Genomic profiling of *Escherichia coli* strains from food producing animals and animal derived food products in South Africa

by

REFILOE PETUNIA MALESA

STUDENT No: 17951704

Submitted in accordance with the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

LIFE SCIENCES

at the

UNIVERSITY OF SOUTH AFRICA

SUPERVISOR: Dr. I Matle

CO-SUPERVISOR: Dr. R Pierneef

DECLARATION

Name: Refiloe Malesa

Student number: 17951704

Degree: PhD in Life Sciences

Title: Genomic profiling of *Escherichia coli* strains from food producing animals and animal-derived food products in South Africa

I, Refiloe Petunia Malesa, hereby declare that the dissertation submitted for the Degree of PhD in Life Sciences at the University of South Africa is my original work. I further declare that this work has not previously been submitted to any other institution of higher education. I further declare that all sources that I have quoted or cited are indicated and acknowledged by means of a complete list of references.

DEDICATION

I dedicate this study to my mother, Puleng, and my late grandparents for their sacrifices over the years in making my education a priority. I also extend my heartfelt gratitude to my husband and children for their endless support and patience throughout my studies.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my heartfelt gratitude to God Almighty for granting me the strength, courage, and perseverance to complete this journey. I am deeply thankful for the knowledge and wisdom He continues to instill in me.

I wish to extend my sincere appreciation to my supervisors, Dr. Itumeleng Matle and Dr. Rian Pierneef, for conceptualizing this project and for their unwavering support and guidance throughout. Their patience in helping me navigate the subject matter, their openness to answer my questions, and their encouragement of my ideas have been invaluable. I am especially grateful for the time they dedicated to reviewing numerous drafts and for their constructive feedback, which greatly enhanced my writing and analytical skills.

A special thank you goes to Dr. Thendo Mafuna for his assistance with genomic analysis, as well as for his valuable inputs and the considerable efforts he invested in my work. I also wish to thank Dr. Kudakwashe Magwedere for his insightful contributions and expertise, which have significantly deepened my understanding of the field.

I am also thankful to the staff of the Agricultural Research Council-Onderstepoort Veterinary Research Feed and Food Laboratory for their meticulous work in processing the samples for this study. My heartfelt appreciation goes to the South African Department of Agriculture (Previously known as, the Department of Agriculture, Land Reform and Rural Development), and the University of South Africa for their financial support, which made this research possible.

Lastly, I would like to express my profound gratitude to my mother, Puleng, for prioritizing my education and inspiring me to strive for greater heights. To my family, Naphtali, Olerato, and Obonolo thank you for your unwavering support, love, and patience throughout my studies. Your encouragement has been my anchor throughout this journey.

ABSTRACT

Escherichia coli is one of the most extensively studied bacterial species worldwide due to its remarkable adaptability, genomic diversity, and zoonotic potential. The increasing prevalence of antimicrobial resistance (AMR) among *E. coli* strains represents a growing public-health concern, particularly in countries such as South Africa where the consumption of animal-derived food products is widespread. However, *E. coli* is not exclusively associated with foods of animal origin only; it is also commonly detected in non-animal-derived products such as fresh produce and water, highlighting the importance of investigating multiple sources. This study specifically focuses on isolates from animal-derived food products and does not attempt to give a full representation of *E. coli*'s ecological distribution within South Africa.

This study employed whole-genome sequencing (WGS) and comparative genomic approaches to characterize *E. coli* isolates from animals, meat, and meat products, focusing on phylogenetic assignment, virulence gene repertoire, mobile genetic elements, plasmid content, AMR determinants, and Type III Secretion System (T3SS) components. Isolates included in this study originated from food-producing animals and animal-derived products from species such as poultry, bovine, ovine and porcine, these products included processed meat, meat itself and faecal material from some of these species.

Genomic DNA was extracted from isolates confirmed as *E. coli* and further sequenced using Illumina sequencing platforms. This was followed by quality assessement of the sequences, genome assembly and annotation. In this study, downstream analysis included serotyping of these isolates, determination of antimicrobial resistance, virulence genes, plasmid detection, and characterisation of CRISPR-Cas system. Comparative genomic analysis in this study included phylogenetic analysis of South African strains with publicly available African genomes.

Based on WGS analysis of 35 isolates, five phylogroups (A, B1, B2, C, and E) were identified, with phylogroup A predominating (46%). Virulence profiling revealed that 69% of isolates were extra-intestinal pathogenic *E. coli* (ExPEC) and 31% intestinal pathogenic (InPEC), encompassing 23 serotypes and 19 sequence types (STs), the most common being ST1858, ST975, and ST10. A wide array of virulence and AMR genes, including those associated with transposons Tn2, IS26, and Tn6196, was detected, while the metal transport system resistance operon sitABCD was ubiquitous. Type I CRISPR systems were observed in all isolates, dominated by subtypes I-E (86%) and I-A (57%).

In-silico analysis of T3SS revealed that only three isolates (11%) carried LEE-encoded structural genes (escV, escN, escD, escU) and effector proteins (espA, espD, espG, eae). Structural modeling confirmed the stability and functional conservation of EscV and EspA, suggesting an intact and potentially functional T3SS apparatus in localized EPEC strains the first such report from South Africa. Comparative genomic analysis of 159 *E. coli* ST10 genomes from 19 African countries revealed extensive intra-lineage diversity, with 71% of isolates being multidrug resistant and 29% carrying *bla_CTX-M-15*. Plasmid replicon analysis identified widespread IncFIB and ColRNAI plasmids strongly associated with AMR determinants. Collectively, these findings highlight the evolutionary plasticity, virulence potential, and One Health significance of *E. coli* lineages circulating in Africa, underscoring the urgent need for integrated genomic surveillance and targeted interventions to mitigate foodborne and zoonotic transmission risks.

Key words: *Escherichia coli*, Whole-genome sequencing, Comparative genomics, Antimicrobial resistance, Multidrug resistance, Virulence factors, Type III secretion system, Plasmids, CRISPR–Cas systems, Sequence types, One Health, Time-scaled phylogeny

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AGP	Antibiotic growth promoters
AIDA-1	Adhesin involved in diffuse adherence 1
AMR	Antimicrobial resistance
ANT	Aminoglycoside nucleotidyltransferases
APEC	Avian pathogenic *Escherichia coli*
APH	Aminoglycoside phosphotransferases
API	Analytical Profile Index
ARG	Antibiotic resistance genes
BFP	Bundle-forming units
BHI	Brain heart infusion
CA	California
CDD	Conserved Domain Database
CF	Colonization factors
CGE	Centre of genomic epidemiology
cgMLST	Core-genome Multi locus sequence typing
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DEC	Diarrheagenic *Escherichia coli*
DHPS	Dihydropteroate synthase
DNA	Deoxyribonucleic acid
eae	*E. coli* attaching and effacing
EAEC	Enteroaggregative *Escherichia coli*
EFSA	European Food Safety Authority
EHEC	Enterohemolytic *Escherichia coli*
EIEC	Enteroinvasive *Escherichia coli*

EPEC	Enteropathogenic *Escherichia coli*
ERIC	Enterobacterial repetitive intergenic consensus
ESBL	Extended-spectrum beta-lactamase
ESC	*Escherichia coli* secretion
ETEC	Enterotoxigenic *Escherichia coli*
ExPEC	Extraintestinal pathogenic *Escherichia coli*
FAO	Food and Agriculture Organization
FDA	Food and Drug Administration
GLASS	Global Antimicrobial Resistance Laboratory & Response Network
GRAVY	Grand Average of Hydropathicity
HGT	Horizontal gene transfer
HMM	Hidden Markov model
HTS	High-throughput sequencing
HUS	Haemolytic uremic syndrome
INNUENDO	Integrated Rapid Infectious Disease Analysis
InPEC	Intestinal pathogenic *Escherichia coli*
IPAs	Invasion plasmid antigens
ISE	Insertion elements
LEE	Locus of enterocyte effacement
LMIC	Low- and middle-income countries
LT	Heat-labile toxin
MALDI-TOF	Matrix-Assisted Laser Desorption/Ionization Time-of-Flight
MDR	Multidrug resistant
MGE	Mobile genetic elements
MITEE	Miniature Inverted-repeat Transposable Element
MLST	Multi locus sequence typing

NCBI	National Centre for Biotechnology Information
NF-T3SS	Non-flagellar-T3SS
NGS	Next generation sequencing
NMEC	Neonatal meningitis *Escherichia coli*
OIE	Office International des Epizooties
ONT	Oxford Nanopore Technologies
PCR	Polymerase chain reaction
PFGE	Pulsed field gel electrophoresis
pH	Potential of hydrogen
RAPD	Random amplified polymorphism DNA
SA	South Africa
Sct	Secretion and cellular translocation
SEPEC	Sepsis-associated *Escherichia coli*
SMART	Simple Modular Architecture Research Tool
SNP	Single Nucleotide Polymorphism
SRSWOR	Simple random sampling without replacement
STa	Heat-stable enterotoxin a
STb	Heat-stable enterotoxin b
ST	Heat-stable toxin
STEC	Shiga toxin-producing *Escherichia coli*
SV	Structural variants
T1SS	Type one secretion system
T2SS	Type two secretions system
T3SS	Type three secretion system
T4SS	Type four secretions system
T6SS	Type six secretion system

TMD	Transmembrane domains
UPEC	Uropathogenic *Escherichia coli*
USA	United States of America
UTI	Urinary tract infections
VAG	Virulence-associated genes
VF	Virulence factors
VFDB	Virulence finder database
WGS	Whole genome sequencing
WHO	World health organisation
WOAH	World Organisation for Animal Health

RESEARCH PUBLICATION OUTCOMES

Publications

1. **Malesa, R**., Pierneef, R., Magwedere, K., Mafuna, T. and Matle, I. (2024). Genomic characterisation of generic *Escherichia coli* from food-producing animals and products of animal origin in South Africa. *Frontiers in Bacteriology*. 3, p.1432292. https://doi.org/10.3389/fbrio.2024.1432292
2. **Malesa, R.,** Pierneef, R., Mafuna, T., Magwedere, K., Seakamela, E. and Matle, I., 2025. Genome-Based In Silico Analysis of the Structural and Functional Characteristics of the Type Three Secretion System (T3SS) and Core Effector Proteins in Enteropathogenic *Escherichia coli* (EPEC) Strains Isolated from Food-Producing Animals and Products of Animal Origin. *Pathogens*, 14(11), p.1099. https://doi.org/10.3390/pathogens14111099
3. **Malesa, R**., Pierneef, R., Magwedere, K., Mafuna, T. and Matle, I. (2025). Shifting Genomes, Rising Threats: Adaptive Evolution and Antimicrobial Resistance of *E. coli* ST10 across African Ecologies. Manuscript number: GEN-D-25-01233, submitted to ScienceDirect, Genomics 06/08/2025. **Status: With Editor**.

Conference presentations

1. **21st Annual Congress SASVPEM Congress** held at Lagoon Beach Hotel, Cape Town on 21 to 23 August 2024: Virtual presentation.
2. **Biennial Congress of the South African Association of Veterinary Technologists** held at 26 Degrees South Muldersdrift, on 9-10 October 2024: Oral presentation.

CHAPTER 1

GENERAL INTRODUCTION

1.1 Background

Escherichia coli (*E. coli*) is a ubiquitous, Gram-negative, facultatively anaerobic, rod-shaped bacterium that inhabits diverse ecological niches. While commonly associated with the gastrointestinal tracts of humans and warm-blooded animals, its presence extends to extraintestinal environments such as soil, water and sewage (Wang et al., 2020; Pakbin, et al., 2021). Although most *E. coli* strains are commensal and integral to maintaining gut homeostasis, numerous pathotypes are responsible for severe infections in both humans and animals. These pathogenic strains pose a significant public health concern, contributing to considerable morbidity and mortality, economic burdens and disruptions in international food trade through foodborne illnesses. When it comes to animal-derived food products and humans, pathogenic *E. coli* often plays a significant role since its exposure occurs in these systems through contamination (Kirk et al., 2015; Almaary, 2023; Li et al., 2024).

Human infection is primarily facilitated by the ingestion of contaminated food or water, as well as direct contact with colonized or infected animals and their by-products (Bolukaoto et al., 2019; Ribeiro et al., 2023; Alhadlaq et al., 2024). Clinical outcomes are highly variable and strain-dependent, ranging from asymptomatic colonization and mild gastroenteritis to severe complications including hemolytic-uremic syndrome (HUS), neonatal meningitis, and urinary tract infections (UTIs) (Anueyiagu et al., 2024; Baruah, 2024). The treatment of such infections relies heavily on antibiotic therapy, however, the escalating prevalence of multidrug-resistant (MDR) strains has rendered many first-line agents including β-lactams, tetracyclines, and macrolides ineffective (Streicher, 2021). Consequently, treatment options are increasingly reliant on last-resort antibiotics such as carbapenems, colistin, and tigecycline (Poirel et al., 2018; Kocsis et al., 2022; Islam et al., 2024). The drivers of this resistance crisis are multifactorial and include the overuse and misuse of antibiotics in human medicine, veterinary practice, and agricultural production systems. In animal production systems, thorough use of antimicrobials often promotes emergence of multi-drug resistance which triggers transmission within the food chain (Van Boeckel et al., 2019; Zeng et al., 2023).

In this context, the One Health framework offers a critical lens for understanding the interconnectedness of human, animal, and environmental health, particularly regarding the emergence and dissemination of zoonotic pathogens and AMR determinants. In South Africa, where livestock farming underpins national food security and rural livelihoods, foodborne pathogens such as *E. coli* pose a substantial threat to public and veterinary health.

Notably, zoonotic *E. coli* pathotypes including Enteropathogenic *E. coli* (EPEC) and Shiga toxin-producing *E. coli* (STEC), particularly the Enterohemorrhagic *E. coli* (EHEC) subgroup has been implicated in numerous outbreaks linked to contaminated animal products (Abdalla et al., 2022; Pokharel et al., 2023). These strains are frequently characterized by MDR and an arsenal of virulence genes, increasing their potential for host colonization and disease progression (Moghadam et al., 2021; Wyrsch et al., 2022).

Central to the pathogenicity of EPEC and EHEC is the locus of enterocyte effacement (LEE), a horizontally acquired genomic island encoding the Type III Secretion System (T3SS). The T3SS is a specialized nanomachine that injects bacterial effector proteins such as Tir, EspF, EspG, Map, and EspZ—directly into host intestinal epithelial cells (Pinaud, Sansonetti and Phalipon, 2018). These effectors collectively manipulate host immune signaling, compromise cytoskeletal integrity, and disrupt epithelial barrier functions, facilitating intimate adherence and the formation of characteristic attaching and effacing (A/E) lesions (Ruano-Gallego et al., 2021; Sanchez-Garrido et al., 2022; Feng & Wang, 2024). The evolutionary acquisition of the LEE pathogenicity island is a critical virulence innovation that enables host adaptation and transmission within both human and animal populations. To assess the zoonotic risks posed by *E. coli*, it is important to understand the structural and functional characteristics of LEE-encoded T3SS components (Hazen et al., 2015; Wagner et al., 2018).

Of mounting concern is the global emergence of high-risk *E. coli* clones that simultaneously exhibit enhanced virulence and extensive antimicrobial resistance. Lineages such as ST131, a predominant extraintestinal pathogenic *E. coli* (ExPEC) clone, and STEC O157:H7 (ST11) have been responsible for recurrent outbreaks and are frequently associated with the carriage of extended-spectrum β-lactamases (ESBLs) and carbapenemase genes (Stoesser et al., 2016; Alqasim, 2021; Wyrsch et al., 2022). The co-occurrence of AMR and virulence factors in such lineages increases the likelihood of treatment failure and imposes significant challenges on infection control practices, especially in resource-limited settings (Bevan et al., 2017).

Whole-genome sequencing (WGS) has emerged as a transformative tool in microbial genomics, enabling high-resolution analysis of pathogen genomes. The WGS facilitates the in-silico prediction of serotypes, sequence types, AMR genes, virulence factors, and mobile genetic elements, as well as robust phylogenetic analysis for source attribution and outbreak investigation (Allard et al., 2018; Montso et al., 2022; Blankenship et al., 2023). The integration of WGS into AMR surveillance and food safety programs enhances our ability to detect emerging threats in real time and supports data-driven public health interventions. In South Africa, the uptake of WGS is steadily increasing, particularly in clinical microbiology laboratories; however, its application remains limited in the agricultural and environmental sectors (Aijuka et al., 2019; Mbanga et al., 2021; Kalule et al., 2024).

A critical gap persists in the genomic characterization of *E. coli* isolates from non-clinical sources including food-producing animals, wastewater, and animal-derived food products and wastewater. This paucity of genomic data significantly limits our understanding of strain diversity, resistance gene reservoirs, and interspecies transmission dynamics. The lack of comprehensive One Health genomic surveillance hampers our ability to quantify the zoonotic potential of MDR *E. coli* strains and undermines efforts to implement effective control measures (Mahfouz et al., 2018; Brockhurst et al., 2019; Denamur et al., 2021). Addressing this knowledge gap is essential for building a predictive framework to combat AMR and ensuring the health of interconnected human, animal, and environmental systems.

1.2 Rationale and significance of the study

Escherichia coli remains a globally significant zoonotic and foodborne pathogen, posing profound public health and socio-economic challenges, including in South Africa. In the South

African context, numerous studies have consistently identified meat and meat products particularly, poultry and beef, as major sources of *E. coli* contamination (Ateba & Mbewe, 2014; Alonso et al., 2017; Onyeka et al., 2020; Jaja et al., 2020). This concern is compounded by the frequent detection of highly virulent strains harbouring critical virulence genes such as *eae, stx1, stx2,* and *hlyA,* which are associated with severe clinical outcomes in humans, including hemorrhagic colitis and hemolytic-uremic syndrome (Madoroba et al., 2022). A notable example includes the rare but significant outbreak of Shiga toxin-producing *E. coli* (STEC) O157:H7 in South Africa in 1990, which underscored the epidemiological role of cattle as reservoirs of STEC (Gambushe et al., 2022; Malahlela et al., 2022).

Beyond virulence, the emergence and proliferation of MDR *E. coli* strains may further exacerbate the disease burden. Meat products that are improperly handled, processed, or cooked serve as critical vehicles for the transmission of these resistant strains from food animals to humans (Jaja et al. 2020). This growing threat has been closely linked to the widespread misuse and overuse of antibiotics in intensive animal farming systems, a key driver of AMR in both commensal and pathogenic *E. coli* strains (Abebe et al., 2020; Alara & Alara, 2024). The One Health implications of such resistance spanning clinical, veterinary, and environmental domains are particularly pronounced in low- and middle-income countries, where surveillance infrastructure remains limited.

Despite the gravity of these threats, several critical knowledge gaps persist that hinder effective mitigation strategies. First, research efforts in South Africa have largely concentrated on *E. coli* strains from human clinical settings, with limited attention directed toward animal reservoirs and food sources. This narrow focus has resulted in a fragmented understanding of the bacterium’s transmission dynamics and its zoonotic interface at the animal-food-human spectrum. Second, there is a marked deficiency in the application of WGS to *E. coli* isolates obtained from food-producing animals and animal-derived products. This limits our capacity to resolve phylogenetic relationships, characterize the accessory genome, and detect the presence of AMR and virulence genes with high resolution (Wee et al., 2020; Muloi et al., 2022). Third, South Africa currently lacks a publicly accessible, representative genomic repository for *E. coli* isolates originating from non-clinical sources. This absence hinders national efforts in real-time pathogen surveillance, outbreak source attribution, and evidence-based policymaking. Lastly, the local and transboundary movement of contaminated animal-derived food products often occurs under substandard hygienic conditions and weak regulatory enforcement further facilitating the persistence and dissemination of virulent and resistant *E. coli* strains (Olanya et al., 2019; Carroll et al., 2022).

In response to these critical gaps in genomic data from non-clinical isolates, this study aims to apply WGS to genetically characterize *E. coli* isolates recovered from food and animal sources in South Africa. By employing advanced genomic tools, the study will (i) elucidate the genomic diversity and evolutionary patterns of *E. coli* circulating at the animal-food interface, (ii) identify key virulence factors and AMR determinants that pose a risk to human health, and (iii) assess the potential for horizontal gene transfer through mobile genetic elements. Furthermore, the study seeks to establish a locally relevant genomic database that can support national AMR surveillance programs, enhance food safety risk assessments, and strengthen outbreak response capacity.

By bridging the gap between clinical and non-clinical *E. coli* research, this study will provide a different genomic evidence that will contribute significantly to South Africa's One Health AMR agenda. The genomic insights generated will inform targeted control measures, promote prudent antibiotic use in agriculture, and guide regulatory policy to curb the spread of MDR pathogens in the food chain. Ultimately, these outcomes strengthen national preparedness for foodborne threats by integrating animal, food, and public health perspectives into evidence-based decision-making.

1.3 Research aim and objectives

The overall aim of this study was to utilize WGS to conduct genomic characterization of *Escherichia coli* strains derived from food-producing animals, and animal-derived food products in South Africa, aligning with the principles of the One Health paradigm. The specific objectives of the study were as follows:

I. To determine the population structure of generic *E. coli* strains isolated from food-producing animals and animal-derived food products such as raw meat and processed meat products from various livestock in South Africa

II. To identify and profile virulence-associated genes, with a focus on the Type III Secretion System (T3SS), in *E. coli* strains isolated from food-producing animals, and animal-derived food products.

III. To undertake a WGS-based structural and functional characterization of the T3SS and its associated core effector proteins in Enteropathogenic *Escherichia coli* (EPEC) strains recovered from food-producing animals and animal-derived food products.

IV. To perform a Pan-African genomic surveillance to reveal the evolutionary plasticity and antimicrobial threat of *E. coli* ST10.

1.4 Research questions

1. What are the predominant pathotypes, serotypes, and sequence types of *E. coli* present in food-producing animal and animal-derived food products such as raw meat and processed meat products from various livestock in South Africa?
2. What are the most common virulence factors and antimicrobial resistance profiles characterized *E. coli* isolates from food-producing animals and animals-derived food products in South Africa?
3. How effectively can WGS be utilized to detect and characterize the T3SS in *E. coli* isolates from food and animal sources?
4. Which T3SS protein effectors are most prevalent in *E. coli* isolates from food and animals in South Africa?
5. How do the genomic characteristics of *E. coli* isolates from this study compare with those of isolates reported in other regions in Africa?

CHAPTER 2

LITERATURE REVIEW

2.1 General characteristics of *Escherichia coli*

Escherichia coli, a cornerstone model organism in microbiology, molecular genetics, and biotechnology, was first isolated in 1885 by Theodor Escherich from the stool of a healthy infant and was initially designated *Bacterium coli* to reflect its predominant niche in the colon. The species was later renamed *Escherichia coli* in 1919 in honour of its discoverer (Foster-Nyarko & Pallen, 2022). It belongs to the family *Enterobacteriaceae,* a diverse group of Gram-negative, facultatively anaerobic, non-spore-forming, rod-shaped bacteria. Members of this family are widely distributed in the environment and intestinal tracts of animals and humans and possess significant clinical and public health importance (Ruiz & Silhavy, 2022).

Morphologically, *E. coli* measures approximately 1-3 µm in length and 0.4-0.7 µm in diameter. Most strains are motile by means of peritrichous flagella, although non-motile variants exist, particularly among pathogenic and laboratory-adapted strains (Gupta & Gupta, 2021; Moi et al., 2022; Ruwandeepika et al., 2023). Its growth is supported across a wide temperature range (4-45°C), with optimal proliferation at 37°C, corresponding to the physiological temperature of its endothermic hosts. The organism exhibits a broad pH tolerance, optimally thriving under near-neutral conditions (pH 6.0-7.5) and demonstrates remarkable resilience in diverse environments due to its capacity to form biofilms and metabolize an extensive array of carbon substrates (Ziege et al., 2021; Meyers & McLellan, 2022).

From a genomic standpoint, *E. coli* is one of the most comprehensively studied prokaryotic organisms. Its genome size typically ranges between 4.5 and 5.5 megabases, with an average guanine-cytosine (G+C) content of ~50% (Wani et al., 2022). The *E. coli* pan-genome is highly plastic and dynamic, shaped by extensive horizontal gene transfer (HGT) events and genome rearrangements. This genomic fluidity underpins its capacity to acquire virulence-associated elements, AMR genes, and mobile genetic elements such as plasmids, transposons, integrons, and pathogenicity islands (Frost et al., 2005; Touchon et al., 2012). These genetic innovations contribute to the species' broad ecological range and clinical adaptability.

