	1	0	0	0	**4**
21	NTRO+LNZD+LVFL+TGC	0	0	0	0	1	**1**
22	AMPI+DOX+ERTH	4	0	0	0	0	**4**
23	AMPI+DOX+TGC	1	0	0	0	0	**1**
24	DOX+ERTH+NTRO	31	0	0	1	0	**32**
25	DOX+ERTH+LNZD	18	0	0	0	0	**18**
26	DOX+ERTH+LVFL	3	0	0	0	0	**3**
27	DOX+ERTH+QUIN	4	1	0	0	1	**6**
28	DOX+ERTH+TGC	40	2	0	0	1	**43**
29	DOX+NTRO+LNZD	4	1	0	0	0	**5**
30	DOX+NTRO+TGC	15	1	0	0	0	**16**
31	DOX+LNZD+TGC	2	0	0	0	0	**2**
32	DOX+LNZD+QUIN	1	0	0	0	0	**1**
33	ERTH+GENT+LNZD	1	0	0	0	0	**1**
34	ERTH+LNZD+TGC	1	0	0	0	0	**1**
35	ERTH+NTRO+STREP	0	0	0	0	1	**1**
36	VNCO+ TEIC	1	0	0	0	0	**1**
37	VNCO + LNZD	0	0	0	0	1	**1**
38	AMPI+DOX	1	0	0	0	0	**1**
39	AMPI+ERTH	1	0	0	0	0	**1**
40	AMPI+TGC	1	0	0	0	0	**1**
41	AMPI+QUIN	1	0	0	0	0	**1**
42	DOX+ERTH	328	18	2	16	0	**364**
43	DOX+NTRO	30	4	1	4	0	**39**
44	DOX+LNZD	14	0	0	0	0	**14**
45	DOX+TGC	25	1	0	4	0	**30**
46	ERTH+NTRO	14	1	0	0	0	**15**

47	ERTH+LNZD	4	0	0	0	0	**4**
48	ERTH+QUIN	5	0	0	0	0	**5**
49	ERTH+TGC	10	1	0	2	0	**13**
50	NTRO+TGC	4	0	0	0	0	**4**
51	DOX	208	22	0	12	0	**242**
52	NTRO	24	2	0	0	1	**27**
3	LNZD	1	1	0	0	0	**2**
54	TGC	6	1	0	0	2	**9**
55	ERTH	183	5	0	3	0	**191**
56	Susceptible to all antibiotics	172	4	0	5	1	**182**
	Total	**1203**	**71**	**3**	**47**	**10**	**1334**

Table S3 Detailed assembly statistics

Assembly Statistics	
Assembly stats	**Average Value**
Min contig length	613
Max contig length	271379
Mean contig length	40316
N50 contig length	110141
Number of contigs	95 (26-235)
Number of contigs >=1kb	87
Number of contigs in N50	13 (04-42)
Number of bases in all contigs	2933987
Number of bases in contigs >=1kb	2927527
GC Content of contigs	36.70%

Table S4 Resistant genotypes identified in *E. hirae* isolates

S.no.	Resistant genotype	No. of isolates	Percentage %
1	*aac 6'-Iid*	12	4.2
2	*aac 6'-Iid + tet L*	47	16.4
3	*aac 6'-Iid + tetM*	6	2.1
4	*aac 6'-Iid + ermB*	5	1.7
5	*aac 6'-Iid + tet O/32/O*	4	1.4
6	*aac 6'-Iid + tetS/M*	2	0.7
7	*aac 6'-Iid + tet O*	1	0.3
8	*aac 6'-Iid + tet L + ermB*	87	30.4
9	*aac 6'-Iid + tetO/32/O +ermB*	9	3.1
10	*aac 6'-Iid + tetO + ermB*	8	2.7
11	*aac 6'-Iid + tetL + tetO*	4	1.4
12	*aac 6'-Iid + tet tetL + tetM*	3	1.0
13	*aac 6'-Iid + tetM + ermB*	1	0.3
14	*aac 6'-Iid + tetS/M + ermB*	1	0.3
15	*aac 6'-Iid + tetL + tetM + ermB*	64	22.3
16	*aac 6'-Iid + tetL + tetO + ermB*	26	9.1
17	*aac 6'-Iid + tetL + tetS/M + ermB*	3	1.0
18	*aac 6'-Iid+ tetL +tetO/32/O +ermB*	2	0.7
19	*aac 6'-Iid+ aph3-III + ant 6-la+ ermB*	1	0.3
		286	100

Following tables are submitted as external tables due to their size

Table S5 Phenotypic and genotypic data of *E. hirae* isolates

Table S6 Phages identified in *E. hirae* isolates through PHASTER analysis

Table S7 Genomic signatures identified through Neptune analysis (cut off score of 0.9)

Appendix 2- Supplementary for Chapter 3: Comparative genomic analysis of enterococci across sectors of the One Health continuum

Following tables are submitted as external tables due to their size

Table S1 Antimicrobial resistance genes of *E. faecium* isolates

Table S2 List of plasmids harboring antimicrobial resistance genes identified in 246 *E. faecium*

Table S3 Multilocus sequence type of 246 *E. faecium* genomes

Table S4 Virulence genes profiling of 246 *E. faecium* genomes

Table S5 Antimicrobial resistance genes profiles of *E. faecalis*

Table S6 List of plasmids harboring AMR genes identified in 376 *E. faecalis* genomes

Table S7 Multilocus sequence types of 376 *E. faecalis* genomes

Table S8 virulence genes profiling of 376 *E. faecalis* genome

Appendix 3- Supplementary for Chapter 4: Genomic characterization of carbapenem resistant bacteria from beef cattle feedlots clinical sources

Following tables are submitted as external tables due to their size

Table S1 Hybrid assemblies quality statistics determine using Quast

Table S2 List of *Pseudomonas aeruginosa* genomes retrieved from public database for analysis

Table S3 AMR gene profiles of *Pseudomonas aeruginosa* genomes retrieved from public database for analysis