In its commensal state, *E. coli* constitutes a dominant member of the gut microbiota, playing essential roles in vitamin K2 synthesis, competitive exclusion of pathogens, and modulation of host immune responses (Bernreiter-hofer et al., 2021; Geurtsen et al., 2022). However, this mutualistic relationship is fragile; under certain conditions such as immune suppression, microbiota dysbiosis, or the acquisition of virulence genes, commensal *E. coli* may transition into an opportunistic pathogen.

Biochemically, *E. coli* exhibits a distinct metabolic fingerprint that facilitates its routine laboratory identification. It is catalase-positive, oxidase-negative, and capable of fermenting lactose, glucose, and mannitol with acid and gas production. It produces indole from tryptophan, is methyl red-positive, Voges-Proskauer-negative, and does not utilise citrate as the sole carbon source in Simmons' medium (Garrity et al., 2005). Additionally, it typically does not produce hydrogen sulfide in triple sugar iron agar. These biochemical characteristics form the basis of phenotypic identification using diagnostic systems such as API 20E and VITEK 2. In contrast, identification using Matrix-Assisted Laser Desorption/Ionization Time-of-Flight (MALDI-TOF) mass spectrometry is based on species-specific protein spectral fingerprints rather than biochemical or phenotypic traits (Janda & Abbott, 2021).

Furthermore, advances in phylogenomics and multi-locus sequence typing (MLST) have elucidated the population structure of *E. coli*, which includes eight major phylogroups (A, B1, B2, C, D, E, F, and G). Each phylogroup harbours unique combinations of virulence determinants, metabolic pathways, and ecological niches (Clermont et al., 2013; Denamur et al., 2021). For instance, phylogroup B2 is commonly associated with ExPEC, whereas B1 is more frequently linked to environmental and commensal strains. These phylogenomic insights have proven invaluable for One Health surveillance initiatives aimed at tracking the zoonotic and environmental transmission of pathogenic *E. coli* lineages.

The dual nature of *E. coli* as both a benign intestinal resident and a versatile pathogen exemplifies the complexity of host-microbe interactions and evolutionary adaptation. The species' tractability in genetic manipulation, rapid growth, and historical significance in biotechnology (e.g., recombinant protein production, CRISPR systems, and synthetic biology platforms) further establish *E. coli* as a keystone organism in life sciences research (Doudna & Charpentier, 2014; Blount, 2015).

Though characteristics of *E. coli* are well documented in literature, comparative data of this bacterium that addresses genomic diversity, antimicrobial resistance, virulence factors and other features of isolates from food-producing animals and animal-derived food products. The imbalance in research limits our overall understanding of *E. coli's* diversity, and in some studies, data is still limited on food products of aquatic origion (Allende et al., 2025, Touati et al., 2025).

2.2 *Escherichia coli* pathogenesis and pathogenicity

Escherichia coli pathogenesis is intricately linked to its ability to colonize and persist within the gastrointestinal tract of both humans and animals. Successful colonization requires the bacterium to overcome innate host defenses, including low pH, bile salts, antimicrobial peptides, and mucosal immunity, while securing essential nutrients for growth and replication (Jang et al., 2017; Ziege et al., 2021; Gieryńska et al., 2022). After colonization, *E. coli* traverses the mucus layer and may either engage in mucosal adherence or remain luminal. Irrespective of its pathogenic potential, *E. coli* is ultimately excreted into the intestinal lumen and reintroduced into the environment via faecal shedding a critical step for transmission and environmental persistence (Rogers et al., 2022). In the environment, the shedding of *E. coli* is the key ecological bridge found between animals as reservoirs, human exposure and food contamination especially where intensive farming is practiced (Dushayeva, 2025).

Pathogenic *E. coli* strains have evolved diverse virulence mechanisms, enabling them to exploit host tissues and circumvent immune responses. These strains are broadly categorized into intestinal pathogenic *E. coli* (IPEC) and ExPEC, with the former implicated in diarrhoeal diseases and the latter associated with systemic infections such as urinary tract infections

(UTIs), meningitis, and sepsis. Among IPEC, six major diarrhoeagenic pathotypes are recognized: enterotoxigenic (ETEC), enteropathogenic (EPEC), Shiga toxin-producing (STEC), enteroinvasive (EIEC), enteroaggregative (EAEC), and diffusely adherent *E. coli* (DAEC) (Denamur et al., 2021).

These pathotypes are defined by distinct virulence determinants and pathophysiological strategies. Enterotoxigenic *E. coli*, a leading cause of traveller's diarrhoea and infantile diarrhoea in low- and middle-income countries, colonizes the small intestine through colonization factors (CFs) and produces heat-labile (LT) and/or heat-stable (STa/STb) enterotoxins, encoded by *elt* and *est* genes respectively, which dysregulate electrolyte transport leading to secretory diarrhoea. Enteropathogenic *E. coli*, traditionally associated with infantile diarrhoea, initiates an attaching and effacing (A/E) lesion on intestinal epithelial cells via the locus of enterocyte effacement (LEE), which encodes a Type III Secretion System (T3SS), intimin (*eae*), and effector proteins such as EspB. Typical EPEC also carries the *bfpA* gene encoding the bundle-forming pilus (BFP), which is essential for localized adherence (Arnold et al., 2024).

Shiga toxin-producing, including the notorious serotype O157:H7, is responsible for bloody diarrhoea and HUS, a life-threatening complication characterized by acute renal failure. Shiga toxin-producing pathogenesis is mediated by phage-encoded Shiga toxins (*stx1*, *stx2*), which inhibit protein synthesis in host cells, and by LEE-encoded adherence factors. Enteroinvasive *E. coli* closely resembles *Shigella* spp. in pathogenesis and genome structure and causes a dysentery-like illness characterized by mucosal invasion, intracellular multiplication, and lateral intercellular spread. This is facilitated by the presence of the large virulence plasmid pINV, encoding T3SS machinery, and a set of invasion plasmid antigens (Ipas) (Kaur & Dudeja, 2023; Lee et al., 2024).

Enteroaggregative *E. coli* represents a genetically heterogeneous group of strains that form a characteristic "stacked-brick" aggregative adherence pattern on HEp-2 cells. It is increasingly implicated in persistent diarrhoea, especially in children, immunocompromised individuals, and travelers. Key virulence factors include the aggregative adherence fimbriae (AAF/I-IV), *aatA* (transporter), and *aaiC* (secreted protein), which promote biofilm formation, intestinal inflammation, and mucus hypersecretion (Llorente et al., 2023). Diffusely adherent *E. coli*, although less well-characterized, has been associated with both diarrhoea and UTIs and expresses adhesins encoded by *afa*, *dra*, and *daa* operons, facilitating diffuse adherence to epithelial cells. It also secretes the autotransporter toxin AIDA-I (Aijuka et al., 2019; Jesser & Levy, 2020).

Molecular epidemiological investigations leveraging WGS and virulome analysis have revealed considerable genetic plasticity and virulence diversity across *E. coli* pathotypes, often influenced by horizontal gene transfer via plasmids, integrative elements, and bacteriophages (Tetzschner et al., 2020). This genomic fluidity contributes to the emergence of hybrid strains exhibiting features of multiple pathotypes, such as enteroaggregative hemorrhagic *E. coli*, which was responsible for the large 2011 outbreak in Germany. Such convergent evolution underscores the significance of One Health surveillance approaches that integrate human, animal, and environmental data.

Although commensal *E. coli* strains generally lack classical virulence genes, they can become opportunistic pathogens in immunocompromised hosts or serve as reservoirs of AMR and virulence determinants. These genes can be mobilized under selective pressures such as antibiotic exposure, thereby blurring the distinction between commensalism and pathogenicity (Zhang et al., 2021; Mengistu & Mengesha, 2023). Consequently, deciphering the interplay between host immunity, microbiota composition, and bacterial genomics remains critical in understanding *E. coli* pathogenesis and guiding therapeutic strategies.

2.3 Management and treatment of *Escherichia coli* infections in humans

Infections caused by *E. coli* are clinically diverse, ranging from superficial gastrointestinal disturbances to life-threatening systemic illnesses. These include urinary tract infections (UTIs), gastroenteritis, neonatal meningitis, bacteremia, and wound infections. Clinical management is contingent on several factors, including the infecting pathotype, anatomical site of infection, immune status of the host, and the antimicrobial susceptibility profile of the isolate (Oladokun & Sharif, 2024). While mild gastrointestinal infections caused by diarrheagenic *E.*

coli are often self-limiting, extraintestinal infections, particularly those caused by ExPEC strains, typically necessitate targeted antimicrobial therapy.

2.4 Understanding Antimicrobial Resistance (AMR)

Antimicrobial resistance (AMR) arises when bacterial populations acquire or develop adaptive mechanisms that render them insensitive to previously effective antibiotics. In *E. coli,* AMR is largely driven by β-lactamase production—enzymes that hydrolyze the β-lactam ring, thereby inactivating penicillins, cephalosporins, and even last-line carbapenems (Pitout et al., 1998; Li et al., 2022). Among these enzymes, extended-spectrum β-lactamases (ESBLs), AmpC β-lactamases, and carbapenemases (e.g., NDM, KPC, OXA-48) have emerged as clinically significant contributors to therapeutic failure.

The genomic plasticity of *E. coli,* coupled with its broad ecological distribution across human, animal, and environmental reservoirs, facilitates inter- and intra-species gene transfer. This HGT occurs through well-characterized mechanisms such as conjugation, transformation, and transduction, with mobile genetic elements (MGEs) including plasmids, integrons, and transposons serving as critical vectors for the dissemination of AMR determinants (Poirel et al., 2018; Nelson et al., 2019; Ramadan, 2022). Notably, integron-associated gene cassettes often encode resistance to multiple antibiotic classes simultaneously, accelerating the emergence of MDR and XDR strains.

Conjugative plasmids self-transmissible extrachromosomal DNA elements play a particularly influential role in AMR propagation. ESBL- and β-lactamase (BL)-producing *E. coli* strains frequently harbour IncF, IncI, or IncX-type plasmids, which enable efficient gene transfer within and across bacterial species. The co-localization of resistance genes with virulence factors and heavy metal resistance determinants on the same plasmids further complicates treatment and containment strategies. This convergence increases fitness under diverse selective pressures, including antibiotic exposure and metal-rich agricultural environments.

The World Health Organization (WHO) has classified ESBL-producing and carbapenem-resistant *E. coli* among its top-tier "critical priority pathogens" due to their limited treatment options and pandemic potential (Darphorn et al., 2021). Their inclusion in this global research agenda reflects a pressing need for novel antimicrobials, rapid diagnostic tools, and coordinated One Health strategies that address AMR holistically across human health, veterinary medicine, agriculture, and environmental sectors. Ultimately, addressing AMR in *E. coli* demands a multi-pronged approach: reducing inappropriate antibiotic usage, strengthening surveillance systems, investing in innovative therapeutics, and enforcing rigorous hygiene and sanitation protocols across the farm-to-fork continuum (Saeed et al., 2023; Bhandari et al., 2024; Alhassan et al., 2025).

2.4.1 Commonly used antibiotics and emerging resistance patterns

The therapeutic management of *E. coli* infections traditionally relies on a range of antimicrobial classes, including β-lactams (such as penicillins, cephalosporins, and carbapenems), fluoroquinolones, aminoglycosides, sulphonamides, tetracyclines, polymyxins, and macrolides (Urban-Chmiel et al., 2022). Among these, β-lactams and fluoroquinolones remain cornerstone agents due to their proven efficacy, favourable pharmacokinetic profiles, and broad-spectrum activity. Nevertheless, the global clinical utility of these antibiotics is being eroded by the alarming and accelerating emergence of MDR *E. coli* strains a trend largely fuelled by the indiscriminate and unregulated use of antimicrobials in both human and veterinary medicine (Roth et al., 2019).

Resistance patterns exhibit pronounced geographic variability, often reflecting local antimicrobial usage practices and infection control capacities. For example, fluoroquinolone resistance among *E. coli* strains implicated in urinary tract infections in India has been reported to exceed 60%, posing serious limitations for empiric therapy (Kumar et al., 2021). In Egypt, resistance to third-generation cephalosporins is widespread, primarily mediated by the dissemination of extended-spectrum β-lactamases (ESBLs), particularly those encoded by *bla_CTX-M* variants (Abdelrahim et al., 2024). In Brazil, the identification of polymyxin-resistant *E. coli* strains carrying the plasmid-mediated *mcr-1* gene in both clinical and animal isolates underscores the transboundary threat posed by mobile resistance determinants (Fernandes et al., 2020).

A similarly concerning picture is emerging from China, where *E. coli* strains have been found to co-harbour multiple high-risk resistance genes, including *blaNDM* (New Delhi metallo-β-lactamase), *blaKPC* (*Klebsiella pneumoniae* carbapenemase), and *blaCTX-M*, resulting in extensively drug-resistant (XDR) phenotypes with limited treatment options (Zhou et al., 2020). These findings highlight the convergence of critical resistance mechanisms on conjugative plasmids, facilitating their rapid horizontal dissemination and interspecies transfer.

In South Africa, resistance profiles largely mirror these global patterns yet are compounded by local ecological pressures. Molecular surveillance studies have shown that *E. coli* isolates from wastewater, food-producing environments, and clinical settings frequently harbour class 1 and class 2 integrons genetic elements that act as assembly platforms for antimicrobial resistance gene cassettes (Mbanga et al., 2021). Moreover, healthcare-associated infections increasingly involve carbapenemase-producing strains, notably those expressing *blaNDM*, *blaOXA-48*, and *blaKPC*, which significantly compromise treatment outcomes and infection control efforts (Sekyere, 2016).

These regional and global resistance dynamics underscore the critical need for integrated, cross-sectoral AMR surveillance. A One Health framework which recognizes the interconnectedness of human health, animal health, and the environment is essential for the effective monitoring, control and mitigation of *E. coli*-associated AMR. Without coordinated efforts across disciplines and borders, the proliferation of MDR and XDR *E. coli* strains may render common infections untreatable, posing a formidable threat to public health and global biosecurity.

Table 1: Antibiotics grouped by their mechanism of action

Mechanism	Antibiotic Class	Selected Examples
Inhibitors of cell wall synthesis (β-lactams)	Penicillins	Penicillin G, Methicillin, Ampicillin
	Cephalosporins	Cefalotin, Cefaloxin, Cefotan
	Carbapenems	Ertapenem, Meropenem
	Monobactams	Aztreonam
	Others	Vancomycin, Bacitracin
Disruptors of cell membranes	Polymyxins	Polymyxin B, Polymyxin E
Nucleic acid synthesis inhibitors	Fluoroquinolones	Nalidixic Acid, Ciprofloxacin
	Rifamycins	Rifampin
Protein synthesis inhibitors	Aminoglycosides (30S)	Gentamicin, Streptomycin
	Tetracyclines	Tetracycline, Doxycycline
	Macrolides (50S)	Erythromycin, Clarithromycin
	Others	Chloramphenicol, Clindamycin
Folic acid synthesis inhibitors	Sulphonamides/Trimethoprim	Sulfamethoxazole, Trimethoprim
	Others	Pyrimethamine
Mycolic acid synthesis inhibitors	Anti-TB agents	Isoniazid

2.5 The role of food and livestock systems in the dissemination antimicrobial resistance spread to animals

Environmental contamination acts as a key reservoir and amplifier of AMR. Suboptimal hygiene and biosecurity practices in intensive poultry production systems contribute to persistent *E. coli* colonization. Studies have shown that incomplete disinfection of hatcheries, feed processing facilities, and water delivery systems allows resistant *E. coli* strains to persist and recolonize chicks during early life stages (Swelum et al., 2021; Yang et al., 2021).

Alarmingly, viable *E. coli* can be recovered even after feed pelleting processes, highlighting the inadequacy of thermal inactivation in some contexts. These environments serve as “AMR incubators,” enabling the accumulation, maintenance, and dissemination of resistance genes through direct contact, faecal contamination, or shared microbial niches.

The emergence and spread of AMR are a pressing global health crisis that transcends clinical boundaries and implicates food production systems as significant reservoirs and transmission pathways (Jian et al., 2021). The gastrointestinal tract of animals constitutes a dynamic microbial ecosystem where HGT events frequently occur, enabling the rapid dissemination of ARGs among commensal and pathogenic bacteria (Wei et al., 2021; Michaelis & Grohmann, 2023). Resistance may arise either through chromosomal mutations or, more frequently, via acquisition of ARGs on MGEs such as plasmids, integrons, and transposons.

As global food demand intensifies with population growth, livestock production has become increasingly intensified and industrialized (Dadgostar, 2019). While such systems are essential for ensuring food security, they also contribute to the amplification and persistence of AMR within the animal microbiota, especially when prophylactic or subtherapeutic antibiotics are used to enhance productivity or prevent disease. A critical environmental consequence of intensive livestock production is the generation of large volumes of manure, which, while nutrient-rich and widely used as fertilizer, also harbours significant loads of antibiotic residues, ARGs, MGEs, and heavy metals (Qian et al., 2018; Lima et al., 2020; Montso et al., 2022).

Heavy metals such as copper and zinc frequently added to animal feed to promote growth exert co-selection pressure, as many MGEs carry both metal resistance genes and ARGs. This co-localization enhances the environmental stability and persistence of resistance determinants, with studies demonstrating ARG retention in manure for over 5-6 months post-excretion (Lima et al., 2020). Consequently, land application of untreated or poorly treated manure may lead to the contamination of agricultural soils and water systems, facilitating ARG entry into crops and the broader food chain.

The transmission of AMR through food systems is a multifaceted process involving several interconnected pathways. It can occur through direct contact with colonized animals or contaminated carcasses during slaughter and handling, as well as through the ingestion of raw or undercooked animal products that harbour resistant bacteria. Additionally, environmental exposure plays a role, particularly via irrigation with wastewater or runoff contaminated with manure. The international trade of food products adds another layer of complexity by enabling the cross-border movement of AMR determinants, thus facilitating their global dissemination (Woolhouse et al., 2015; Nelson et al., 2019; de Mesquita Souza Saraiva et al., 2021; Koutsoumanis et al., 2022; Nastasijevic et al., 2023).

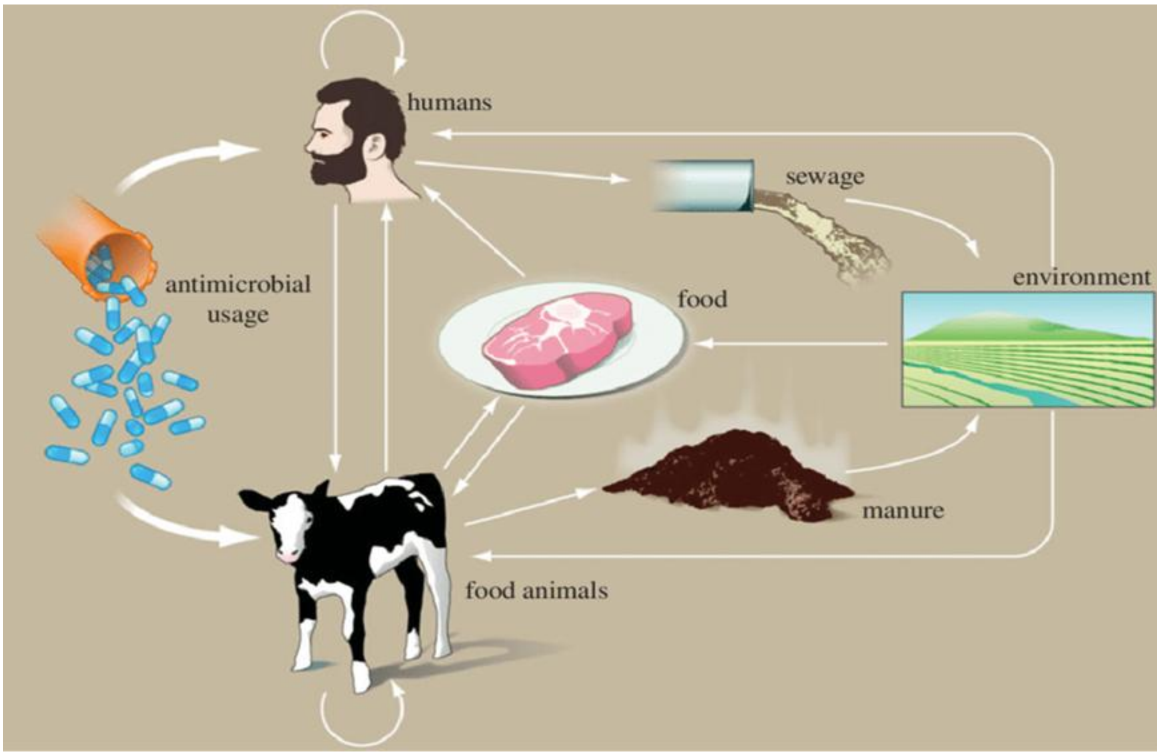


Figure 1: Schematic representation of the transmission pathways of antimicrobial resistant *Escherichia coli* and resistance genes between animals, humans, food, and the environment (Woolhouse et al., 2015).

Despite the inherent challenges in tracing precise transmission routes, systematic reviews suggest that AMR transfer from animals to humans occurs frequently. A meta-analysis of studies published between 1940 and 2016 revealed that 18% (n = 8) of studies identified bidirectional AMR transfer between food-producing animals and humans, and 56% (n = 25) showed unidirectional transmission from animals to humans (Muloi et al., 2018). While methodological constraints limit definitive conclusions regarding directionality, these findings provide strong support for the One Health paradigm, which emphasizes the interdependence of human, animal, and environmental health. South Africa and other low- and middle-income countries (LMICs) often face additional challenges in enforcing biosafety and biosecurity measures at abattoirs, farms, and processing facilities. Inconsistent regulatory oversight, inadequate sanitation, and limited access to routine AMR surveillance exacerbate the risk of foodborne AMR transmission in these contexts (Gambushe et al., 2022).

Escherichia coli has emerged as a sentinel organism for AMR monitoring due to its ubiquity in the gastrointestinal tracts of both humans and animals, as well as its capacity to acquire and disseminate ARGs across ecological boundaries (Ahmad et al., 2023; Sabença et al., 2024). Its dual role as a commensal and a pathogen, combined with its genetic plasticity and ease of detection, makes it an ideal model organism for assessing the impact of food production systems on AMR ecology.

2.6 Trends of antimicrobial-resistant *Escherichia coli* in South Africa

South Africa, like many LMICs, faces a growing challenge in managing AMR, particularly in *E. coli*, a pathogen frequently implicated in both community and healthcare-associated infections. However, the majority of AMR surveillance and research efforts in the country have historically focused on human clinical settings, contributing to a substantial knowledge gap regarding environmental and agricultural drivers of resistance (Fatoba et al., 2022).

Agricultural antibiotic use, particularly in the poultry and pork industries, constitutes a major source of selective pressure for AMR. These sectors are among the highest consumers of antibiotics, often using antimicrobials for growth promotion and disease prophylaxis. The agricultural sector is the biggest contributor to maintenance and dissemination of resistant *E. coli* strains at the animal, food and human interface. Surveillance along the "farm-to-fork" continuum has revealed that over 60% of *E. coli* isolates from poultry and swine environments harbour resistance to multiple antibiotic classes, including β-lactams, sulphonamides, and tetracyclines (Founou et al., 2018; Abdalla et al., 2021; Ramtahal et al., 2022). The prevalence of resistance is exacerbated by the widespread use of animal waste as fertilizer, often without adequate treatment, facilitating the environmental dissemination of resistant bacteria and resistance genes.

The environmental interface particularly wastewater treatment systems plays a crucial role in the ecology of AMR. In South Africa, there are approximately 986 municipal wastewater treatment facilities, of which at least 46% discharge untreated or partially treated sewage into rivers and streams (Omar & Barnard, 2010; Igwaran et al., 2018, Jiyane, Hoorzook and van Rensburg, 2025). These inadequacies allow the survival and propagation of resistant *E. coli* strains, including diarrheagenic pathotypes, in surface waters. Notably, several studies have isolated multidrug-resistant *E. coli* from surface water bodies used by rural and peri-urban communities for domestic purposes, raising significant public health concerns (Mbanga et al., 2021; Baloyi et al., 2022).

Compounding this issue is the documented presence of resistant *E. coli* strains on fresh produce sold at informal markets. Contamination may occur via irrigation with untreated wastewater or handling by colonized individuals, underscoring the importance of hygiene practices and water safety in food systems (Bell et al., 2021). These environmental and foodborne pathways suggest the urgent need to expand AMR surveillance beyond hospital settings into the agricultural and ecological domains. Of concern is the rising resistance to third-generation cephalosporins such as ceftriaxone and ceftazidime and carbapenems, including imipenem and ertapenem, in both clinical and non-clinical *E. coli* isolates (Jaja, et al., 2020). Interestingly, cephalosporins are not commonly used in South African veterinary medicine, suggesting that the observed resistance is more likely due to selective pressure from clinical usage in human healthcare. The detection of carbapenem-resistant *E. coli* in environmental and animal sources further points to cross-sectoral dissemination, possibly facilitated by plasmid-mediated resistance genes such as *blaNDM* and *blaOXA-48*.

2.7 Managing and controlling antimicrobial resistance: A One Health perspective

The widespread and often indiscriminate use of antibiotics in human medicine, animal husbandry, and agricultural practices has led to a global surge in AMR, now recognized by the WHO as one of the top 10 threats to global public health. The consequences are particularly acute in LMICs, where weak regulatory enforcement, limited surveillance infrastructure, and over-the-counter antibiotic access exacerbate the emergence and transmission of resistant pathogens (Caniça et al., 2019).

In the context of animal agriculture, rising global food demand has accelerated the use of antibiotics as both therapeutic and non-therapeutic agents. Antibiotic growth promoters (AGPs), administered at subtherapeutic doses, have historically been used in food-producing animals to enhance feed conversion efficiency, promote growth, and reduce disease incidence in intensive production systems. While AGPs have improved productivity, a growing body of evidence indicates that their continued use contributes significantly to the selection and dissemination of ARGs among zoonotic and commensal bacteria (Abdalla et al., 2021).

Recognizing these risks, the European Union implemented a ban on the use of AGPs in livestock production as early as 2006. Though these regulatory interventions demonstrate success in reducing the use of antibiotics, their effectiveness in low-middle income countries is often restricted by economic and infrastructural challenges. This policy shift marked a critical milestone in the global fight against AMR and inspired similar regulatory actions in other regions (Abdalla et al., 2021). However, implementation of such bans remains inconsistent worldwide, particularly in countries lacking alternative disease control strategies, economic incentives for compliance, or awareness of the long-term consequences of AMR.

To guide rational antimicrobial use across sectors, both the WHO and the World Organisation for Animal Health (WOAH, formerly OIE) have developed comprehensive frameworks such as the Global Action Plan on AMR and the List of Critically Important Antimicrobials for Human Medicine, which categorizes antimicrobials based on their therapeutic value and potential for resistance development. These guidelines promote prudent antimicrobial stewardship and advocate for the restriction of certain critical drugs in veterinary medicine when they are also vital for human health (Gehring, Mochel and Schmerold, 2023; Robbins et al., 2024).

Among the most overused antibiotics with high levels of resistance observed in *E. coli* and other bacterial species are tetracycline, chloramphenicol, sulfamethoxazole, and ampicillin

(Zhang et al., 2017; Abdalla et al., 2021). These agents, while historically effective, are now frequently compromised by widespread ARGs such as *tetA*, *catA1*, *sul1/sul2*, and *blaTEM*. These genes are often embedded in MGEs, enhancing their potential for cross-species and cross-environmental dissemination.

A holistic, One Health-based approach is required to mitigate AMR effectively. Key interventions include which includes strengthening the AMR Surveillance Systems through robust, harmonized surveillance of antimicrobial use and resistance trends across the human–animal–environment interface is critical. This includes genomic surveillance to detect emerging resistance genes and track their spread. Implementing Antimicrobial Stewardship Programs through stewardship efforts in hospitals, veterinary clinics, and farms can optimize antibiotic prescribing practices, reduce misuse, and ensure treatment guidelines are aligned with resistance profiles. Phasing out non-therapeutic use of antibiotics in agriculture and encouraging the transition away from AGPs through policy reform, education, and incentives for sustainable farming practices can significantly reduce selection pressure.

Furthermore, promoting alternatives to antibiotics such as vaccines, probiotics, prebiotics, phage therapy, and enhanced biosecurity measures to reduce reliance on antimicrobials for disease prevention. Educating and building capacity through training stakeholders from healthcare workers to farmers and policymakers on the risks of AMR and responsible antibiotic use is essential to ensure long-term behavioural and systemic change. It is also important that international collaboration and policy harmonization be introduced because AMR does not respect borders, regional and global cooperation is vital. Initiatives such as the Global Antimicrobial Resistance Surveillance System (GLASS) and the Tripartite AMR partnership between WHO, WOAH, and FAO aim to promote coordinated responses.

Ultimately, addressing AMR requires not only scientific innovation but also political will, economic investment, and public engagement. Without decisive, cross-sectoral action, the world risks entering a post-antibiotic era where even common infections may become untreatable.

2.8 Genomic diversity of *Escherichia coli*

Escherichia coli exhibits remarkable genomic plasticity, which underpins its ecological versatility and pathogenic diversity. Its genomic flexibility has allowed it to easily adapt to diverse environments (Saini et al., 2024). The first complete genome sequence of the laboratory strain *E. coli* K-12 MG1655, published by (Lawrence & Ochman, 1998), laid the foundation for understanding the evolutionary mechanisms driving HGT, genome evolution, and strain differentiation. Since then, comparative genomics has revealed that *E. coli* harbours a core genome of approximately 1,400–1,500 conserved genes and a pan-genome comprising an ever-expanding repertoire of accessory genes, whose total number is considered virtually infinite due to the organism's constant acquisition of MGEs and niche-specific genes.

The substantial gap between core and pan-genome sizes highlights *E. coli*'s extraordinary genomic diversity, which is shaped by factors such as ecological source, host type, and selective pressures. For instance, isolates recovered from wastewater environments often containing a mix of commensal, pathogenic, and environmental strains tend to exhibit higher genomic variability than those from clinical settings (Mahfouz et al., 2018). This diversity complicates outbreak investigations and calls for high-resolution molecular tools to track the evolution and transmission of AMR and virulence traits.

Several molecular typing techniques have been employed to characterize the genetic diversity of *E. coli* populations. Among these are Multilocus Sequence typing (MLST), Pulsed-field gel electrophoresis (PFGE), Random Amplified Polymorphic DNA (RAPD), Enterobacterial Repetitive Intergenic Consensus (ERIC) PCR and Whole genome sequencing.

Multilocus Sequence typing is a robust and reproducible method that has sequences of internal fragments of seven housekeeping genes (~450–500 bp each). It assigns unique allele numbers to each gene and therefore allocates the corresponding sequence type (ST) to a unique allelic profile. MLST has become a gold standard in epidemiological surveillance, enabling unambiguous strain tracking and phylogenetic analyses through platforms such as PubMLST and EnteroBase (Castellanos et al., 2017; Ramadan et al., 2020).

Pulsed-Field Gel Electrophoresis (PFGE) is widely used for bacterial subtyping in foodborne outbreak investigations, PFGE resolves large DNA fragments generated by restriction enzyme digestion, allowing differentiation between closely related isolates. Despite its discriminatory power, PFGE is labor-intensive and less portable across laboratories compared to sequence-based methods (Im et al., 2021).

Random Amplified Polymorphic DNA (RAPD) and Enterobacterial Repetitive Intergenic Consensus (ERIC)-PCR, these fingerprinting techniques amplify variable regions in bacterial genomes and are cost-effective for preliminary screening but are less standardized and reproducible compared to MLST and whole genome sequencing (WGS) (Bumunang et al., 2019; Montso et al., 2022).

Whole Genome Sequencing (WGS) has emerged as the most powerful tool for strain-level characterization, enabling comprehensive analysis of single-nucleotidepolymorphisms (SNPs), AMR determinants, virulence factors, plasmid profiles, and genomic islands. The integration of WGS into surveillance frameworks enhances resolution and supports outbreak tracing, source attribution, and risk assessment (Stratakos et al., 2020; Montso et al., 2022).

2.8.1 Serotyping and surface antigens

Serotyping remains a valuable classical method for differentiating *E. coli* strains, especially in the context of pathogenicity. It is based on the identification of four key surface antigens. The O-antigen (somatic) is located on the outer membrane lipopolysaccharide (LPS), the O-antigen plays a central role in immune evasion and virulence. Genes responsible for its biosynthesis are flanked by the housekeeping genes *galF* and *gnd*. Loss or mutation in O-antigen biosynthesis can attenuate virulence (Liu et al., 2020). The H-antigen (flagellar) is an encoded by *fliC*, the H-antigen determines motility and contributes to serotype designation (e.g., O157:H7).

The capsular antigen (K-antigen), are polysaccharide capsules that are classified into four groups (I–IV) and contribute to serum resistance and immune evasion. The fimbrial (F-antigen) are less commonly used in routine serotyping, but it contributes to adhesion and colonization, particularly in uropathogenic strains, through its fimbriae. Serotyping is particularly important in distinguishing diarrheagenic *E. coli* (DEC) pathotypes such as EHEC O157:H7 and EPEC O111:H2, which are associated with specific virulence gene profiles and clinical syndromes.

2.8.2 Implications for public health and One Health surveillance

The genomic diversity study of *E. coli* is not merely academic, it has profound implications for public health, food safety, and environmental monitoring. Pathogenic lineages such as ST131, ST10, and ST69 exhibit distinct resistance and virulence profiles that are increasingly linked to zoonotic reservoirs and foodborne transmission. In South Africa and other LMICs, where *E. coli* isolates from wastewater, animal products, and produce show high genomic variability and multidrug resistance, molecular surveillance is essential for tracking inter-sectoral AMR transmission. Incorporating WGS-based genotyping into national AMR surveillance programs enhances early warning systems, improves risk assessment models, and supports the design of targeted interventions to mitigate the spread of high-risk clones across the One Health spectrum.

2.9 Overview of bacterial secretion systems

Bacterial secretion systems are complex molecular machineries that mediate the transport of a wide array of substrates including proteins, DNA, and small molecules across cell membranes. These systems are essential for bacterial survival, niche adaptation, host colonization, and, in the case of pathogenic bacteria, virulence expression. The nature and function of secreted substrates vary across environments, bacterial species, and secretion systems, playing roles in adhesion, biofilm formation, immune evasion, and direct host cell manipulation (Makarova et al., 2016; Grossman et al., 2021).

Pathogenic bacteria exploit secretion systems to deliver effector molecules that manipulate host cell processes, often subverting innate immunity and promoting intracellular survival. The ability to invade, persist, and damage mammalian hosts via protein secretion is a fundamental hallmark of bacterial virulence (Green & Mecsas, 2016). To date, eleven types of bacterial secretion systems (T1SS–T11SS) have been characterized in Gram-negative and Gram-positive bacteria, each differing in structure, mechanism, and substrate specificity (Rapisarda & Fronzes, 2018). These systems can secrete enzymes such as proteases, lipases, amidases, adhesins, or toxins, and in some cases, they facilitate direct injection of virulence proteins into host cells.

2.9.1 Functional architecture and mechanistic diversity

Each secretion system is adapted to its biological niche whereby T1SS, T2SS, T4SS, and T6SS utilize homologous hexameric AAA+ ATPases for energy transduction and are involved in a broad range of functions from toxin secretion to DNA exchange (Milne-Davies et al., 2021). T3SS (injectisome) and flagellar secretion systems rely on cytosolic hexameric ATPases structurally and functionally reminiscent of the F_0F_1 ATP synthase complex, highlighting evolutionary conservation among energy-dependent export systems (Milne-Davies et al., 2021). T7SS–T11SS, though less well-characterized in *E. coli*, demonstrate further specialization in Gram-positive bacteria and mycobacteria. The systems may secrete substrates across the inner membrane, outer membrane, or directly into host cells depending on system type. For example, T3SS, T4SS, and T6SS function as injectisomes capable of directly translocating effectors into the cytosol of host cells.

2.9.2 The Type III Secretion System: A key pathogenicity factor in *Escherichia coli*

Among the best-studied secretion systems is the Type III Secretion System (T3SS), particularly important in enteropathogenic EPEC and EHEC. T3SSs are needle-like structures composed of ~20–30 conserved membrane-associated proteins, forming a basal body, inner membrane rings, and a hollow needle that spans bacterial and host membranes (Hueck, 1998; Wagner et al., 2018) (Figure 2).

T3SSs are activated upon contact with host cell surfaces and facilitate the translocation of unfolded effector proteins directly into host cytoplasm. Effector secretion is directed by an N-terminal signal sequence, often with the aid of specific chaperones that maintain proteins in an unfolded state, ensuring passage through the narrow channel of the injectisome (Wagner et al., 2018). Despite the diversity of effector structures and functions ranging from cytoskeletal rearrangement to apoptotic inhibition, the T3SS apparatus itself is remarkably conserved across many pathogens, including *Salmonella*, *Shigella*, *Yersinia*, and *E. coli*. Pathogenicity islands such as the locus of enterocyte effacement (LEE) in EPEC and EHEC encode the structural genes of T3SS along with key effectors like EspB, EspF, Tir, and intimin, which together mediate intimate attachment and actin pedestal formation on host enterocytes.

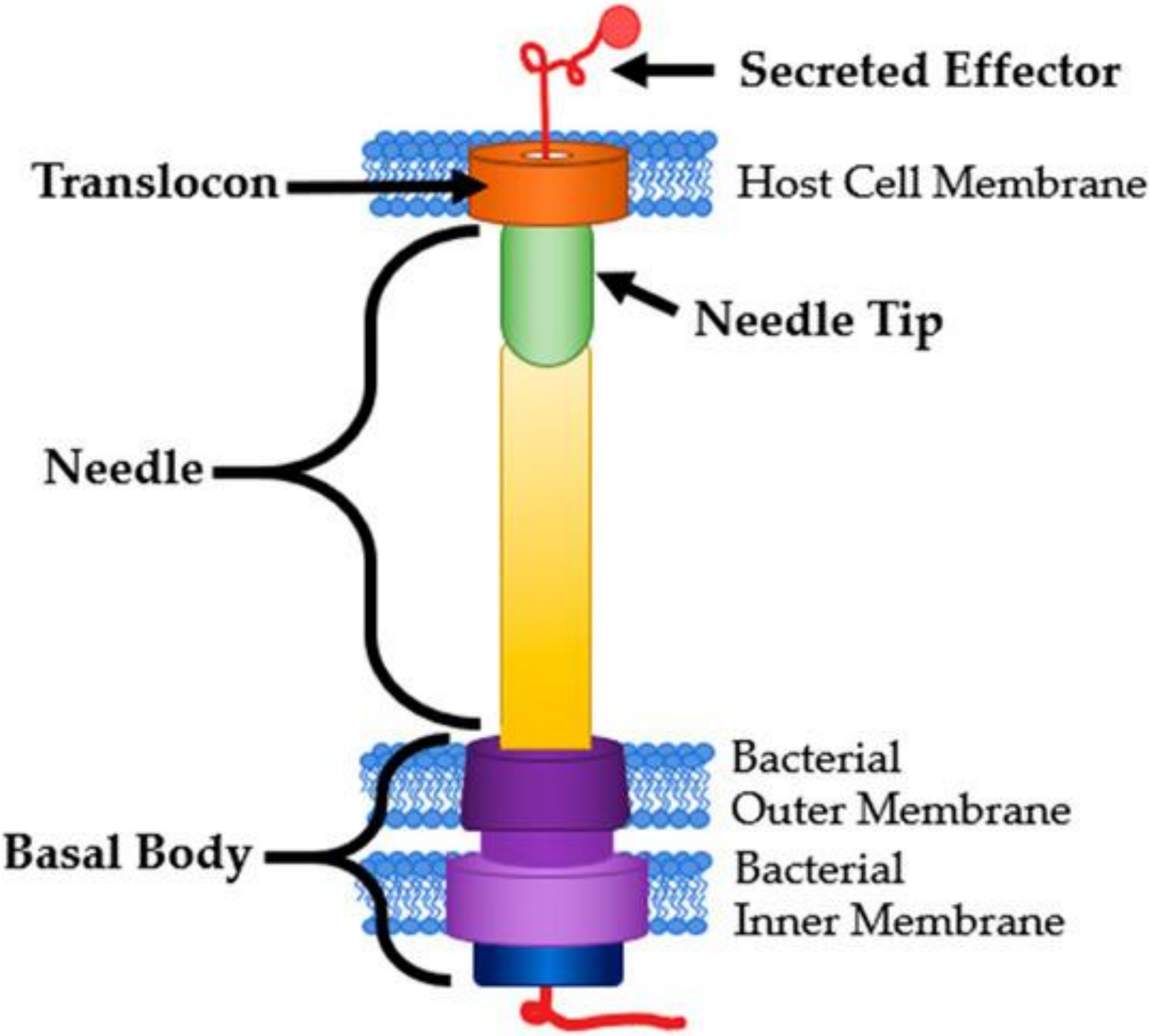


Figure 2: Mechanism and structure of the Type III Secretion System (T3SS)

2.9.3 Biological and ecological implications

Bacterial secretion systems do more than just support virulence. They contribute to symbiotic interactions, nutrient acquisition, bacterial competition, and biofilm formation, making them central to bacterial ecology (Tosi et al., 2013; Murphy & Brinkworth, 2021). In polymicrobial environments such as the gut, soil, or wastewater—secretion systems also play a defensive role, targeting competing bacteria through effector-mediated lysis (T6SS). The evolutionary diversification of secretion systems reflects bacterial adaptation to host pressures, environmental niches, and interspecies competition. Their study, particularly through genomic and structural biology tools, continues to unveil potential therapeutic targets for antivirulence drugs, offering an alternative strategy in the fight against AMR.

2.10 Historical Background and Technological Evolution

The origins of DNA sequencing trace back to the Sanger chain-termination method (Sanger & Coulson, 1975) and the Maxam-Gilbert chemical cleavage method, both pioneered in the late 1970s. Sanger sequencing, despite its high accuracy, is inherently low-throughput, allowing only one DNA fragment to be sequenced per reaction. This limitation significantly constrained its utility in large-scale genomic studies (Hu et al., 2021).

The NGS technologies, introduced commercially between 2004 and 2006, overcame these limitations by enabling massively parallel sequencing the simultaneous processing of millions of DNA fragments. This exponential leap in throughput was made possible through advances in nanotechnology, microfluidics, and automated imaging systems (Mardis, 2013; Slatko et al., 2018).

2.10.1 Applications of next-generation sequencing

Next-generation sequencing (NGS), also known as high-throughput sequencing (HTS), represents a transformative leap in genomic analysis by enabling the simultaneous sequencing of millions to billions of DNA fragments. Since its emergence in the early 2000s, NGS has revolutionized biomedical research, diagnostics, epidemiology, and microbial genomics by enabling rapid, cost-effective, and scalable whole-genome and targeted sequencing (Slatko et al., 2018).

2.10.2 Overview of NGS Platforms and Sequencing Strategies

Second-generation sequencing technologies, primarily dominated by Illumina and Ion Torrent, utilize short-read sequencing strategies. DNA is sheared into fragments (typically 200–500 bp), end-repaired, ligated to sequencing adapters, and then amplified *in situ*. Sequencing-by-synthesis is used to read each fragment, producing paired-end reads of 150–300 bp, which are then computationally aligned to a reference genome (Gong et al., 2021; Li et al., 2022).

Despite its robustness and accuracy, short-read sequencing poses challenges in resolving complex genomic regions such as structural variants (SVs), repetitive elements, and insertion sequences. These limitations are addressed by third-generation (long-read) sequencing platforms such as Pacific Biosciences (PacBio) and Oxford Nanopore Technologies (ONT). These technologies generate reads up to tens of kilobases in length, enabling direct sequencing of large genomic regions and real-time data acquisition with minimal library preparation (Chen et al., 2019; Hu et al., 2021). While long-read platforms improve genome assembly and variant detection, they have historically struggled with lower per-base accuracy compared to Illumina. However, recent advances such as PacBio HiFi and Nanopore Q20+ chemistry have significantly improved read fidelity, making long-read sequencing increasingly viable for clinical and regulatory applications.

2.10.3 Applications in microbial genomics and public health

Next-generation sequencing has become an indispensable tool in microbial genomics, with far-reaching implications for environmental, animal health and the One Health surveillance. Whole genome sequencing of foodborne pathogens offers unparalleled resolution for several key applications. It enables outbreak investigation and source attribution by comparing single nucleotide polymorphisms (SNPs) among isolates, thereby elucidating the clonal relationships between clinical, food, and environmental strains with high precision (Brown et al., 2019). InAMR profiling, NGS allows for the detection of resistance genes, integrons, plasmids, and MGEs directly from genomic data, aided by databases such as CARD and ResFinder for in silico annotation. Genomic analysis also supports virulence and serotype prediction, enabling the identification of key virulence factors (e.g., *stx*, *eae*, *ipaH*) and O:H antigen combinations without the need for phenotypic assays. Furthermore, high-resolution phylogenetic approaches based on WGS facilitate the study of pathogen evolution and transmission dynamics, providing insights into the emergence of clones such as *E. coli* ST131 and ST10, or hypervirulent *Klebsiella pneumoniae* (Zheng et al., 2021).

2.10.4 Integration into global surveillance and One Health frameworks

Regulatory agencies such as the CDC, Food and drug administration (FDA), European Food Safety Authority (EFSA), WHO, and FAO have integrated whole genome sequencing into global surveillance initiatives including platforms like PulseNet, GenomeTrakr, Global Antimicrobial Resistance Laboratory & Response Network (GLASS), and Integrated Rapid Infectious Disease Analysis (INNUENDO). These systems enhance early detection of emerging pathogens and AMR threats, promote global data sharing, and support the harmonization of genomic surveillance practices. Additionally, they play a crucial role in building surveillance capacity in LMICs. In South Africa and other LMICs, the adoption of WGS for monitoring pathogens such as *E. coli*, *Salmonella*, and *Listeria* is catalyzing a shift toward evidence-based decision-making in food safety and public health policy.

In Africa, the Africa Genomic Initiative has placed some of its efforts to strengthen the current molecular diagnostic efforts along with its genomic surveillance in this continent. The initiatives priority includes building on the skills development for molecular diagnostics across the African Union member states. Other priorities includes developing early-warning systems, sharing of genomic data as well as strengthing of mechanisms that enable systems that can detect emerging biosecurity threats.

2.10.5 Future prospects

As NGS costs continue to decline and analysis pipelines become increasingly automated and cloud-based, real-time metagenomic surveillance of pathogens in clinical, food, and environmental samples is becoming a reality. The ongoing development of portable sequencers (e.g., ONT MinION) and AI-powered genome analytics promises even broader accessibility and integration of NGS into field diagnostics, antimicrobial stewardship, and One Health-based interventions.

2. 11 Conclusion

Escherichia coli remains a well-studied bacterial organism worldwide; its remarkable adaptability to various environments and hosts has allowed this bacterium to continually advance its pathogenic potential. Over the years, with advancements in molecular diagnostics such as next-generation sequencing, this organism has been characterised to understand the mechanisms it uses to invade and cause infections in both humans and animals. To date, various pathotypes, serotypes and virulence factors are being determined with *in silico* techniques that do not require the use of traditionaldetection methods. This technology has revolutionised detection and surveillance techniques which allow in-depth analysis of such bacterial organisms.

Despite many advances with technologies, the literature we have reviewed in this chapter shows a critical gap in the current knowledge, though there are studies in animals and food products, but the data is still not significantly large enough to give a clear representation of non-clinical isolates. Addressing the gaps that exists is essential to inform evidence-based interventions for One-health surveillance frameworks. Within this context, our study plans to address some of these gaps in order to contribute to some of knowledge gaps that exists.

CHAPTER 3

Genomic characterization of generic *Escherichia coli* from food-producing animals and products of animal origin in South Africa

Research article published from this chapter:

Malesa, R., Pierneef, R., Magwedere, K., Mafuna, T. and Matle, I. (2024). Genomic characterisation of generic *Escherichia coli* from food-producing animals and products of animal origin in South Africa. *Frontiers in Bacteriology*. 3, p.1432292. https://doi.org/10.3389/fbrio.2024.1432292

Conference presentation from this thesis:

21st Annual Congress SASVPEM Congress held at Lagoon Beach Hotel, Cape Town on 21 to 23 August 2024: Virtual presentation.

Published Article

frontiers | Frontiers in Bacteriology

TYPE Original Research
PUBLISHED 16 July 2024
DOI 10.3389/fbrio.2024.1432292

Check for updates

OPEN ACCESS

EDITED BY
Kumaragurubaran Karthik,
Tamil Nadu Veterinary and Animal Sciences University, India

REVIEWED BY
Anca Farkas,
Babeş-Bolyai University, Romania
Dalia Hamza,
Cairo University, Egypt
Armando Navarro,
National Autonomous University of Mexico, Mexico

*CORRESPONDENCE
Refiloe Malesa
✉ malokotsar@arc.agric.za
Itumeleng Matle
✉ matlei@arc.agric.za

Genomic characterisation of generic *Escherichia coli* from food-producing animals and products of animal origin in South Africa

Refiloe Malesa[1]*, Rian Pierneef[2,3], Kudakwashe Magwedere[4], Thendo Mafuna[5] and Itumeleng Matle[1]*

[1]Bacteriology Division, Agricultural Research Council-Onderstepoort Veterinary Research, Onderstepoort, South Africa, [2]Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Pretoria, South Africa, [3]Centre for Bioinformatics and Computational Biology, University of Pretoria, Pretoria, South Africa, [4]Unaffiliated Global Health Security Research Scholar, Pretoria, South Africa, [5]Department of Biochemistry, University of Johannesburg, Johannesburg, South Africa

3.1 Abstract

Escherichia coli is found in diverse environmental niches, including meat and meat products, and is known for its significance in both food safety and public health. In South Africa, whole genomic sequencing (WGS) efforts for *E. coli* are overwhelmingly skewed toward clinical isolates with limited studies conducted on non-clinical isolates. Therefore, the aim of this study was to use WGS to characterize generic strains of *E. coli* isolated from animal specimens, meat, and meat-based products in South Africa. Based on WGS analysis, a total of 35 *E. coli* strains were grouped into five phylogroups (A, B1, B2, C, and E), with A (46 %) being the most predominant. Virulence-associated genes identified the isolates as either extra-intestinal pathogenic *E. coli* (EIEC) (69%) or intestinal pathogenic *E. coli* (IPEC) (31%) pathotypes. Twenty-three different serotypes were identified, with O101:H37, O2:H4 (17 %), O6:H5 (13 %), and O64:H19 (9 %) being the predominant ones. Among the 19 different sequence types (STs), ST1858, ST975, and ST10 were the most prevalent (11% each). Various virulence genes, antimicrobial resistance genes, and genetic mobile elements carrying Tn2, IS26, and Tn6196 elements were detected, with the metal transport system *sitABCD* being the most predominant. The type 1 CRISPR system, which functions by storing records of previous invasions to provide immunological memory for a rapid and robust response upon subsequent viral infections, was detected in all isolates, consisting of subtypes I-E (86%), I-A (57%), and I-F (11%). The findings of this study provide an insight into the genetic diversity of generic *E. coli* isolates from animal species, meat, and meat-based products in South Africa.

Keywords: Phylogroups, extra-intestinal pathogenic *E. coli*, intestinal pathogenic *E. coli*, sequence type, virulence genes, antimicrobial resistance genes, CRISPR system.

3.2 Introduction

Food-producing animal slaughter establishments and processors test for *E. coli* Biotype I (generic *E. coli*) to verify the effectiveness of their process controls and sanitation practices. Faecal contamination, one of the primary sources of pathogenic and non-pathogenic organisms that contaminate animal protein foods, is commonly indicated by generic *E. coli* (Gekenidis et al., 2018). This bacterium is prevalent since it inhabits the gastrointestinal tract of both humans and animals (Allahyari, et al., 2021; Pakbin et al., 2021; Ali & Alsayeqh, 2022). The performance criteria for generic *E. coli* are not enforceable, as the bacteria numbers simply represent microbial loads used to monitor and verify whether the slaughter and/or production process was adequately controlled. These criteria provide guidance to livestock slaughter establishments on the effectiveness of their processes in preventing faecal contamination. Test results serve as evidence that the slaughter or production maintained sufficient process controls for hygienic dressing. *E. coli* is ubiquitous, inhabiting diverse environments such as water sources, animals, and food (Lupindu, 2017; Galindo-Méndez, 2020). However, the characterization of generic *E. coli* and its potential pathogenicity to humans is often overlooked, despite *E. coli* being a significant cause of serious diseases in humans and animals.

In sub-Saharan Africa, *E. coli* diarrheal infections represent a significant public health challenge, with a high incidence attributed to factors such as limited access to clean water, sanitation and inadequate hygiene (Robert et al., 2021). *E. coli* strains causing intestinal infections are known as intestinal pathogenic *E. coli* (InPEC), while those responsible for extraintestinal infections are termed extraintestinal pathogenic *E. coli* (ExPEC). These infections encompass various pathotypes, each characterized by specific traits (Johnson & Russo, 2018; Meena et al., 2021, 2023). InPEC is linked to pathotypes such as enterotoxigenic *E. coli* (ETEC), enteropathogenic *E. coli* (EPEC), enteroaggregative *E. coli* (EAEC), shiga toxin-producing *E. coli* (STEC), diffusely adherent *E. coli* (DAEC), adherent-invasive *E. coli* (AIEC), and enteroinvasive *E. coli* (EIEC). Extraintestinal pathogenic *E. coli* infections are associated with pathotypes such as avian pathogenic *E. coli* (APEC), uropathogenic *E. coli* (UPEC), neonatal meningitis *E. coli* (NMEC), and sepsis-associated *E. coli* (SEPEC), primarily affecting humans (Martinez-Medina, 2021).

There is a compelling need to closely monitor the spread of *E. coli* in animals and food derived from animal sources (Ramos et al., 2020). Furthermore, there's a growing apprehension regarding the potential for *E. coli* to acquireAMR traits within livestock environments. This evolution could not only complicate but also escalate the cost of treating infections in both humans and animals (Palma et al., 2020). Such antimicrobial resistance poses a significant public health concern, as it reduces the effectiveness of antibiotics, thereby limiting treatment options and potentially increasing the severity and duration of illnesses caused by *E. coli* infections.

Studies have increasingly highlighted the intriguing link between the virulence of *E. coli* strains and their Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR-Cas) systems. CRISPR-Cas systems, originally identified as adaptive immune mechanisms in bacteria and archaea, play a crucial role in defending against foreign genetic elements such as bacteriophages and plasmids (Koonin et al., 2017; Murugan et al., 2017). Several studies have suggested that certain *E. coli* strains with more robust CRISPR-Cas systems may exhibit decreased virulence due to their enhanced ability to fend off invading genetic elements (García-Gutiérrez et al., 2015). Conversely, strains with compromised or less effective CRISPR-Cas systems might be associated with higher virulence as they struggle to combat invasive genetic elements, potentially including virulence factors (Louwen et al., 2014). Furthermore, the interplay between CRISPR and virulence in *E. coli* is complex and multifaceted. It involves various factors such as the specific composition and activity of CRISPR systems, the presence of virulence genes, and the environmental context in which the bacteria reside. Understanding this interrelationship can offer insights into the evolution of *E. coli* pathogenicity and potentially inform strategies for combating *E. coli*related infections (Kang & Lee, 2022).

Cases and outbreaks of *E. coli* in both animals and humans are well-documented in South Africa (Gambushe et al., 2022; Khabo-Mmekoa et al., 2022; Manyi-Loh & Lues, 2023). However, the primary focus of WGS endeavours has predominantly targeted human clinical strains associated with outbreaks (Muloi et al., 2018; Massella et al., 2020). Additionally, attention has been devoted to analysing environmental samples (Nadalian et al., 2021). Nonetheless, genomic data regarding the complete diversity of *E. coli* colonizing food and animals, their array of antimicrobial resistance genes (ARGs), associated mobile genetic elements (MGEs), virulence-associated genes (VAGs), and the possible presence of ExPEC is lacking in South Africa. This study utilised WGS to analyse 35 *E. coli* isolates collected from animals, meat, and meat products in South Africa to highlight the significance and requirement of such studies within South Africa over extended period, few of these samples are international samples which were collected from ports of entry in South Africa.

3.3 Material and Methods

3.3.1 Isolate Selection

The isolates utilized in this study were obtained from samples processed between 1988 and 2018 at the Bacteriology Laboratory of the Onderstepoort Veterinary Research Institute in South Africa, as part of their routine diagnostic testing. Hence, for this research, a selection of 35 *E. coli* isolates was made, encompassing diverse geographical regions within the country (Gauteng, Free State, North-West, Limpopo, and Mpumalanga), as well as international locales, different isolation sources (animal, meat and meat products such as a traditional sausage called wors), and a range of animal species (poultry, porcine, bovine and ovine) (Table 1), a water sample was included in this analysis in order to trace the source of *E. coli* contamination. These international samples are isolates from imported meat samples intended for local consumption. The isolates were stored in a lyophilized form and then reconstituted by inoculating them into brain heart infusion (BHI) broth, followed by incubation at 37 °C for 18 to 24. Although this study included isolates that span from 1976 to 2018 from various sources, the sample size is relatively small and the diverse sampling used does limit our ability to generalise the context in a wider context. Predominance of certain species does introduce bias, therefore our findings should be interpreted based on their relation to the sample size not as a full representation of isolates circulating in South Africa.

Table 2: *Escherichia coli* isolates sequenced in this study (n = 35) and their corresponding metadata

SAMPLE NO.	Accession no.	GEOGRAPHIC LOCATION	REGION	SAMPLE TYPE	Source of isolation	ANIMAL SPP	SAMPLING YEAR
S1	SAMN41920844	South Africa	North West	Beef and Pork Wors	Processed meat	Bovine and Porcine	2016
S2	SAMN41920845	North America	Canada	Turkey drumstick	Raw poultry	Poultry	2016
S3	SAMN41920846	South Africa	Mpumalanga	Chicken drumstick	Raw poultry	Poultry	2016
S4	SAMN41920847	Netherlands	Holland	Chicken leg quarter	Raw poultry	Poultry	2016
S5	SAMN41920848	South Africa	Free State	Chicken wings	Raw poultry	Poultry	2016
S6	SAMN41920849	South Africa	Free State	Chicken drumstick	Raw poultry	Poultry	2016
S7	SAMN41920850	South Africa	Free State	Chicken fillet	Raw poultry	Poultry	2016
S8	SAMN41920851	South Africa	Free State	Chicken leg quarter	Raw poultry	Poultry	2016
S9	SAMN41920852	South Africa	Free State	Chicken drumstick	Raw poultry	Poultry	2016
S10	SAMN41920853	South Africa	Free State	Chicken wings	Raw poultry	Poultry	2016
S11	SAMN41920854	South Africa	Free State	Chicken drumstick	Raw poultry	Poultry	2016
S12	SAMN41920855	South Africa	Free State	Pork wors	Processed pork	Porcine	2016
S13	SAMN41920856	South Africa	Free State	Lamb leash	Raw lamb	Ovine	2016
S14	SAMN41920857	South Africa	Free State	Pork shoulder	Raw pork	Porcine	2016
S15	SAMN41920858	South Africa	Free State	Pork chops	Raw pork	Porcine	2016
S17	SAMN41920859	South Africa	North West	Chicken drumstick and wings	Raw poultry	Poultry	2016
S18	SAMN41920860	South Africa	Free State	Chicken thigh	Raw poultry	Poultry	2016
S19	SAMN41920861	South Africa	Gauteng	Animal faeces	Digestive system	Porcine	1988
S20	SAMN41920862	South Africa	Gauteng	Beef mince	Processed beef	Bovine	2018
S21	SAMN41920863	South Africa	Gauteng	Animal faeces	Digestive system	Ovine	1992
S22	SAMN41920864	South Africa	Gauteng	Animal faeces	Digestive system	Porcine	1998
S23	SAMN41920865	South Africa	Gauteng	Animal faeces	Digestive system	Unknown	1988
S24	SAMN41920866	South Africa	Free State	Beef wors	Processed beef	Bovine	2018
S25	SAMN41920867	South Africa	Gauteng	Animal faeces	Digestive system	Unknown	1998
S26	SAMN41920868	South Africa	Gauteng	Animal faeces	Digestive system	Porcine	1976
S27	SAMN41920869	South Africa	Gauteng	Animal faeces	Digestive system	Unknown	1988
S28	SAMN41920870	Europe	Belgium	Chicken leg quarter	Raw poultry	Poultry	2018
S29	SAMN41920871	South Africa	Gauteng	Animal faeces	Digestive system	Porcine	1988
S30	SAMN41920872	South Africa	Gauteng	Animal faeces	Digestive system	Unknown	1992
S32	SAMN41920873	South Africa	Gauteng	Animal faeces	Digestive system	Porcine	1988

S33	SAMN41920874	South Africa	Gauteng	Animal faeces	Digestive system	Unknown	2003
S34	SAMN41920875	South Africa	Gauteng	Animal faeces	Digestive system	Porcine	1996
S35	SAMN41920876	South Africa	Gauteng	Water	Water	N/A	2004
S36	SAMN41920877	South Africa	Limpopo	Beef biltong	RTE beef	Bovine	2018
S37	SAMN41920878	South Africa	Gauteng	Animal faeces	Digestive system	Unknown	1992

3.3.2 Genomic DNA extraction and whole-genome sequencing

Genomic DNA was extracted from overnight cultures using the High Pure PCR Template Preparation Kit (Roche, Germany) according tothe manufacturer's instructions. Purity and concentration of the DNA were assessed using a Nanodrop 1000 spectrophotometer. Subsequently, WGS was performed at the Biotechnology Platform Agricultural Research Council, Onderstepoort, South Africa, using a HiSeq 2500 instrument (Illumina, San Diego, CA, USA). The construction of DNA libraries was accomplished using TruSeq DNA library preparation kits (Illumina, San Diego, CA, USA). Many HiSeq 2500 platform runs produce paired-end reads that are over ~150 bp (Gong et al., 2021). The sequencing depth for bacterial genomes generated using this platform generally exceeds 30× coverage, providing sufficient resolution for reliable genome assembly and downstream comparative genomic analyses

3.3.3 Data pre-processing, quality control

The raw read quality was assessed with FastQC v.0.11.9 (Andrews, 2010) and the adapters and low-quality reads were trimmed using Trimmomatic v.0.39 (Bolger et al., 2014). SPAdes (v3.15.3) was used for assembly of each isolate (Prjibelski et al., 2020) and assembly quality was analysed using Quast v4.4 (Gurevich et al., 2013). Genome quality along with contamination levels were assessed using CheckM v1.0.18 (Parks et al., 2015). The isolates were annotated using Prokka v1.13.7 (Seemann, 2014). To determine the strains taxonomic classification, a portion of the complete nucleotide sequences assemblies in this study were aligned using the Basic local alignment tool (BLASTN) against the nucleotide sequences on the NCBI database. Isolate with a percentage identity of >90% to *E. coli* on the NCBI database were accepted (Peker et al., 2019).

3.3.4 Detection of phylogroups, serotypes and pathotypes

Clermont quadruplex phylo-group assignment technique was employed to determine *E. coli* phylogroups. Fasta files containing *E. coli* contigs were uploaded on the Clermont Typing website (http://clermontyping.iame-research.center/) and sequences were analysed using the default settings (Clermont et al., 2019). The assembled genomes of *E. coli* isolates were used to perform *in silico* serotyping of the O and H antigens, employing SerotypeFinder gene database hosted at the Center for Genomic Epidemiology (CGE), accessible at https://cge.food.dtu.dk/services/SerotypeFinder/. Specifically, for the O antigen, the database analyzed the *wzx*, *wzy*, *wzm*, and *wzt* genes, while for flagellin H-antigen, it processed the *fliC*, *flkA*, *flmA*, *flnA*, and *fllA* genes (Joensen et al., 2015). The analysis was conducted with the default parameters specified on the website, including a minimum sequence length of 60% and a threshold of 85%.

Escherichia coli pathotypes in this study were classified based on their virulence factor characteristics as seen in literature (Jesser & Levy, 2020; Riley, 2020; Robins-Browne & Hartland, 2002; Enciso-Martínez et al., 2022; Geurtsen et al., 2022) with isolates classified as either InPEC or ExPEC (Supplementary data Table S 1). Each isolate's virulence characteristics were analysed and assigned a pathotype based on the presence of the target gene. Identification was based on a combination of the main virulence factor genes which are capable of causing disease and these included genes responsible for attachment, production of toxins or hemolysis. In order to classify InPEC according to the corresponding pathotype, ETEC must contain either LT or ST enterotoxin, STEC must contain any *stx* gene and the EPEC must contain the intimin *eae.* On the ExPEC classification, the UPEC, APEC and NMEC pathotype assignment depends on presence of any of the two target genes listed on **Error! Reference source not found.**

3.3.5 Multi-locus sequence typing

Sequence types were identified through multi-locus sequence typing (MLST) using version 2.0 of the CGE tool, which is accessible at https://cge.food.dtu.dk/services/MLST/. On the CGE tool, select the MLST configuration (*Escherichia coli #1*), select the minimum depth for allele of 5X, input the assembled genomes (Larsen et al., 2012). Retrieve and analyse the results based on the sequence types identified, confidence scores as well the allele profiles.

3.3.6 Determination of virulence factors

As part of the publicly accessible web-based tools for WGS analysis offered by the CGE, virulence factors within this study were determined using this web-based platform. The assembled genomes of *E. coli* isolates were submitted on the VirulenceFinder database (https://cge.cbs.dtu.dk/services/VirulenceFinder/) to detect virulence genes, *E. coli* species was selected, a threshold of over 90% identity with a minimum length of 60% was selected and assembled genomes were analysed (Joensen et al., 2015). In order to select significant genes which, encode virulence for InPEC and ExPEC, (Chapman et al., 2006) and (Badi et al., 2018) were used as a reference guide to select specific genes (Supplementary dataTable S 1).

3.3.7 Determination of human-associated pathogenicity

The isolates underwent analysis using PathogenFinder v. 1.1 (https://cge.food.dtu.dk/services/PathogenFinder/) to assess their potential human-associated pathogenicity. Assembled genomes were uploaded, and the phylum selection included all the classes of bacteria that can be detectable by the system. The system predicts the number of pathogenic and non-pathogenic bacterial families (Zhang & Ye, 2017).

3.3.8 Detection of CRISPR-associated genes (*Cas*)

The presence and characteristics of a cluster of regularly spaced short palindromic repeats (CRISPR) were determined using the CRISPRone online tool (https://omics.informatics.indiana.edu/CRISPRone/). This tool allows searching for CRISPR-Cas system genes and proteins, as well as class types and subtypes of the system. It also returns the number of loci, length, and nucleotide sequences of repeat spacers (Zhang & Ye, 2017).

3.3.9 Detection of resistance genes, plasmids and mobile genetic elements

Resistance genes and antibiotic phenotypes were obtained from ResFinder v4.1 (https://cge.food.dtu.dk/services/ResFinder/). Plasmid-associated genes and mobile genetic elements genes were obtained from PlasmidFinder v2.1 (https://cge.food.dtu.dk/services/PlasmidFinder/ and MobileElementFinder v1.0.3 (https://cge.food.dtu.dk/services/MobileElementFinder/), respectively. Plasmid Inc types hosted by pathogenwatch (https://pathogen.watch/) was used to verify the result of PlasmidFinder (Argimón et al., 2021).

3.3.10 Determination of antimicrobial resistance phenotypes

Assembled genomes of isolates within this study were analysed for the presence of antimicrobial resistance phenotypes using the ResFinder v4 tool with default parameters. This tool is able to analyse antibiograms *in silico*. Using ResFinder, a threshold percentage identity of 90% with a minimum length of 60% was used for antimicrobial resistance genes (Bortolaia et al., 2020). Presence of antimicrobial resistance genes within the isolates and the antimicrobials identified by the tool were used to infer presence of antimicrobial resistance phenotypes.

3.4 Results

3.4.1 Identification of *Escherichia coli* strains

A total of 35 isolates were obtained from animal specimens, meat and meat products in South Africa, as well as international samples from the country's ports of entryinto the country. Table S5 provides an overall summary of the results, grouped by species. These isolates underwent WGS, and all were identified as *E. coli* using NCBI BLASTN tool. The BLASTN results showed that all the genomes from this study clustered among publicly available *E. coli* genomes with a percentage identity of over 95% to those on the NCBI database (Table S3) (Peker et al., 2019).

3.4.2 Detection of *Escherichia coli* phylogroups, serotypes and pathotypes

Upon phylogenetic classification, phylogroup A was revealed as the most prevalent among poultry isolates, accounting for 46% of the total isolates examined in this study, followed by phylogroup B1 (20%), B2 (6%), with phylogroups C and E contributing 3%, these other phylogroups were widely distributed among various animal species (bovine, ovine and porcine). Phylogroup C was detected in both poultry and porcine samples, while phylogroup E was exclusively isolated from poultry samples. Notably, all phylogroups were observed across different provinces in South Africa. Twenty-three different serotypes were identified with O101:H37 (17%), O2:H4 (17%), O6:H5 (13%) and O64:H19 (9%) being the predominant (Table S2 and S5). The distribution of serotypes across provinces showed that majority of isolates originated from the Free State (43%) and Gauteng province (36%), followed by some from Northwest (14%) and only one (7%) from a port of entry.

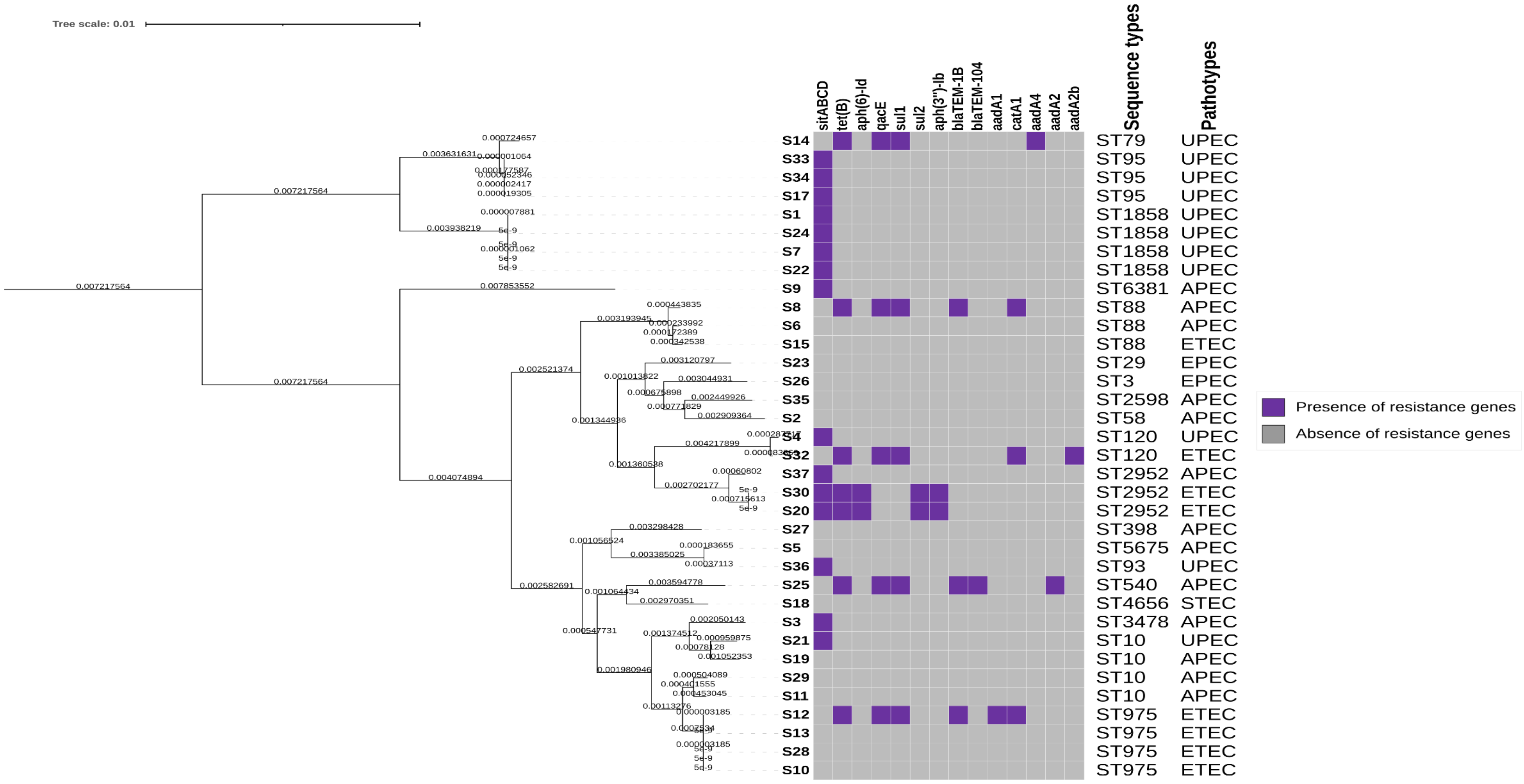
In this study, virulence-associated genes of *E. coli* pathotypes were utilized to characterize isolates based on the pathotype they belong to, thereby classifying them as either InPEC or ExPEC. The majority of the isolates belonged to the ExPEC pathotype, which comprised 69%, with APEC accounting for 37% and UPEC for 31%. Among the APEC pathotypes, pathotypes, poultry contributed 20%, while the UPEC pathotype contributed 9% for bovine, and poultry each and 3% for porcine. In contrast, InPEC pathotypes comprised only 31% of the isolates, with ETEC representing 23%. The ETEC pathotype in poultry, and porcine contributed 6% in each, in bovine and ovine contributed 3%. The STEC contributed 3% and was found in poultry, EPEC was only found in porcine and contributed 6% of the isolates, Table S5 and

Figure ***3*** lists the pathotype assignment for each isolate.

3.4.3 Determination of *Escherichia coli* MLSTs diversity

The isolates were classified into a total of 19 distinct STs. Among these, ST1858, ST975, and ST10 emerged as the most prevalent, each accounting for 11% of the isolates. ST95, ST88, and ST2952, constituted 9%, while ST120 was 6% of the total isolates, respectively. All the identified STs are presented in

Figure ***4***, Figure S1 and Table S2.



	sitABCD	tet(B)	aph(6)-Id	qacE	sul1	sul2	aph(3'')-Ib	blaTEM-1B	blaTEM-104	aadA1	catA1	aadA4	aadA2	aadA2b	Sequence types	Pathotypes
S14	□	■	□	■	■	□	□	□	□	□	□	■	□	□	ST79	UPEC
S33	■	□	□	□	□	□	□	□	□	□	□	□	□	□	ST95	UPEC
S34	■	□	□	□	□	□	□	□	□	□	□	□	□	□	ST95	UPEC
S17	■	□	□	□	□	□	□	□	□	□	□	□	□	□	ST95	UPEC
S1	■	□	□	□	□	□	□	□	□	□	□	□	□	□	ST1858	UPEC
S24	■	□	□	□	□	□	□	□	□	□	□	□	□	□	ST1858	UPEC
S7	■	□	□	□	□	□	□	□	□	□	□	□	□	□	ST1858	UPEC
S22	■	□	□	□	□	□	□	□	□	□	□	□	□	□	ST1858	UPEC
S9	■	□	□	□	□	□	□	□	□	□	□	□	□	□	ST6381	APEC
S8	□	■	□	■	■	□	□	■	□	□	■	□	□	□	ST88	APEC
S6	□	□	□	□	□	□	□	□	□	□	□	□	□	□	ST88	APEC
S15	□	□	□	□	□	□	□	□	□	□	□	□	□	□	ST88	ETEC
S23	□	□	□	□	□	□	□	□	□	□	□	□	□	□	ST29	EPEC
S26	□	□	□	□	□	□	□	□	□	□	□	□	□	□	ST3	EPEC
S35	□	□	□	□	□	□	□	□	□	□	□	□	□	□	ST2598	APEC
S2	□	□	□	□	□	□	□	□	□	□	□	□	□	□	ST58	APEC
S4	■	□	□	□	□	□	□	□	□	□	□	□	□	□	ST120	UPEC
S32	□	■	□	■	■	□	□	□	□	□	■	□	□	■	ST120	ETEC
S37	■	□	□	□	□	□	□	□	□	□	□	□	□	□	ST2952	APEC
S30	■	■	■	□	□	■	■	□	□	□	□	□	□	□	ST2952	ETEC
S20	■	■	■	□	□	■	■	□	□	□	□	□	□	□	ST2952	ETEC
S27	□	□	□	□	□	□	□	□	□	□	□	□	□	□	ST398	APEC
S5	□	□	□	□	□	□	□	□	□	□	□	□	□	□	ST5675	APEC
S36	■	□	□	□	□	□	□	□	□	□	□	□	□	□	ST93	UPEC
S25	□	■	□	■	■	□	□	■	■	□	□	□	■	□	ST540	APEC
S18	□	□	□	□	□	□	□	□	□	□	□	□	□	□	ST4656	STEC
S3	■	□	□	□	□	□	□	□	□	□	□	□	□	□	ST3478	APEC
S21	■	□	□	□	□	□	□	□	□	□	□	□	□	□	ST10	UPEC
S19	□	□	□	□	□	□	□	□	□	□	□	□	□	□	ST10	APEC
S29	□	□	□	□	□	□	□	□	□	□	□	□	□	□	ST10	APEC
S11	□	□	□	□	□	□	□	□	□	□	□	□	□	□	ST10	APEC
S12	□	■	□	■	■	□	□	■	□	■	■	□	□	□	ST975	ETEC
S13	□	□	□	□	□	□	□	□	□	□	□	□	□	□	ST975	ETEC
S28	□	□	□	□	□	□	□	□	□	□	□	□	□	□	ST975	ETEC
S10	□	□	□	□	□	□	□	□	□	□	□	□	□	□	ST975	ETEC

Figure 3: Heat map showing the distribution of *E. coli* antimicrobial resistance genes, violet colour indicates presence of resistance genes and grey indicates absence of resistance genes. The top panel of the x-axis indicates the antimicrobial resistance genes detected. The right panel of the map indicates the sequence types and pathotypes associated with these isolates. The heatmap was generated using

sequencing–derived antimicrobial resistance gene profiles of isolates in this study. Individual isolates are represented in each row. The AMR genes were identified using ResFinder bioinformatics tool. The heatmap was constructed using Itol, the figure allows for comparison of resistance genes across various sequence types and pathotypes.

3.4.4 Detection of virulence genes

Figure ***4*** shows distribution of virulence genes detected in all the isolates. The predominant virulence genes detected in the isolates encompassed a variety of crucial factors. Notably, *csgA* (100%), encoding the curlin major subunit, and *nlpI* (100%), responsible for encoding the lipoprotein NlpI precursor, were universally present. This was followed closely by *terC* (associated with tellurium ion resistance) detected at 94%, and *fimH* (involved in type 1 fimbriae formation) at 89%. The YHD fimbrial clusters were also prevalent, with 86% for *yehA* encoding the outer membrane lipoprotein, 88.6% for *yehB* encoding the usher, *yehC* (89%) encoding the chaperone, and *yehD* (86%) encoding the major pilin subunit. Among other notable genes were *fdeC* (86%), an intimin-like adhesion, and *hlyE* (77%), encoding avian *E. coli* haemolysin. The *iss* gene (54%), associated with increased serum survival, and *irp2* (46%), encoding a high molecular weight protein 2 non-ribosomal peptide synthetase, were also detected in the isolates. Less frequently observed genes included *stx2* (3%), a Shiga-toxin gene, and various type three secretion effector proteins such as *espA*, *espB, espF, espJ*, and *cif,* collectively found in 6% of the isolates.

Major genes clusters were observed within the study. Genes contributing to *E. coli* pathogenicity were identified which included adherence genes (*csgA, fimH, lpfA, pap, afa*), iron-uptake genes (*fyuA, chuA, iucC, iroN*), the capsule synthesis genes (*kpsE, kpsMII, kpsMIII*), genes responsible for invasion and survival (*ompT, ibeA, tia*) and those responsible for the secretion system (*esp* genes). Other virulence genes which could not be classified into a group were also identified.

3.4.5 Determination of human-associated pathogenicity

All the isolates within this study were confirmed as human pathogens that belong to the Gammaproteobacteria class.

3.4.6 Detection of CRISPR-associated genes (*Cas*)

The type 1 CRISPR system was found across all isolates, with a noteworthy distribution of subtypes. Specifically, 86% exhibited subtype I-E, 57% subtype I-A, and 11% subtype I-F. Among the subtypes, 80% featured a singular CRISPR loci, while the remaining 20% exhibited two loci. Subtype I-E demonstrated distinctive characteristics, including the presence of *cas3* (82%), *cas6e* (51%), and *cas8e* (77%). Subtype I-A was characterized by *cas5* (51%) and *cas7* (51%), whereas Subtype I-F exhibited *cas5f*, *cas6f*, *cas7f*, and *cas8f*, each detected in 11% of the isolates. Within the CRISPR system, Cas genes featuring a nuclease with the DEDDh motif were identified in 66% of the isolates. Additionally, the universal *cas1* and *cas2* elements were present in 63% of the isolates, adding a foundational element to the diversity observed within the CRISPR loci across the studied isolates.

Figure 4: A heatmap of *E. coli* virulence genes, the map shows presence/absence of genes. Violet colour indicates presence of genes and grey indicates absence of genes in isolates from animal specimen, meat and meat products in South Africa. The genes are listed on the top panel of the x-axis. The right panel indicates the sequence types and pathotypes of each of the isolates represented on the phylogenetic tress. The heatmap was generated using sequencing-derived virulence gene profiles of isolates included in this study. Individual isolates are represented in each row. Virulence genes were identified using the

VirulenceFinder bioinformatics tool. The heatmap was constructed using iTOL, and the figure allows for comparison of virulence gene distribution across different sequence types and pathotypes.

3.4.7 Predicted antimicrobial resistance genes and antibiotic resistance

Antimicrobial resistance profiles were inferred in silico on the presence of known antimicrobial resistance genes, therefore, these predictions represent genotype-baased resistance rather than those confirmed phenotypically. The distribution of AMR gene presence among the isolates is presented in Table S 4,,

Figure ***3***, and Figure S1. The metal transport system gene *sitABCD* (43%) emerged as the most detected resistance gene among these *E. coli* isolates. Within the subset of isolates exhibiting *sitABCD* resistance, additional genes were identified, including the tetracycline resistance gene *tet(*B*)* (23%), and the disinfectant resistance gene *qacE* (14%). Sulfonamide resistance genes *sul1* (9%) and *sul2* (6%) were also identified, along with the chloramphenicol resistance gene *catA1* (6%). Beta-lactam resistance genes *blaTEM-1B* (9%) and *blaTEM-104* (3%) were detected. Furthermore, aminoglycoside resistance genes *aadA1* (6%), *aadA2b* (6%), *aph(6)-Id*, and *aph(3'')-Ib* (6%) were observed. Notably, at least fifteen isolates, comprising 43% of the total, did not exhibit any resistance genes.

3.4.8 Predicted antimicrobial resistance profiles

The putative AMR phenotypes of these isolates were determined *in silico* using Resfinder tool, this was employed due to lack of the minimum inhibitory concentration (MIC) results and also putting into consideration that these isolates are primarily historic samples. The results revealed that all isolates predicted resistance to cotrimoxazole (100%), while 97% were resistant to penicillin, Table S5. Additionally, resistance to the chemical disinfectant hydrogen peroxide was observed in 37% of the isolates. Furthermore, resistance to amoxicillin (26%), sulfamethazine (17%), tetracycline (14%), chloramphenicol (11%), ceftazidime (14%), aztreonam (9%), piperacillin (9%), ampicillin (6%), and azithromycin (3%) was also identified (Table S5). Multidrug resistance (MDR) was observed in seven isolates (20%). MDR in this study is defined as those isolates that conferred resistance to five or seven antibiotics. The most common class of antibiotics among these multi-drug-resistant isolates were beta-lactams, sulphonamides, tetracycline, and chloramphenicol.

3.4.9 Mobile genetic elements and plasmids

Mobile elements detected in the isolates from this study exhibited a diverse array, showcasing various transposable elements (IS elements), ISEc elements, transposons, and the presence of a specific Miniature Inverted-repeat Transposable Element (MITEEc1) of 123bp in each isolate. Among the detected IS elements, a spectrum of types was identified, including IS4, IS5, IS100, IS609, IS682, IS911, and IS45. The ISEc elements displayed similar diversity, encompassing ISEc1, ISEc42, ISEc39, ISEc38, ISEc10, ISEc45, ISEc30, and ISEc52. Notably, transposons such as Tn2 and Tn6196 were also identified, further highlighting the complexity of mobile genetic elements in these isolates.

Antimicrobial resistance genes detected within the multidrug resistant isolate were found to be in different contig positions. Sample S8 isolated from poultry meat was found to have resistance genes on node 245, resistance genes were *sul1, aadA1, qacE* with IS26, *blaTEM-1B* was found on node 252 with MGE Tn2 and *tet(*B*)* was found on node 222 with no MGE. Sample S12 from porcine meat had resistance on node 121 which carried *tet(*B*)* and *blaTEM-1B* with MGE Tn2, node 143 carried *aadA1, qacE* and *sul1* with no MGE, node 146 carried *catA1* with no MGE. Sample S14 from porcine meat had resistance on node 24 which carried *aadA4, sul1* and *qacE,* on the same contig Tn6196 was detected. Node 170 carried *tet(*B*).* Sample S20 isolated from animal faeces had a resistance gene *aph(6)-Id, aph(3'')-Ib* and *sul2* on node 117 and *tet(*B*)* was on node 112, MGE were not detected in this isolate. Sample S25 from animal faeces had no MGE, node 180 carried *tet(*B*),* node 187 had *qacE, sul1* and *aadA2,* node 295 had *blaTEM-1B* and *blaTEM-104* and node 128 carried *catA1.* Sample S30 isolated from animal faeces had no MGE, while node 113 had *sul2, aph(3'')-Ib* and *aph(6)-Id* and node 121 carried *tet(*B*).* Sample S32 from animal faeces node 228 carried *tet(B),* node 225 *qacE, sul1* and *aadA2b,* node 185 carried *catA1.* Node 150 carried *blaTEM-1B* with Tn2.

As illustrated in Figure 5, the analysis of plasmid replicons using in silico WGS revealed a predominant presence of incompatibility groups, particularly IncFII (n=23, 66%) and IncFIB (n=21, 60%), across various isolates. Additionally, Col plasmid groups were also detected, with Col440I (n=2, 6%), Col156 (n=5, 14%), ColRNAI (n=2, 6%), ColE10 (n=1, 3%), and Col440II (n=2, 6%) exhibiting distribution among the isolates. Other plasmid types, such as IncFIA (n=7, 20%), IncB/O/K/Z (n=1, 3%), IncI2 (Delta) (n=1, 3%), and IncHI2A (n=1, 3%), were also identified, contributing to the overall plasmid landscape. Furthermore, the study identified isolates displaying multidrug resistance that harbored the class 1 integron (*intI1*), characterized by cassette arrays including *aadA1* and *aadA2b*, thereby adding another layer of genetic complexity to the mobile elements within this *E. coli* population. These isolates were isolated from porcine and bovine.

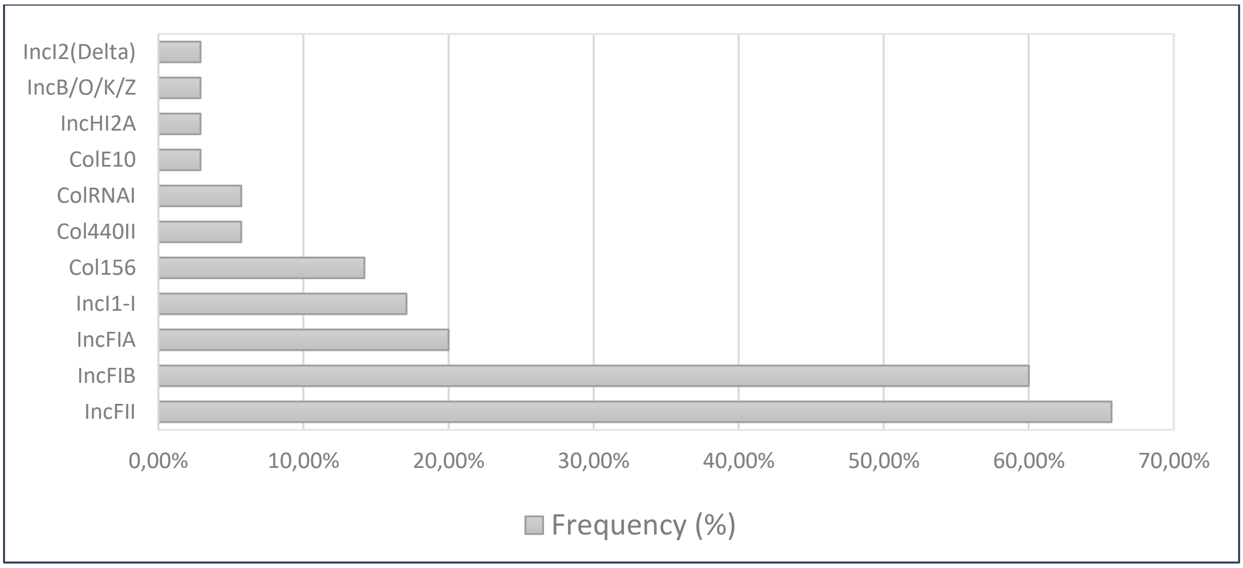


Figure 5: Bar graph of the plasmid Inc Types detected. Eleven plasmids were detected within the thirty-five *E. coli* isolates. On the left side of the chart, the y-axis indicates the percentage of each plasmid.

3.5 Discussion

A total of 35 generic *E. coli* isolates from various sources underwent characterization using WGS. The highly discriminative nature of WGS enables comparison of genetic relatedness among bacteria, even at the sub-species level, thus establishing it as the gold standard for typing bacterial isolates (Uelze et al., 2020). Moreover, WGS facilitates monitoring of antimicrobial resistance, virulence, and pathogenicity profiling, as well as source tracing, root cause analysis of contamination events, and enforcement of quality checks for bacterial analysis (Allard et al., 2018). The application of WGS in this study revealed that the generic isolates of *E. coli* belonged to five phylogroups: A, B1, B2, C and E, representing 2, 23, and 19 different pathotypes, serotypesand STs, respectively. Therefore, the information population structure of

E. coli circulating in the country of non-human clinical isolates provides critical information for epidemiological purposes.

Understanding the dynamics of the *E. coli* population is crucial for various reasons, including public health management and food safety. Phylogrouping, a method used to categorize *E. coli* strains based on genetic relatedness, plays a pivotal role in this regard. By classifying strains into different phylogroups, researchers can gain insights into how these strains are associated with the diseases they cause (Halaji et al., 2022). In the current study, the most prevalent phylogroups identified were A, B2 and B1, with the less common groups being C and E. This finding aligns with previous research indicating that phylogroups A, B1 and B2 are often the most predominant among *E. coli* strains found in various contexts, including clinical human cases, as well as in food and animal products (Pakbin et al., 2021; Aguirre-Sánchez et al., 2022; Zhao et al., 2022). Of particular note, phylogroup B2 has garnered attention for its association with extra-intestinal infections, indicating its potential role in causing a majority of such infections. This highlights the importance of understanding the distribution and prevalence of different phylogroups in various environments, as it can inform strategies for disease prevention and control.

In the current study, ExPEC strains were identified as the most predominant, collectively contributing at least 69% of the isolates. These strains were primarily classified into two pathotypes: APEC at 37% and UPEC at 31%. ExPEC strains are recognized for their propensity to cause diseases beyond the intestinal tract, including meningitis, urinary tract infections (UTIs), and sepsis (Tonini Da Rocha et al., 2021). In human populations, UPEC is infamous for its association with urinary tract infections, a prevalent and often recurrent medical issue (Whelan et al., 2023). Conversely, APEC is predominantly linked with avian colibacillosis, a significant bacterial infection affecting poultry industries worldwide with serious economic losses and welfare concerns in poultry farming (Kathayat et al., 2021). Most APEC strains identified in this study were found to exhibit diverse phylogenetic origins, with group A being the most prevalent. In contrast, UPEC strains primarily grouped within phylogenetic group B2. This pattern of phylogenetic distribution among APEC and UPEC strains has been consistently observed across multiple research studies (Malema et al., 2018; Tonini Da Rocha et al., 2021; Ghorbani et al., 2022). This phenomenon highlights the importance of phylogenetic characterization in understanding the epidemiology and pathogenesis of different pathotypes of *E. coli*, shedding light on the potential evolutionary and ecological dynamics that shape their distribution and virulence.

Interestingly, InPEC strains were found in lower numbers (31.3%) compared to ExPEC, with ETEC (23%), STEC (3%), and EPEC (6%) pathotypes being the most predominant. Intestinal diseases caused by InPEC play a significant role in burdening low-income countries with infections (Rojas-Lopez et al., 2018). The ETEC, EPEC, and STEC are known to be found in various animal species, including bovine, porcine, ovine or caprine, cats, and dogs, with bovine recognized as the major reservoir of these pathotypes. These three pathotypes are members of the DEC group; the majority of the samples thatwere identified were from animal meat that is used as a source of protein. It has been observed that ETEC and EPEC pathotypes tend to be the most frequently isolated pathotypes which are known to cause severe diarrhoea, this has been observed in South Africa were farming of domestic animals occurs (García et al., 2018; Abdalla et al., 2022). Presence of these pathotypes constitutes as a major public health risk since majority of these pathotypes displayed resistance to various antibiotics including tetracycline, cotrimoxazole, chloramphenicol, penicillin, sulfamethazine and amoxycillin (Eagar & Naidoo, 2017). It has been observed that high prevalence of these pathotypes is due to contamination of food products during processing (Tanih et al., 2015).

The occurrence of both InPEC and ExPEC pathotypes has been described in other studies (Masters et al., 2011; Omolajaiye et al., 2020). Although pathotypes were detected based on presence of their virulence gene markers, it is important to note that many virulence factors are not always exclusive to a single pathotype. Other virulence factors can be shared across various pathotypes. Therefore, in our study, pathotype assignment were interepreted based on predominant virulence genes detected. In this study, a higher prevalence of ExPEC may indicate that pathogenic *E. coli* is currently persisting in various sources of food and animals, suggesting a higher potential for health risks associated with foods consumed by humans. In South Africa, the presence of both InPEC and ExPEC has been detected in treated effluents (Omolajaiye et al., 2020), highlighting the importance of comprehensive surveillance and control measures to mitigate the spread of pathogenic *E. coli* strains in both environmental and food contexts.

In this study, all the *E. coli* strains were subjected to serotyping using an *in-silico* technique called serotype finder. In the past, O (somatic) and H (flagellar) antigens were utilized for serotyping *E. coli* isolates as either pathogenic or non-pathogenic (Joensen et al., 2015). Presently, there are 187 O antigens and 53 H antigens (Mare et al., 2021). The O antigen constitutes part of the lipopolysaccharide layer, which forms the outer membrane of *E. coli*, while the H antigen is the flagellar antigen responsible for *E. coli's* motility. Both are considered major surface antigens (Royer et al., 2021). In this study, 21 O antigens were identified: O2, O6, O9, O50, and O101 each contributed to 19% of the total O antigens, making them the most prevalent. This was followed by O8, O11, and O64, which accounted for 10%. Nineteen H antigens were detected, with H4 being the most prevalent, followed by H37 (21%), H16 (16%), and H9, H10, H19, and H27 each contributing 10.5% of the total H antigen. Predominant combinations of serotypes in this study included O6:H5, O9:H37 and O50:H4. Serotype O6:H5 has been linked to causing UTI in humans and domestic animals such as cats and dogs were O6 is found to be most predominant antigen (Johnson & Russo, 2018; Ksiezarek et al., 2022). There are not that many reports of serotype O9:H37 and O50:H4, but both have been isolated from poultry farms in China that keep goose and ducks (Hu et al., 2022; Shawa et al., 2022).

Among the predominant STs, all detected ST1858 strains were found to belong to the UPEC pathotype. One study conducted in Germany detected ST1858 from a catheter in a hospital with patients treated for UTI (Toval et al., 2014). Sequence type (ST)975 belonged to the ETEC pathotype and was found to belong to phylogroup A. ST2952 belonged to phylogroup B1 and belonged to the ETEC pathotype, while ST10 belonged to phylogroup A and was associated with both APEC and UPEC. ST95, a group B2, was associated with the UPEC pathotype. Sequence type (ST)88 belonged to phylogroups A and C and was associated with both APEC and ETEC. Sequence type (ST)120 belonged to groups A and B1 and was associated with both UPEC and ETEC.

In porcine, ST88 belonging to phylogroup C has been observed, while phylogroup A associated with ST88 has been isolated in avian and are often commensal strains (Abraham et al., 2014; Maluta et al., 2014). Isolates which belonging to A and B1 are often associated with commensal

E. coli which is often found in the gut of animals and humans, ST120 is considered a commensal that belongs to phylogroup A (Bahgat et al., 2024). In immunocompromised humans, phylogroup B1 has been associated with resistance to antibiotics and some clones of ST120 have been found to be susceptible (Das et al., 2013; Liu et al., 2015).

This study also detected various virulence factors associated with *E. coli* pathogenesis. Important virulence factor genes such as *fimH, papA/papG, hlyA, usp, cnf1, iutA, fyuA, afa, ompT, sfa, chuA, vat,* and *fycV* were detected in these isolates, and these genes are associated with the UPEC pathotype (Khairy et al., 2019). Virulence genes associated with the APEC pathotype were also detected, including *iroN, iss, ompT, tsh, hlyF, cvaC, iutA, sfa, papGII, fimH,* and *nlpI.* The ETEC pathotype presented genes such as *elt* and *est* genes, which are associated with diarrhoea in children in low-income countries. The ETEC isolates in this study possessed only the *est* gene (Singh et al., 2019; Higginson et al., 2022).

The STEC pathotype was detected in very low proportions and presented with the *stx2* gene; *stx1* was not detected in any of the isolates. The presence of the *stx2* gene is often associated with undercooked food or contaminated water, and the disease caused by STEC is characterized by haemolytic uremic syndrome, bloody diarrhoea, and haemorrhagic colitis (Alfinete et al., 2022). The EPEC pathotype presented with the *eae* gene, which is known for causing severe diarrhoea in children in developing countries, and mortalities in children have also been observed (Bolukaoto et al., 2021).

PathogenFinder, a webserver that can identify genetic characteristics linked to pathogenic and non-pathogenic isolates was used to determine if the isolates in this study have the potential to be pathogenic to humans. It is noted that some isolates may appear to be non-pathogenic, when introduced into favourable environments they may become opportunistic and become pathogenic (Cosentino et al., 2013). All the isolates in this study matched to various pathogenic of *E. coli* strains with various protein families identified which are linked to *E. coli* pathogenicity. This led to the prediction that all the isolates in this study have a very high probability (>0.90) of being human pathogens even though only a few indicated to be resistant to multiple antibiotics. In addition, these findings were supported by the virulence characteristics identified using virulencefinder (Montso et al., 2022; Wang et al., 2024).

Four CRISPR loci were found in *E. coli*: CRISPR 1, 2, 3, and 4. Depending on the presence of the corresponding *cas* genes, these loci are categorized as Type I-E (CRISPR 1 and 2) or Type I-F (CRISPR 3 and 4) (Xue & Sashital, 2019). In this study, CRISPR 1 was the most prevalent (86%) system, and it had a variable number of CRISPR arrays. CRISPR 3 loci were detected in very low numbers (11%). CRISPR2 and CRISPR4 systems were not detected. These arrays provide RNA molecules with a pattern to follow, enabling CRISPR-associated (Cas) proteins to precisely destroy viruses or bacteria upon re-infection. Only Cas1 and Cas2 are needed for the genetic recording of infections by obtaining spacers from DNA invaders. Despite the diversity observed between mobile genetic elements and CRISPR systems, nearly all known CRISPR-Cas systems share significant conservation in *Cas*1 and *Cas*2. Currently, CRISPR-*Cas* systems are classified into class 1 and 2, which are further subdivided into six types, type I-VI, and thirty-three subtypes (Makarova et al., 2016; Koonin et al., 2017; Murugan et al., 2017).

Clustered regularly interspaced short palindromic repeats (CRISPR-*Cas*) systems play a crucial role in restricting phage infection and proliferation, serving as a crucial component of bacteriophage resistance mechanisms. These mechanisms enable bacteria to detect and remove phage DNA upon reinfection by gathering and preserving genetic material from previous phage interactions (Oluwarinde et al., 2023). The CRISPR-cas3 protein was the most predominant (83 %) of all the CRISPR-cas proteins. Due to the helicase and nuclease activity offered by the CRISPR-*cas*3 protein, bacteria are well protected against phage attacks, increasing bacteria's chances of survival and prolonging its viability (Montso et al., 2022).

All the isolates presented with either the CRISPR or *cas* genes, there was no significant difference observed within the CRISPR-cas system between isolates which showed MDR (20%) and those that did not present with MDR. In this study, it was observed that 80% of the isolates were resistant to fewer antibiotics, while 20% of the were deemed to be multidrug resistant isolates, but no clear difference could be identified. It has been investigated in several studies with *E. coli* that there is no significant correlation observed between the CRISPR-cas system as well as antibiotic resistance (Touchon et al., 2012; Toro et al., 2014). Studies have shown that the type 1 CRISPR present in *E. coli* isolates has the ability to prevent a pathogen from acquiring plasmids that are resistant to antibiotics (Tao et al., 2022).

The type CRISPR type I-F constitutes 11% of the strains in the study, isolates belonging to this type were found to group with phylogroup B2 and the isolates were resistant to three or more antibiotics, other B2 phylogroups were found in CRISPR type I-E. Strains within type I-F are often associated with UPEC pathotype and are associated with causing urinary tract infections. Only one isolate within the type I-F which forms part of the MDR strains had plasmid IncFII, other isolates within this type did not present with any plasmid. This may support the finding that the type I-F CRISPR system does interfere with survival of plasmids responsible for antimicrobial resistance (Almendros et al., 2012; Aydin et al., 2017). The Type I-E CRISPR system is the most common type found in *E. coli*; however, it is considered inactive in this bacterium since its spacers are not suitable for targeting the viruses that bacteria encounter (Dion et al., 2024). In this study, only four isolates were found to contain both type I-E and I-F CRISPR. This is very uncommon, and this may be attributed to gene flow or ecological diversity of these bacteria (García-Gutiérrez et al., 2015). Although subtype I-A is uncommon in E. coli isolates, its presence was verified using CRISPRone online tool.

Twelve resistance genes were detected in this study, with the most predominant resistance gene being the *sitABCD* metal transport system gene, which builds the ABC transporter system responsible for causing resistance against hydrogen peroxide (Al-Mustapha et al., 2022; Sabença et al., 2024). The disinfectant *qacE* was found in 14% of the isolates on IS26, and it is responsible for resistance to quaternary ammonium compounds. Resistance against *qacE* has been observed in the food industry, where such chemical agents are used (Zou et al., 2014). The resistance gene *tet(*B*)* was detected in 23% of the isolates, with no MGE surrounding it. This gene confers resistance to tetracycline and encodes an efflux pump that plays a crucial role in the ability to cause resistance to antibiotics (Arredondo et al., 2019).

The high occurrence of resistance against tetracycline in these isolates may suggest misuse of antibiotics (Jaja, et al., 2020). Sulphonamide resistance genes *sul1* (9%) and *sul2* (6%) were also detected, along with beta-lactam resistance genes *blaTEM-1B* (9%) and *blaTEM-104* (3%). Aminoglycoside resistance genes *aadA1* (6%), *aadA2b* (6%), *aph(6)-Id*, and *aph(3'')-Ib*

(6%) were also observed. All the *blaTEM-1B* genes were found on the transposable element Tn2. Any form of *TEM-1* is encoded by either Tn1, Tn2, or Tn3 resistance transposons (Partridge & Hall, 2005; Stephens et al., 2020). Resistance *gene sul1, aadA1, qacE* were found either on insertion elements IS26 or Tn6196. Tn6196 were found on plasmid IncHI2A, IS26 is one of the elements known to speed up transmission of antimicrobial resistance genes in various communities of microorganisms (Behera et al., 2023).

Sulfonamide resistance genes, such as the *sul1* and *sul2*, are known to confer resistance to sulfonamide antibiotics. These genes can be transferred from commensal bacteria through mobile genetic elements to humans by consumption of meat-based products and, in turn, persist in more virulent bacteria in the human gut (Soufi et al., 2011). Sulfonamide antibiotics are known to inhibit the enzyme dihydropteroate synthase which is important for synthesis of folate in bacteria. Through the process of horizontal gene transfer, resistance to sulphonamides often occurs (Capasso & Supuran, 2014). Though in this study there was less prevalence of these genes, they are commonly known to be found in high prevalence in domesticated animals, humans and in environments were aquatic animals are cultivated (Jiang et al., 2019). Aminoglycosides are important antibiotics used for various treatments of bacterial infections in humans and animals. Resistance to this antibiotic occurs when the Aminoglycoside nucleotidyltransferases (ANTs) enzymes inactivate these antibiotics. Aminoglycosides molecules have different adenylation positions and there are five classes of ANT enzymes which are responsible for targeting these positions for inactivation. There are also Aminoglycoside phosphotransferases (APHs) enzyme with a specific focus on disabling the antibiotics ability to bind to bacteria (Van Duijkeren et al., 2019).

The mobile genetic elements and plasmids detected in these isolates exhibited significant diversity. Among them, the IncFII, IncFIB, Col440, and Col156 plasmid replicons were the most common. Notably, the IncFIB plasmid replicon is known to carry genes conferring resistance to cephalosporin, a phenomenon observed particularly among poultry farmers in Nigeria (Al-Mustapha et al., 2022), *in silico* analysis of antibiograms within this study did not detect any isolates resistance to cephalosporins or carbapenems. Additionally, the IS5 element is associated with *mcr-9* and *blaCTX-M55* resistance mechanisms (Mbanga et al., 2021). Mobile genetic elements identified in this study were mapped to specific contigs within our genomes, the localization of resistance genes with presence of insertion sequences such as IS26 suggests there is HGT and dissemination of antimicrobial resistance determinants.

The need for heightened surveillance is underscored by the emergence of antibiotic resistance in diseases affecting both humans and animals. *Escherichia coli* has long been recognized as a gram-negative indicator bacterium for antibiotic resistance due to its abundance in humans and various animal species, making it a potential vehicle for the spread of resistance genes between organisms (Chantziaras et al., 2014; Jaja, et al., 2020). Significant phenotypic resistance was observed, particularly for penicillin, cotrimoxazole, amoxicillin, hydrogen peroxide, and sulfamethazine. The fact that these antibiotics are easily accessible could explain the degree of resistance observed , especially in penicillin, in settings where it is used as prophylactics to treat diseases or as a growth promoter (Kazemnia et al., 2014; Deekshit & Srikumar, 2022). Among the list of priority pathogens that the World health organization has, *E. coli* forms part of the pathogens that are known to harbour cryptic resistance genes, these genes are present on the bacteria but do not exhibit the corresponding phenotypic resistance. This suggests that under certain conditions some genes remain silent but when transferred to a new host they become activated (Deekshit & Srikumar, 2022). This phenomenon of silent resistance genes has been observed in a number of studies relating *E. coli* (Zhao et al., 2001; Enne et al., 2006; Xu et al., 2014). Antibiotic resistance patterns may vary between animal populations and regions. Studies conducted in Africa have highlighted high levels of antimicrobial resistance in foods derived from animal products, particularly against antibiotics such as tetracycline and sulfamethazine, which are crucial in both human and veterinary medicine (Wesonga et al., 2010; Donkor et al., 2012; Alonso et al., 2017; Jaja, et al., 2020).

3.6 Conclusion

This study focused on a small subset of generic *E. coli*; thus, its findings may not be generalized for all *E. coli* pathotypes. However, this study adds to the knowledge that pathogenic *E. coli* can survive and be disseminated in the animal specimen, meat and meat products, which is a public health concern. The findings on the mobile genetic elements detected suggests a potential for horizontal gene transfer and spread of resistance genes of *E. coli* of which some may encode most of the significant virulence factors which affects humans. Some of the virulence factors identified include genes associated with adhesion, invasion and survival and they mainly belong to the ExPEC pathotype which are indicative of a pathogenic potential. The study also identified the prevalence of antimicrobial resistance genes, which encode for beta-lactams, tetracycline, sulphonamides, and disinfectants. These findings suggest that active surveillance on drug usage is required, and new strategies for control measures for *E. coli* infections are required. Results presented on this study were mainly from domestic animals that are used as a source of protein, these results demonstrated that there is a high prevalence rate of *E. coli* isolates, and it is mainly from animal-based food products, mainly poultry, porcine and bovine. More comprehensive studies are required to characterise resistant *E. coli* from animal-based food products with specific focus on monitoring virulence traits and the genetic traits associated with pathogenicity as well the risk factors posed by this bacterium.

3.7 Data Availability

Dataset analysed in this study is available in the NCBI Sequence Read Archive repository with accession number: PRJNA1126085.

CHAPTER 4

Genome-Based *In Silico* Analysis of the Structural and Functional Characteristics of the Type Three Secretion System (T3SS) and Core Effector Proteins in Enteropathogenic *Escherichia coli* (EPEC) Strains Isolated from Food-Producing Animals and Products of Animal Origin

Research article published from this chapter:

Malesa, Refiloe, Rian Pierneef, Thendo Mafuna, Kudakwashe Magwedere, Emmanuel Seakamela, and Itumeleng Matle. "Genome-Based *In Silico* Analysis of the Structural and Functional Characteristics of the Type Three Secretion System (T3SS) and Core Effector Proteins in Enteropathogenic *Escherichia coli* (EPEC) Strains Isolated from Food-Producing Animals and Products of Animal Origin." *Pathogens* 14, no. 11 (2025): 1099. doi: 10.3390/pathogens14111099.

Conference presentation from this chapter:

Biennial Congress of the South African Association of Veterinary Technologists held at 26 Degrees South Muldersdrift, on 9-10 October 2024: Oral presentation.

Published Article

pathogens

MDPI

Article

Genome-Based In Silico Analysis of the Structural and Functional Characteristics of the Type Three Secretion System (T3SS) and Core Effector Proteins in Enteropathogenic *Escherichia coli* (EPEC) Strains Isolated from Food-Producing Animals and Products of Animal Origin

Refiloe Malesa [1,2,*], Rian Pierneef [3,4], Thendo Mafuna [5], Kudakwashe Magwedere [6], Emmanuel Seakamela [1] and Itumeleng Matle [2]

[1] Bacteriology Division, Agricultural Research Council-Onderstepoort Veterinary Research, Onderstepoort 0110, South Africa; seakamelae@arc.agric.za
[2] Department of Agriculture and Animal Health, Science Campus, University of South Africa, Florida 1709, South Africa; ematlei@unisa.ac.za
[3] Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Pretoria 0001, South Africa; rian.pierneef@up.ac.za
[4] Centre for Bioinformatics and Computational Biology, University of Pretoria, Pretoria 0001, South Africa
[5] Department of Biochemistry, University of Johannesburg, Auckland Park, Johannesburg 2001, South Africa; tmafuna@uj.ac.za
[6] Center on Emerging Infectious Diseases, Boston University, 111 Cummington Mall, Suite 140, Boston, MA 02215, USA; gwedas@yahoo.co.uk
* Correspondence: malokotsar@arc.agric.za

4.1 Abstract

Enteropathogenic *Escherichia coli* (EPEC) is a significant diarrheagenic pathotype responsible for severe gastrointestinal infections, particularly in vulnerable populations. The aim of this study is to utilize genome-based *in silico* analysis to study the structural and functional characteristics of the Type III Secretion System (T3SS) and its core effector proteins in EPEC strains. Representative proteins were selected, with particular interest placed on EscV and EscD as major parts of the export apparatus and the basal body, and the EspA effector protein which forms the filamentous structure. Several *in silico*-based techniques were employed, revealing key structural proteins, core effectors, and adhesion-related proteins among the sequenced isolates. Of the 27 isolates analyzed, only 3 (11%) were found to carry LEE-encoded proteins associated with T3SS structural components (escV, escN, escD, and escU) and core effector proteins (espA, espD, espG, and eae). Structural predictions and Ramachandran plot validations suggested stability and potential functional conservation of T3SS proteins, with EscV and EspA selected for detailed 3D structural modelling. Insights into transmembrane domains, protein–protein interaction, and secondary structures were obtained, providing a comprehensive understanding of T3SS assembly and function. These findings suggest that the T3SS in EPEC consists of stable proteins that enable the system to remain functional. The structural and functional properties of the LEE genes encoding the T3SS in the EPEC pathotype represent promising targets for developing virulence blockers to disrupt the pathogenesis of a broad range of bacteria. This study is the first to report EPEC strains with functional T3SS in South Africa, emphasizing the importance of continued surveillance and molecular characterization of EPEC strains. The findings contribute to the development of targeted interventions to mitigate foodborne infections and improve public health.

Keywords: enteropathogenic *E. coli*, T3SS, protein domains, transmembrane domains, locus of enterocyte effacement, protein-protein interaction, protein structure

4.2 Introduction

A 40% reduction in the global average incidence of foodborne diarrhoeal diseases, calculated per 100,000 population, is one of the key indicators outlined in the World Health Organization (WHO) Global Strategy for Food Safety. This ambitious target, set to be achieved by 2030, reflects a global commitment to improving public health by reducing the burden of foodborne diseases. The reduction target was measured relative to the baseline global incidence estimates derived from the WHO Foodborne Disease Burden Epidemiology Reference Group (FERG), these are often based on data originated from 2010 to 2015. The reduction analysis specifically considers the incidence of diarrhoeal diseases caused by five major pathogens: Shiga toxin-producing *Escherichia coli* (STEC), *Campylobacter spp*., Enteropathogenic *E. coli* (EPEC), Enterotoxigenic *E. coli* (ETEC), and Non-typhoidal *Salmonella enterica*. These pathogens were selected due to their significant contribution to the global burden of foodborne illness, as highlighted in the WHO's 2021 report.

Enteropathogenic *Escherichia coli* (EPEC) is a significant diarrheagenic pathotype known for causing severe infant diarrheal outbreaks, particularly in developing countries where public health infrastructure may be limited (Shen et al., 2022; Lee et al., 2023). EPEC infections are typically transmitted to humans through the consumption of contaminated food, often because of inadequate sanitation and hygiene practices in food production and handling (Luna-Guevara et al., 2019). This mode of transmission poses a substantial risk to public health, especially in regions prone to diarrheal disease outbreaks.

Enteropathogenic *Escherichia coli* (EPEC) has been isolated from a diverse array of food sources, including meat, seafood, vegetables, fruits, and dairy products, underscoring the widespread nature of its contamination potential (Lindberg et al., 1998; Nataro and Kaper, 1998; Norazah et al., 1998; Silva et al., 2001; Araújo et al., 2002). Contaminated animal products pose a significant risk, given their frequent consumption, particularly among vulnerable populations (Bloomfield et al., 2007). Animals, particularly livestock, can act as reservoirs for the EPEC, thereby facilitating transmission to humans. As a result, surveillance of EPEC in animal populations and animal-derived products becomes essential for public health risk mitigation.

Given EPEC's pathogenicity and its health implications, there is a strong need for ongoing surveillance of its transmission pathways, virulence factors, and strains, as well as understanding the mechanisms by which these strains contrib-ute to disease outbreaks. Such efforts are critical for developing effective prevention strategies and safeguarding public health from EPEC-related infections (Ramos et al., 2020, Abdalla et al., 2022). Upon attachment to the host intestinal epithelium, EPEC employs a type III secretion system (T3SS) to inject effector proteins into the host cells. The T3SS is known as a transport apparatus primarily used by Gram-negative bacteria to insert effector proteins into host cytoplas-mic cells. This secretion system allows bacteria to manipulate host cellular processes, helping them establish infections, evade the immune response, and promote their own survival (Cepeda-Molero et al., 2020, Serapio-Palacios, and Finlay, 2020). The effector proteins delivered by the T3SS can disrupt cellular signalling, induce inflammation, and facilitate bacterial entry into host cells (Wagner et al., 2018).

Typical EPEC has the locus of enterocyte effacement (LEE) region, which is a 35.6 kb pathogenicity island that encodes the T3SS (Elliott et al., 2000; Franzin and Sircili, 2015; Tejeda-Dominguez et al., 2017). The LEE pathogenicity region is categorized into five gene clusters, named LEE1 to LEE5 (Franzin and Sircili, 2015), which harbour 41 genes that encode structural components of the T3SS, including six translocated effectors, regulators, and chaperones (Serapio-Palacios and Finlay, 2020). The esc genes found in the inner and outer membranes of the T3SS are encoded by the LEE1 to LEE3 gene clusters and contribute to the formation of the core structures of the system. The LEE4 con-sists of translocator proteins EspD and EspB, which are responsible for the assembly of the T3SS. The Tir effector is encoded by LEE5 (Tejeda-Dominguez et al., 2017). Effector proteins encoded by the T3SS are found within and outside the LEE, alongside non-LEE effectors (Cepeda-Molero et al., 2020).

Non-LEE effector proteins play a significant role in enhancing inflammation and inducing cell death within the epithelium (Cepeda-Molero et al., 2020). Non-LEE encoded T3SS effectors, such as Cif, NleA to NleH, EspL, EspJ, and EspO, perform distinct functions within the host cellular system to facilitate infection (Tejeda-Dominguez et al., 2017). The presence of these non-LEE effector proteins varies among different EPEC serotypes, with some effectors being absent in certain strains (Santos and Finlay, 2015).

The T3SS comprises diversified proteins with varied functions. For instance, escU, escV, escR, escT, and escS form the export apparatus, while escN, escL, and escO comprise the cytosolic ATPase complex. Additionally, the effector protein espA forms a filament, and espB and espD create the translocation pore. The assembly and translocation of effectors within the T3SS rely on these proteins (Díaz-Guerrero et al., 2021). Among the LEE-encoded proteins, escV contributes to the structural stability of the T3SS, escD forms part of the injectisome, allowing effectors to pass through, and escU is involved in the secretion of translocator proteins. Meanwhile, escN hydrolyzes ATP to provide the energy needed for effector secretion (Mitrović, Lezerovich, and Sal-Man, 2021).

The export apparatus is assembled from five highly conserved membrane proteins: escR, escS, escT, escU, and escV. EscV consists of two large domains: a cytoplasmic domain at the C-terminal end and an N-terminal pathway with seven to eight transmembrane domains (TMDs). The recruitment of T3SS sub-strates, chaperones, and proteins from the “gatekeeper” family to the T3SS apparatus is linked to escV and its homologs in both pathogenic and flagellar T3SS (Portaliou et al., 2017; Kuhlen et al., 2018; Wagner and Diepold, 2020; Mitrović, Lezerovich, and Sal-Man, 2021). EscD is a member of the single-span transmembrane proteins in the EscD/PrgH/YscD family, characterized by a larger C-terminal periplasmic domain and a smaller N-terminal cytoplasmic region (Deng et al., 2004; Ogino et al., 2006; Tseytin et al., 2018).

The use of genome-based analysis has allowed analysis of various gene clusters within the pathogenic islands found in bacteria that are responsible for the assembly and functioning of the T3SS (Zakaria et al., 2023). Structural prediction of the T3SS using *in silico*-based techniques, such as the Swiss-model, enables the construction of three-dimensional structures of proteins using their amino acid sequences, with these tools, comparative analysis of various pro-teins within the T3SS can be employed (Arnold et al., 2006, Katz et al., 2023). This study employed *in silico* genome-based approaches to predict the functional characteristics of T3SS structural proteins as well as core effector proteins found in EPEC strains isolated from food-producing animals and products of animal origin. The study specifically focused on core proteins, EscV and EscD only, as they play a major role in exporting proteins and formation of the basal body, while the EspA protein was selected due to its key role in the assembly of the filament and an effector protein that interacts directly with the host.

4.3 Material and Methods

4.3.1. Bacterial strains

A total of 27 previously identified EPEC strains (please refer to chapter 3) isolated from food-producing animals and products of animal origin in South Africa and imported meat products from North America and Europe were utilized in the current study (Table 3). These isolates were selected to represent a range of geographic regions, and the sample sources were selected to ensure that both hosts and environments where *E. coli* occurs are represented. DNA extraction, data pre-processing, and whole genome se-quencing of the isolates were performed as previously described in chapter 3.

4.3.2 Determination of LEE effector proteins

To screen for locus of enterocyte effacement (LEE), a query coverage of 85% and 95% identity was used as criteria for selecting T3SS genes for the study. DNA sequences were then analyzed using ORFFinder (https://www.ncbi.nlm.nih.gov/orffinder?tdsourcetag=s_pcqq_aiomsg, accessed on 10 September 2024) to confirm the presence of open reading frames.

The Virulence Factor Database (VFDB) (Chen et al., 2016) was used to identify the viru-lence factors associated with our isolates. Based on the VFDB results, various contigs were analyzed using MEGA v11 to pinpoint the locations where these genes were har-boured. Nucleotide sequences of each contig were selected and translated into protein sequences using MEGA v11, focusing on a subset of conserved proteins included in the study, escV, escD, escU, escN, espA, espD, espG, and eae.

Table 3: Metadata Summary of EPEC Strains (*n* = 27) utilized for whole genome analysis in this study

ACCESSION NO.	GEOGRAPHIC LOCATION	REGION	SOURCE OF ISOLATION
SAMN41920845	North America	North America	Raw poultry
SAMN41920846	South Africa	Mpumalanga	Raw poultry
SAMN41920847	Netherlands	Europe	Raw poultry
SAMN41920848	South Africa	Free State	Raw poultry
SAMN41920849	South Africa	Free State	Raw poultry
SAMN41920851	South Africa	Free State	Raw poultry
SAMN41920852	South Africa	Free State	Raw poultry
SAMN41920853	South Africa	Free State	Raw poultry
SAMN41920854	South Africa	Free State	Raw poultry
SAMN41920855	South Africa	Free State	Processed pork
SAMN41920856	South Africa	Free State	Raw lamb
SAMN41920858	South Africa	Free State	Raw pork
SAMN41920860	South Africa	Free State	Raw poultry
SAMN41920861	South Africa	Gauteng	Digestive system
SAMN41920862	South Africa	Gauteng	Processed beef
SAMN41920863	South Africa	Gauteng	Digestive system
SAMN41920865	South Africa	Gauteng	Digestive system
SAMN41920867	South Africa	Gauteng	Digestive system
SAMN41920868	South Africa	Gauteng	Digestive system
SAMN41920869	South Africa	Gauteng	Digestive system
SAMN41920870	Europe	Europe	Raw poultry
SAMN41920871	South Africa	Gauteng	Digestive system
SAMN41920872	South Africa	Gauteng	Digestive system
SAMN41920873	South Africa	Gauteng	Digestive system
SAMN41920876	South Africa	Gauteng	Water
SAMN41920877	South Africa	Limpopo	RTE beef

SAMN41920878	South Africa	Gauteng	Digestive system

4.3.3 Functional annotation of proteins

Further analysis involved examining the protein sequences using the Simple Modular Architecture Research Tool (SMART) based on Hidden Markov Models to identify and annotate signalling domains (Schultz et al., 1998). The Genomic SMART mode (https://smart.embl.de/smart/change_mode.cgi, accessed on 10 September 2024) was employed for protein analysis in this study, with searches conducted using default settings. Conserved protein domains were verified through the NCBI Conserved Do-main Database (CDD) (Wang et al., 2023), with input protein sequences processed under default conditions.

4.3.4 Validation of SMART

To assess sequence similarities for each protein, a BLASTp search was conducted for each protein sequence (Schultz et al., 1998). The query protein sequences were analyzed using default settings, with results considered valid if they exhibited a query coverage of 90% and a percentage identity exceeding 95%.

4.3.5 Analysis of Physiochemical properties

To assess protein stability and composition, the physicochemical properties of each protein were determined using the ProtParam tool on ExPASy (https://web.expasy.org/protparam/, accessed September 12, 2024). Each protein se-quence was analyzed under default parameters.

4.3.6 Protein-protein interactions

The STRING database (https://string-db.org/, accessed September 12, 2024) was used to construct a network of interactions between distinct known and predicted proteins. The analysis was performed using a full STRING network, with network edges adjusted to confidence mode, a minimum interaction score of 0.4, and a maximum of 20 interactions.

4.3.7 Determination of Transmembrane domains

Transmembrane domains (TMDs) were identified using the TMHMM v2.0 tool (https://services.healthtech.dtu.dk/services/TMHMM-2.0/, accessed on 12 September 2024)) for predicting transmembrane protein helices, with confirmation of TMDs per-formed using the PHOBIUS tool (https://phobius.sbc.su.se/, accessed on 10 September 2024). The topology of the sequences was then determined using Protter (https://wlab.ethz.ch/protter/start/, accessed on 10 September 2024). For each tool, protein sequences were input and analyzed using default settings.

4.3.8 Protein Structure Predicton

Protein structures of the components of the T3SS were predicted using Swiss-Model (https://swissmodel.expasy.org/, accessed on 14 September 2024). The target sequence was submitted to the server, which automatically searched for templates that were selected based on sequence identity. Model quality was assessed using the Ramachandran plot and MolProbity server version 4.2. To enhance the robustness and reliability of the predictions, additional tools were employed. The SOPMA tool (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html, accessed on 14 September 2024) was used for accurate prediction of the secondary structure of each protein.

4.3.9 Multiple sequence alignment

Multiple sequence alignment of the three (escV, escD, espA) proteins analysed ana-lyzed in this study were was aligned with homologous sequences retrieved from the NCBI database based on sequence similarities. The COBALT multiple alignment tool from NCBI was used for alignment (https://www.ncbi.nlm.nih.gov/tools/cobalt/re_cobalt.cgi, accessed on 29 September 2025) using default parameters. The alignment was visualized using the Jalview tool version 2.11.5 (Clamp et al., 2004).

4.4 Results

4.4.1 Analysis of virulence genes

Upon analyzing the VFDB results, various genes encoding T3SS proteins were identified. Notably, only 3 isolates (SAMN41920865, SAMN41920868, and SAMN41920872) carried LEE proteins associated with T3SS structural properties (Table S6); however, proteins from LEE1 through LEE5 were detected.

4.4.2 Determination of protein domains

Several LEE-encoding T3SS proteins were analyzed using SMART and subsequently confirmed with the CDD tool. Genes with a significant impact on T3SS function were selected and categorized based on their roles. The detected proteins included those as-sociated with structural formation, such as escV, escN, escD, and escU; core effector proteins related to pathogenicity, including espA, espD, and espB; and genes responsi-ble for adhesion and attachment, such as eae (Table 4). According to the SMART analy-sis, fifteen protein domains related to the T3SS were identified, with the majority clas-sified using the Protein Families Database (Pfam). Of the eleven proteins analyzed, the nleB1 protein could not be linked to a specific protein domain. Each protein was associ-ated with a single protein domain, except for the eae protein, which exhibited at least six distinct domains (Table 4).

4.4.3 Analysis of SMART validation

The selected proteins showed high similarity to known *E. coli* proteins in the NCBI BLASTp database. Each protein analyzed displayed 100% query coverage and over 99% sequence identity with existing protein sequences (Table 5).

Table 4: Summary of protein domain annotations and their functions

Proteins	**Protein Domain**	**Position**	**Function of each protein**
LEE			
Structural			
escV	Pfam:FHIPEP	26 to 663	It is important for translocation of effector proteins
escD	Pfam:Yop-YscD_ppl	157 to 404	It is involved in assembly and function of the T3SS
escU	Pfam:Bac_export_2	2 to 241	It is important for regulation and stabilization of the apparatus
escN	AAA	169 to 349	It hydrolyses ATP to generate energy required for operation
Pathogenicity			
Eae	LysM Pfam: IAT_beta BID_1 BID_1 BID_2 Pfam: Intimin_C	64 to 113 166 to 442 559 to 648 659 to 746 757 to 835 838 to 939	It is important for mediation of attachment of *E. coli* to the intestinal epithelium
Core effectors			
espA	Pfam: EspA	4 to 186	It is important for formation of pilus-like structures that facilitates movement of effectors
espD	Pfam: SseC	117 to 216	It forms the pore on the host membrane that allows effectors to move from bacteria into the host
espG	Pfam: EspG	15 to 397	It disrupts the host cell processes
non-LEE			
Cif	Pfam: CIF	81 to 215	Interferes with the host cell cycle
nleE	NleE_fam_methyl	13-168	Alters immunological responses and host cell signalling
nleB1	None		Alters proteins in host cells, affecting immune responses

Table 5: Results of SMART validation

Proteins	**Query coverage**	**E-value**	**Percentage identity**	**Reference. Accession no.**
LEE				
escV	100%	0.0	100%	WP_001037814.1
escD	100%	0.0	99.8%	ELP0616342.1
escU	100%	5×10^{-102}	99.4%	WP_063856070.1
escN	100%	0.0	99.8%	WP_000622546.1
Eae	100%	0.0	100%	WP_000627895.1
espA	100%	2×10^{-132}	100%	WP_000381555.1
espD	100%	0.0	100%	WP_000935768.1
espG	100%	0.0	100%	AAC31534.1
non-LEE				
Cif	100%	0.0	100%	WP_000652080.1
nleE	100%	1×10^{-119}	100%	WP_000609738.1
nleB1	100%	0.0	100%	WP_012578998.1

4.4.4 Physiochemical properties of encoded T3SS proteins

Analysis of the LEE encoding proteins revealed significant variability in protein lengths. The proteins eae, escV, escN, and escD were among the longest (675, 446 and 380 amino acids), while espA and escU were considerably shorter (192 and 250 amino acids). This variation in length correlated with molecular weight: eae, with 939 amino acids, had a molecular weight of 101,684.77 Da, whereas espA, with 192 amino acids, had a molecular weight of 20,534.02 Da. The Grand Average of Hydropathicity (GRAVY) results indicated that escU, escN, and espG are hydrophobic proteins, whereas eae, espA, espD, and escV are hydrophilic. The instability index further revealed that escU, escN, and eae are relatively stable proteins, with indices below 40, In contrast, espD and escV showed indices close to the instability threshold, but these proteins are still below 40 and are considered relatively stable (Table S7). All non-LEE encoding genes were associ-ated with proteins classified as hydrophilic based on their GRAVY results. These proteins exhibited variability in length and molecular weight, but their aliphatic indices were similar, ranging from 82 to 88.

4.4.5 Analysis of protein-protein interactions

Protein interactions were analyzed using the STRING database (Figure 6). All proteins studied showed high identity to the *E. coli O157* strain (EDL933) with a taxonomic identifier of 155,864. The protein networks for both LEE and non-LEE proteins demonstrated that they are associated with other known bacterial secretion proteins. The T3SS complex was found to include proteins such as FliI, EprK, EprS, and EscJ. Proton-transporting ATPase activity is associated with FliI, EivC, and EscN. Most of these proteins are involved in various biological processes, including protein secretion, protein transport, and other cellular functions. Specifically, the T3SS complex is present in EscC, EscJ, and EscN.

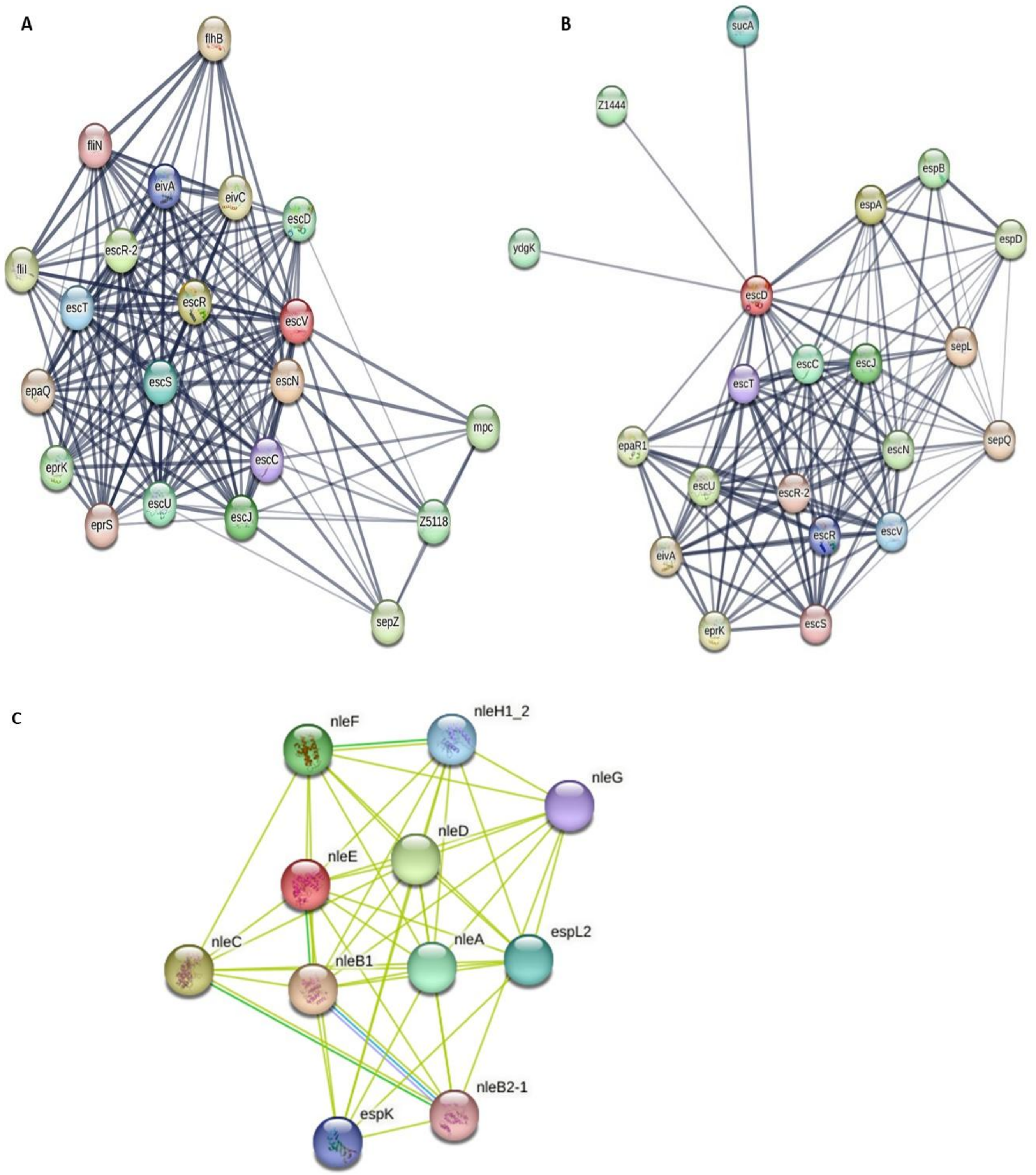


Figure 6: Protein-protein interaction network was generated using the STRING database. **(A)** represents the escV, escN proteins, **(B)** represents the escD, escU, espA, espD, **(C)** represents the non-LEE proteins, nleE, nleB1, nleG. The grey and yellow lines in between the nodes represent the interactions between proteins.

4.4.6 Analysis of Transmembrane domains (TMD)

Transmembrane domains (TMDs) were analyzed using Phobius, and the structural topology was determined with Protter. The structural protein escV was found to have seven TMDs, escU had two TMDs, and escD had one TMD, (Figure 7 and Figure S3). Additionally, N-glyco motifs were identified in these sequences, while signal peptides were absent. Analysis of the T3SS core effector proteins revealed that the espA protein had one TMD, and espD had three TMDs, (Figure 8). Non-LEE proteins were found not to contain any TMDs. Topology predictions in this study indicated that EscV protein contains seven transmembrane domains, which is consistent with the role that it plays as a core inner-membrane component of the T3SS. But, the EscD and EscU protein contained fewer transmembrane regions which reflects their structural roles in stabilizing the secretion apparatus within the bacterial inner membrane. Our predicted topologies of these proteins agrees with previously reported structural organization of T3SS components in enteropathogenic *E. coli*.

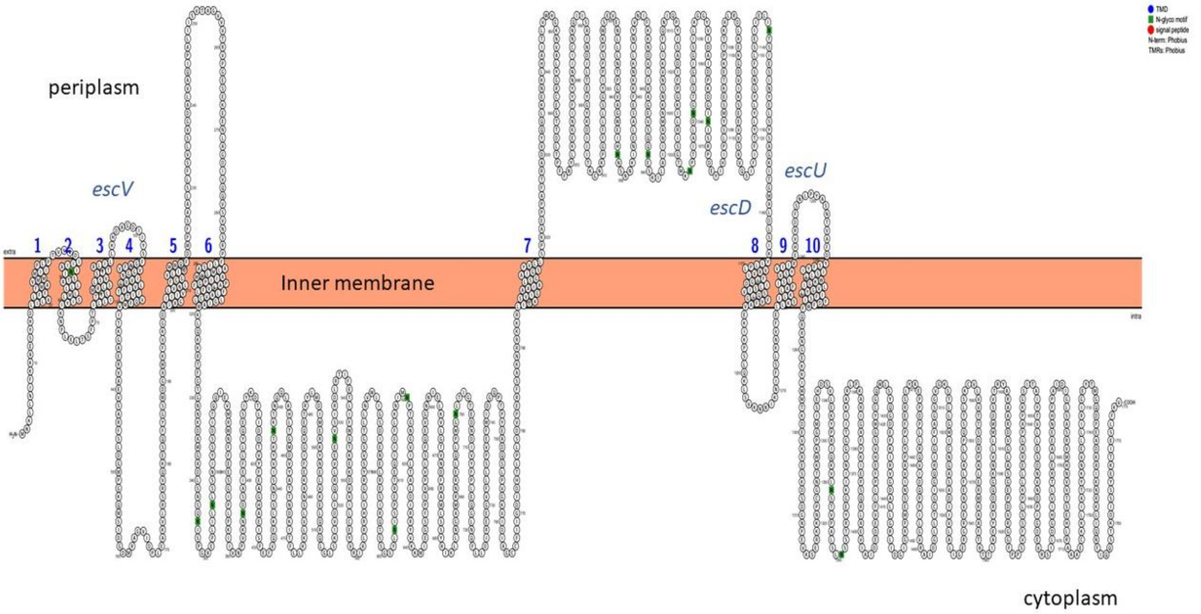


Figure 7: Predicted topology of T3SS proteins that are composed of transmembrane domains, structure determined using Protter. Structural protein escV consists of seven (1-7) TMD, protein escD consists of one (8) TMD, protein escU consists of two (9-10). Inner membrane approximated with an orange rectangle. Letter N in green represents the N-glyco-motif sequences.

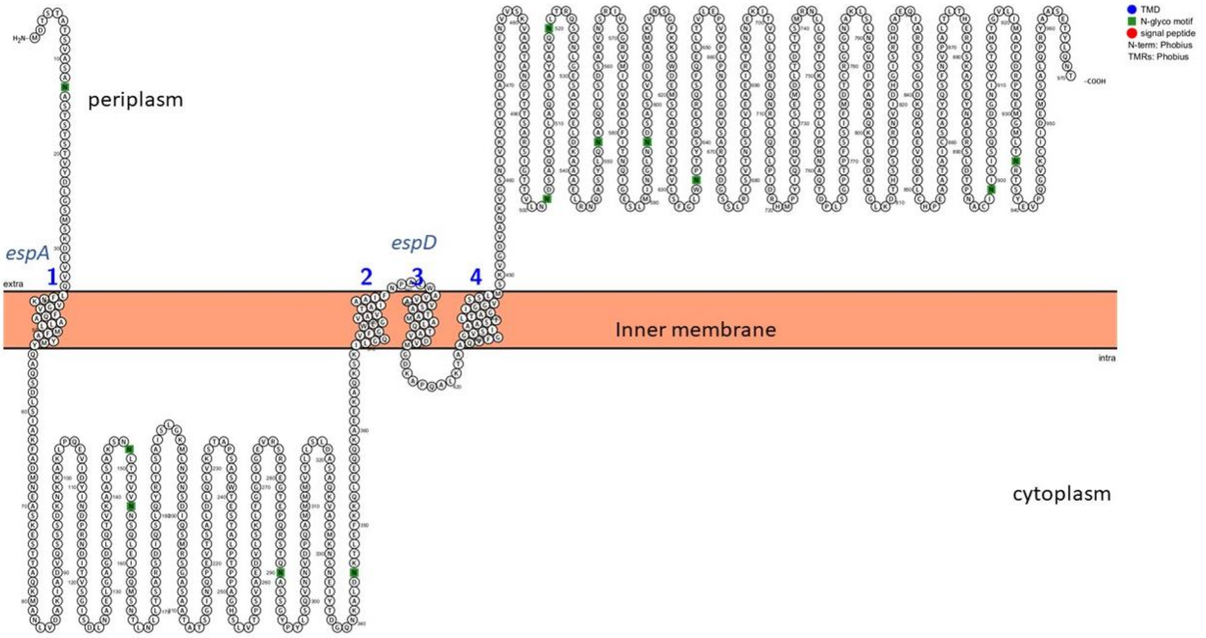


Figure 8: Predicted topology of T3SS Core effector protein espA consisting of one **(1)** TMD and protein espD consisting of three **(2-4)** TMDs. Inner membrane approximated with an orange rectangle. Letter N in green represents the N-glyco-motif sequences.

4.4.7 Predicted Protein Structures

Three proteins were selected for three-dimensional structural analysis: escV and escD to represent structural proteins and espA to represent core effector proteins (Figures 9 and 10). Based on sequence identity, coverage, and Ramachandran plot analysis, the most appropriate models were chosen. The template selected for escV was A0A3I8FV28.1A, which showed a sequence identity of 78.22% to the EscV/YscV/HrcV family T3SS export apparatus protein, and the structure consisted of a single subunit (monomer). The selected model for espA was 7khw.1.J, a translocon protein with 88.54% sequence identity, featuring an oligomeric state of homo-50-mer. For the escD protein, the most appropriate template selected was B1EHF2.1.A, which had 95.07% sequence identity.

4.4.8 Quality assessment of 3D-structures

The Ramachandran plot quality assessment for the escV and espA proteins yielded MolProbity scores of 1.38 and 1.35, respectively. The majority of residues observed in the Ramachandran plot were well-folded, with 96.14% of residues for escV and 92.36% for espA (Table 6).

Table 6: Validation summary of the escV and espA protein structures by the MolProbity software

Ramachandran Plot Analysis	**escV**	**espA**
MolProbity Score	1.38	1.35
Clash Score	2.41	1.40
Ramachandran Favoured	96.14%	92.36%
Ramachandran Outliers	1.19%	0.64%
Rotamer Outliers	1.35%	0.01%
C-Beta Deviations	0	177
Bad Bonds	0/5384	4/66450
Bad Angles	24/7285	350/89750
Twisted Non-Proline	5/652	50/8550

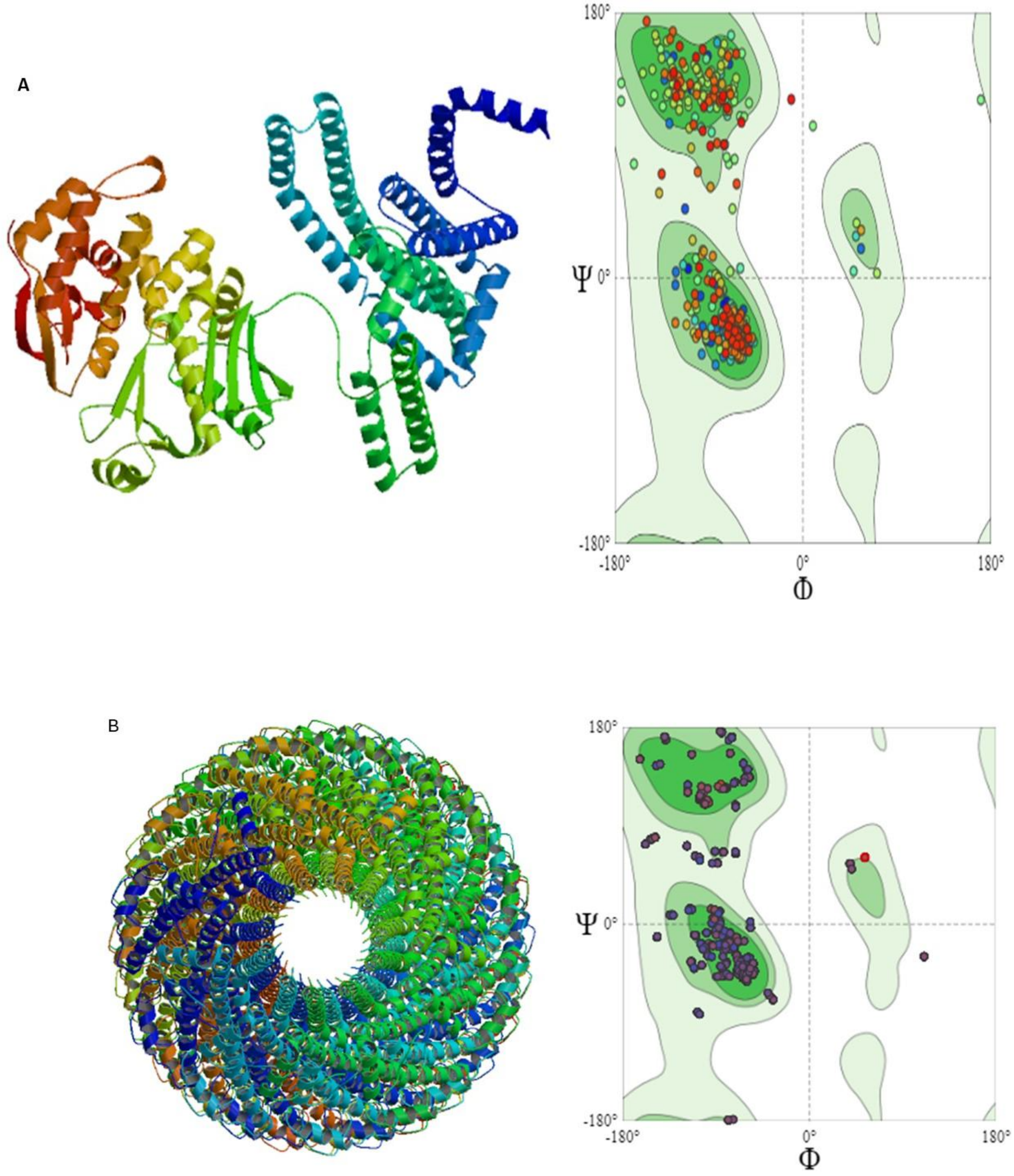


Figure 9: Three-dimensional structure of proteins and their Ramachandran plots, **(A)** escV protein, 96.14% of amino acids are on the favoured region (beta-sheet), 1.19% amino acids are outliers in this region **(B)** espA protein with 92.36% amino acid residues on the favoured region with 0.64% outliers.

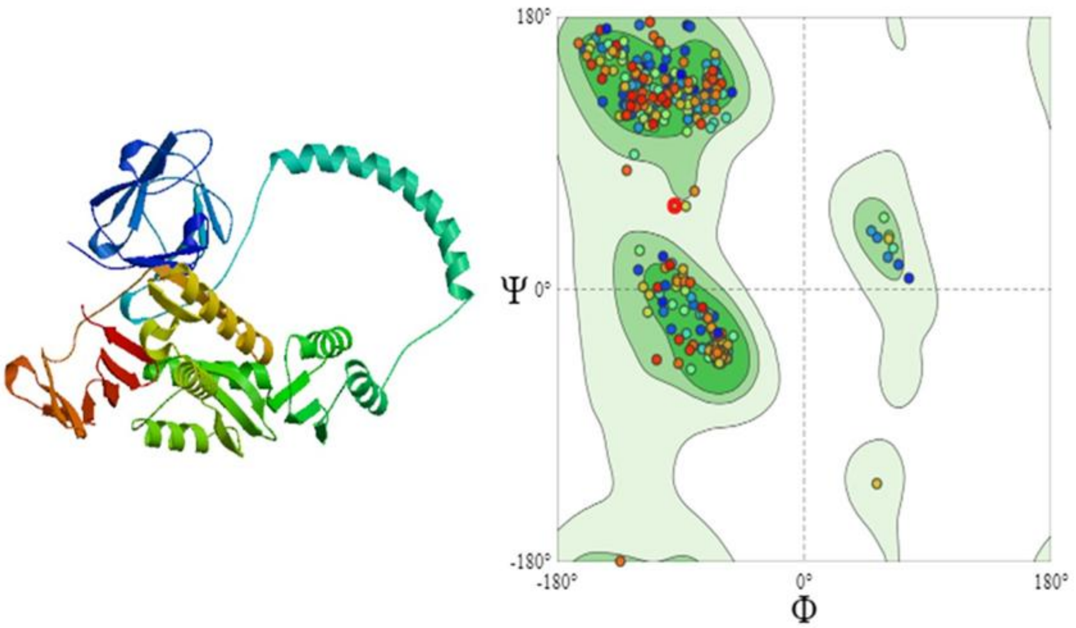


Figure 10: Three-dimensional structure of protein escD and its Ramachandran plot, 94.9% of amino acids are in the favoured region (beta-sheet), while 1.81% amino acids are outliers.

4.4.9 Secondary structure of the proteins

The analysis of the secondary structure of the proteins in terms of alpha-helix, beta-sheet, and coil states (Table 7) was conducted. Protein escV, consisting of 675 amino acids, exhibited 52.59% alpha-helix, 17.48% extended strand, 5.48% beta-turn, and 24.44% random coil. Protein espA, with 192 amino acids, had 68.75% in alpha-helix, 2.08% in extended strand, no beta-turn, and 29.17% in random coil.

4.4.10 Conservation of Proteins

Sequence alignment revealed high conservation across protein sequences included in this study. Conserved motifs were observed in various positions 430–490 (Figure S5). The C-terminal region also exhibited conserved segments, which indicates structural stability.

Table 7: Elements of the secondary structure

	escV		**espA**	
Structure	**No. of Residues**	**Percentage (%)**	**No. of Residues**	**Percentage (%)**
Alpha helix (Hh)	355	52.59	132	68.75
3_{10} helix (Gg)	0	0.00	0	0.00
Pi helix (Ii)	0	0.00	0	0.00
Beta bridge (Bb)	0	0.00	0	0.00
Extended strand (Ee)	118	17.48	4	2.08
Beta turn (Tt)	37	5.48	0	0.00
Bend region (Ss)	0	0.00	0	0.00
Random coil (Cc)	165	24.44	56	29.17
Ambiguous states	0	0.00	0	0.00
Other states	0	0.00	0	0.00

4.5 Discussion

To thoroughly analyze the various components of the type three secretion system detected in EPEC strains in South Africa, EscV, EscD, and EspA were intentionally selected. The EscV protein was selected to represent the inner membrane export apparatus, the EscD protein for its importance in maintaining the structure of the system, and EspA as the needle component, which is the extracellular filament of the system (Hotinger, Pendergrass, and May, 2021).

Enteropathogenic *E. coli* is one of the most important causes of diarrhoeal disease in both humans and animals, often resulting in high mortality rates in children, especially in developing countries (Bolukaoto et al., 2021). Typical EPEC strains often consist of the locus of enterocyte effacement (LEE), which is a region used for attaching and effacing to colonize the host's intestines. This region encodes for the T3SS (Whelan, McVicker, and Leo, 2020). This study was carried out to predict the structural and functional characteristics of the T3SS in EPEC strains using various in silico techniques. To the best of our knowledge, in South Africa, there is currently no published data that investigates the T3SS in EPEC strains.

In other countries, the occurrence of the T3SS has been identified in EPEC strains from various sources, and these isolates have been found to possess various virulence genes that are crucial for EPEC's pathogenicity. In African countries, there are very few studies that are sequenced, which are focused on characterizing EPEC strains. In a study conducted in Nigeria, it was found that one EPEC isolate contained a hemolysin gene, *ehx,* which is commonly found in EHEC pathotypes, and the presence of non-LEE genes was also detected (Akinlabi et al., 2025).

A study was conducted in 2023, which analyzed three diarrheagenic *E. coli* pathotypes (EAEC, EPEC, and STEC) in three African countries, Gambia, Kenya, and Mali, over a period of three years. This study highlighted atypical EPEC as the second most common pathotype to occur, which often results in high mortalities (Ochieng et al., 2023). In Zambia, Kenya, Mozambique, Rwanda, and Sudan, EPEC has been reported as the most prevalent pathotype (Saeed, Abd, and Sandstrom, 2015; Webale et al., 2020; Mwape et al., 2023).

In our study, several virulence-associated genes were detected among the *E. coli* isolates analysed, indicating the presence of potentially pathogenic strains across the sample matrices investigated.Moreover, a systematic review of *E. coli* virulence and antibiotic resistance profiles in various African countries revealed that in Africa, antibiotics are abused, and this has led to widespread resistance. Also, virulence determinants of *E. coli,* such as *eae*, *stx1*, *stx2*, *ast*, *fliC07*, *papC,* and *eagg* virulence genes have been found to occur across various *E. coli* matrices. This observation was found to be evident in South Africa and Nigeria, as they tend to test *E. coli* on various matrices (Hounkpe et al., 2023).

Protein domains play a crucial role in the functioning of the Type III Secretion System (T3SS). The flagellum/hypersensitive response/invasion protein export pore (FHIPEP) domain, which is part of the conserved export protein domains, has not been fully characterized, but it is thought to play a role in sorting substrates into the pore of the export apparatus to aid in the delivery of effector proteins (Barker et al., 2016). In proteobacteria, the Yop-YscD_ppl domain is part of the periplasmic domain of Yop proteins and forms part of the inner membrane of the injectisome (Kuebutornye et al., 2018). Bacterial transmembrane domains, which are essential for bacteria to move along host cellular membranes, contain proteins that target the host, playing a key role in infection (Gershberg, Braverman, and Sal-Man, 2021).

The effector proteins encoded by the LEE pathogenicity island in T3SS include espF, espG, espH, espZ, and map, along with translocator proteins espA, espB, and espD (Mainil, 2013; Xu et al., 2017). As one of the largest export apparatus proteins, escV, is responsible for controlling the export gate of the T3SS, while escD forms part of the inner membrane of the basal body and plays a critical role in secretion regulation within the inner membrane (Worrall, Bergeron, and Strynadka, 2013). Additionally, the ATPase activity of escN drives its interactions with T3SS effectors and chaperones (Sirisaengtaksin et al., 2023).

EscV consists of 675 amino acids and has a molecular weight of 75,150.08 Da, while escD has 406 amino acids and a molecular weight of 45,316.26 Da. The aliphatic index, calculated based on the percentage of alanine, valine, isoleucine, and leucine, was high for both proteins, indicating increased thermostability. Both proteins were also determined to be stable, with instability index values below 40 (escV 39.61, escD 39.83), suggesting they are less prone to denaturation under physiological conditions (Guruprasad, Reddy, and Pandit, 1990).

The theoretical isoelectric point (pI) of the proteins ranged from 4 to 10, suggesting electrical neutrality under normal conditions. The Grand Average of Hydropathicity (GRAVY) for escV was positive (0.355), indicating hydrophobic characteristics, while escD had a negative GRAVY value (-0.087), suggesting it contains hydrophilic amino acids (Kyte and Doolittle, 1982). Though hydrophobicity is only one factor, it plays a key role in understanding protein solubility and interactions (Yarabbi et al., 2020).

Protein interactions for escV and escD were analyzed using the STRING database, with *E. coli* O157 EDL933, a highly pathogenic strain, selected as a reference. The escV protein shared 97.9% identity with the reference strain, while escD shared 96.8%. Both proteins were found to interact with other related T3SS proteins such as escT, escR, escS, and escU, which form part of the conserved export apparatus between the flagellar system and the injectisome (Wagner et al., 2010; Wagner et al., 2018). EscV forms a pore in the inner membrane through which effectors are translocated, while escD is part of the base structure of the injectisome, interacting with escC and escJ proteins to maintain its structural integrity (Wagner et al., 2018).

The conserved domains of these proteins were identified using the Conserved Domain Database (CDD), while the transmembrane domains (TMD) were predicted with TMHMM v2.0. EscV was found to have seven TMDs, each with varying residues. This protein plays a critical role in recruiting T3SS chaperones, substrates, and other proteins, facilitating the assembly and function of the T3SS (Mitrović, Lezerovich, and Sal-Man, 2021). Both proteins 3D models were validated using Ramachandran plots, with over 90% of the residues falling within favoured regions, indicating high model quality (Nazarian et al., 2012). The MolProbity score for these structures was <1.8, further confirming their reliability (Cardoso et al., 2018).

In this study, secondary structure prediction of EscV and EscD from the strains analysed revealed the presence of beta strands, beta turns, alpha helices, and random coils, as predicted using the SOPMA tool. Secondary structures, including beta strands, beta turns, alpha helices, and random coils, are crucial for protein stability, serving as binding sites for other molecules and influencing protein function (Angamuthu et al., 2017). The predominance of these structural elements in EscV and EscD supports their structural and functional roles as core components of the Type III Secretion System, where protein stability and interaction are essential for secretion apparatus assembly. This detailed characterization of EscV and EscD enhances our understanding of their role in the T3SS and provides insights into the molecular mechanisms governing their function, consistent with previous reports describing the conserved structural features of T3SS basal body proteins.

In the present study, the LEE-encoded T3SS genes predicted in this study were found to be highly conserved with the secondary structure elements showing stable folding patterns which suggests that the key structural genes are well preserved. Although these structural characteristics were conserved across these isolates, minor variations in motifs were also observed, and this is consistent with previously described T3SS proteins in EPEC strains (Mitrović et al., 2021).

The identification and structural characterisation of key Type III Secretion System (T3SS) components, including EscV and EscD, highlight the functional importance of this system in the pathogenic potential of the strains analysed. The T3SS comprises various components with diverse functions; these include causing cell death, disrupting the host cytoskeleton, and manipulating cellular activities. Several studies have developed strategies to inhibit activities of the T3SS; some of these include suppressing activities of specific effector proteins, which then results in the system being less virulent (Hotinger, Pendergrass, and May, 2021). For Enteropathogenic *E. coli* (EPEC) and Enterohemorrhagic *E. coli* (EHEC) infections, a mouse model using Citrobacter rodentium was developed to inhibit the effector protein responsible for adhesion of the T3SS. Although other effector proteins were not inhibited and released into the host, this resulted in minimal infections (Deng et al., 2017; Hotinger, Pendergrass, and May, 2021). These findings support the relevance of the T3SS components identified in this study, as interference with T3SS-associated proteins can significantly reduce bacterial virulence.

Studies focused on therapeutics have targeted the use of the T3SS since some of the structural and functional characteristics of this system are able to facilitate the design of mechanism-based antibiotics; such studies are required to overcome antibiotic resistance (Deng et al., 2017). In the context of our study, the conserved structural features observed in EscV and EscD further support their suitability as potential targets for T3SS-directed therapeutic interventions. Other inhibitors of the T3SS include the use of chemical compounds such as 2-imino-5-arylidene thiazolidinone, which blocked the system as well as its virulence functions (Felise et al., 2008; Blasey et al., 2023). The first reported natural inhibitor of the T3SS was glycolipid caminoside A, which reduced the secretion of EPEC. Additionally, lactoferrin, a compound present in body fluids, has been shown to target T3SS virulence mechanisms. However, the limitation of some natural inhibitors is their associated toxicity to eukaryotic cells (Blasey et al., 2023), highlighting the need for alternative strategies informed by structural insights such as those indicated in this study.

Sequence alignment of proteins included in this study showed a high conservation in comparison to other *E. coli* EPEC strains, which contain the same proteins. Apparatus such as the escV and escD in the T3SS structure are known to be well conserved (Mitrović, Lezerovich, and Sal-Man, 2021). The espA protein remains conserved due to its helical assembly (Zheng et al., 2021).

4.6 Conclusion

This study represents the first in silico structural characterisation of LEE-encoded Type Three secretion system components in EPEC isolates which originated from food-producing animals and animal-derived food products in South Africa. Therefore, no experimental techniques were employed for this analysis. The findings revealed that the isolates with a functional T3SS identified in this study belong to the EPEC pathotype. Notably, the presence of the Locus of Enterocyte Effacement (LEE) pathogenic island underscores the virulence of these strains, affirming their pathogenic potential. The high-resolution structural analysis, along with robust validation data, indicates that the proteins in question are stable and likely to maintain their functional integrity.

Although there was a lack of experimental studies conducted for this analysis, to the best of our knowledge, this is the first report documenting EPEC strains that harbor a potentially functional T3SS in South Africa. This highlights the importance of continued surveillance and characterization of pathogenic *E. coli* strains within the region. Future research should involve a larger sample size, and experimental validations should be carried out to further support these predictions and to better understand the distribution and epidemiology of EPEC pathotypes carrying functional secretion systems across the country. Expanding the scope of study to include different regions and environmental sources would provide a more comprehensive picture of the pathogenic landscape and contribute to the development of targeted public health interventions.

CHAPTER 5

Shifting Genomes, Rising Threats: Adaptive Evolution and Antimicrobial Resistance of *E. coli* ST10 across African Ecologies

Research article submitted for publication from this article:

Malesa, R., Pierneef, R., Magwedere, K., Mafuna, T. and Matle, I. (2025). Shifting Genomes, Rising Threats: Adaptive Evolution and Antimicrobial Resistance of *E. coli* ST10 across African Ecologies. Manuscript number: GEN-D-25-01233, submitted to ScienceDirect, Genomics 06/08/2025. **Status: With Editor**.

Conference presentations

Second Para-Veterinary Professionals Conference held at AHA Gateway Hotel, Umhlanga, KwaZulu-Natal, on 21–23 October 2025: Abstract submitted.

5.1 Abstract

Escherichia coli sequence type 10 (ST10) is an emerging zoonotic lineage with global distribution and a high antimicrobial resistance (AMR) profile. Despite its epidemiological importance, the population dynamics of this disease across Africa remain largely uncharacterized. This study analyzed a total of 159 *E. coli* ST10 genomes, comprising isolates from this study (n = 4) and publicly available data from 19 other African countries (n = 155), encompassing human, animal, food, and environmental origins, using whole genome sequencing (WGS). The average nucleotide revealed high species-level similarity (mean 99.44%), yet a single-nucleotide polymorphism (SNP)-based phylogeny identified 122 discrete clusters, indicating substantial intra-lineage diversity. Over 71% of isolates were multidrug-resistant, and 29% harboured the extended-spectrum β-lactamase gene *blaCTX-M-15*. Plasmid replicon analysis revealed widespread IncFIB and ColRNAI plasmids, with strong associations to AMR determinants. Human-derived isolates predominated both phylogenetic clusters and resistance profiles, highlighting the centrality of humans as reservoirs. SNP divergence between intra- and inter-country comparisons suggested both local evolution and long-distance dissemination. These findings illuminate the evolutionary plasticity and public health risks of ST10 across Africa, supporting the urgent need for integrated One Health surveillance systems.

Keywords: *Escherichia coli* ST10, whole genome sequencing, antimicrobial resistance, Extended-spectrum beta-lactamase (ESBL), SNP-based phylogeny, Timescale phylogeny

5.2 Introduction

The global demand for *E. coli* testing has intensified across food systems, clinical settings, and environmental monitoring, driven by the bacterium's multifaceted role as both a commensal and a high-risk pathogen (Djordjevic et al., 2024). The problem of *E. coli* extends far beyond food safety, as *E. coli* in clinical settings remains the leading cause of urinary tract infections and a major driver of sepsis and neonatal meningitis (Manges, 2016; Ekici & Dümen, 2019). The rise of multidrug-resistant *E. coli* strains, particularly extended-spectrum beta-lactamase (ESBL)-producing variants, poses an escalating threat to global health, prompting the WHO to list it among the priority pathogens for new antibiotic development (Ramos et al., 2020; WHO, 2024). Concurrently, agricultural runoff, wastewater reuse, and sub-therapeutic antibiotic use in livestock have created environmental reservoirs that perpetuate the circulation and horizontal gene transfer of virulence and resistance determinants (Endale, Mathewos & Abdeta, 2023, Martak, Henriot & Hocquet, 2024).

Escherichia coli is a genetically heterogeneous species that inhabits the gastrointestinal tracts of warm-blooded animal and humans (Geurtsen et al., 2022, Naidoo & Zishiri, 2025; Omotoriogun et al., 2025), however, several phylogenetic lineages have evolved pathogenic characteristics, giving rise to diarrhoeagenic *E. coli* and extraintestinal pathogenic *E. coli* (ExPEC) (Geurtsen et al., 2022, Silva et al., 2024). Among the ExPEC, *E. coli* sequence types 10 (ST10)-a member of phylogroup A, has garnered increasing attention due to its widespread distribution, genomic plasticity, and multidrug-resistant (MDR) phenotypes (Santos-Neto et al., 2022). Unlike pandemic clones such as ST131 or ST648, which are commonly associated with clinical settings, *E. coli* ST10 has a broader ecological footprint and is frequently isolated from asymptomatic carriers, livestock, companion animals, food products, and environmental reservoirs (Schaufler et al., 2019; Kocsis et al., 2022; Messele et al., 2023).

Escherichia coli ST10 is a key representative of the clonal complex 10 (CC10), a globally distributed and evolutionarily successful lineage often implicated in zoonotic transmission and foodborne infections (Abdalla et al., 2025). It is among the most prevalent ExPEC sequence types, capable of colonising and persisting across host species without causing overt disease. This asymptomatic carriage facilitates the silent dissemination of antimicrobial resistance

(AMR) determinants across One Health domains. Sequence Type 10 strains have been linked to urinary tract infections, bloodstream infections, and neonatal sepsis, particularly in low-resource settings where diagnostic infrastructure is limited and empirical antibiotic use is widespread (Flament-Simon et al., 2020; Nielsen et al., 2021; Sands et al., 2021).

A defining feature of the *E. coli* ST10 lineage is its extensive repertoire of AMR genes, many of which are located on mobile genetic elements such as plasmids, integrons, and transposons (Iwan et al., 2024). These include β-lactamase genes (*blaCTX-M*, *blaSHV*, *blaOXA*, *blaCMY-2*), carbapenemases (*blaNDM*), and plasmid-mediated quinolone resistance determinants (*qnr* gene, as well as colistin resistance genes (*mcr*-type) (Aldea et al., 2022; Anyanwu et al., 2023). The increasing detection of *mcr-1* to *mcr-9* variants, particularly in Enterobacterales from animal sources, has raised global alarms given colistin's role as a last-resort antibiotic (Ling et al., 2020; Ejaz et al., 2021). Furthermore, ST10 frequently harbours IncF, IncI, and Col-type plasmids, which facilitate horizontal gene transfer and co-selection of resistance and virulence traits under antibiotic pressure (Li et al., 2023; Wolde et al., 2024).

Recent studies have suggested that E. coli ST10 may function as a genetic reservoir for AMR genes that can be mobilised into more virulent ExPEC clones (Massella et al., 2021; Cummins et al., 2022; Li et al., 2023). Its presence in poultry, swine, and cattle has been associated with the use of antibiotics for growth promotion and prophylaxis practices that are still common in parts of Africa (Anyanwu et al., 2023; Ogundare et al., 2024). Surveillance data from European and Asian countries have shown high rates of MDR and extended-spectrum β-lactamase (ESBL)-producing *E. coli* ST10 isolates in animal and environmental samples (Kuan et al., 2024), but equivalent data from the African continent remain sparse. The lack of harmonised AMR surveillance and whole genome sequencing (WGS)-based monitoring in Africa has limited our understanding of the population structure, resistance mechanisms, and zoonotic potential of *E. coli* ST10 in the region.

Addressing this knowledge gap is critical, particularly in the context of rising AMR-related mortality and the increasing recognition of the interconnectedness of human, animal, and environmental health. The One Health framework underscores the importance of characterising high-risk clones such as *E. coli* ST10 across reservoirs to mitigate their contribution to the

AMR burden. In this study, we perform a comprehensive genomic analysis of *E. coli* ST10 sequence data collected from 19 African countries, spanning human, animal, food, and environmental sources. We employ WGS, antimicrobial resistance gene profiling, plasmid replicon typing, and high-resolution SNP-based phylogenetics to unravel the population structure, evolutionary dynamics, and public health risks associated with this lineage. By integrating our isolates with publicly available African *E. coli* ST10 genomes, we provide the first pan-continental view of *E. coli* ST10 diversity and AMR potential, informing regional and global One Health strategies.

5.3. Materials and Methods

5.3.1 Study design, acquisition of publicly available whole genome sequence data and metadata

This study employed a comprehensive genomic analysis of *E. coli* ST10 isolates, comprising both newly analyzed strains (n = 4) from this study (Refer to Chapter 3; BioProject ID: PRJNA1126085) and publicly available datasets. Therefore, all publicly available ST10 genomes and associated metadata were retrieved from Enterobase (n = 327 ST10 genomes, accessed 06 May 2025). Rigorous quality and inclusion criteria were applied to ensure analytical consistency. Genomes were retained only if they were associated with a valid BioProject and had available NCBI Assembly accessions, along with complete metadata including data source, host origin, sample type, collection year, and country of isolation. Further filtering was based on assembly quality, retaining only genomes that met the following thresholds: (i) an N50 > 50 kbp, (ii) fewer than 200 contigs ≥200 bp, (iii) confirmed ST10 assignment, and (iv) ≥85% species identity (as annotated in EnteroBase). This refined selection process resulted in a dataset of 155 high-quality *E. coli* ST10 genomes spanning 19 African countries, which were downloaded from the NCBI Assembly database (accessed 06 May 2025).

To preliminarily assess genetic relatedness across the dataset, we constructed a Mash distance-based phylogenetic tree using mashtree v1.4.6 (Ondov et al., 2016), applying a genome size parameter of 4,716,019 bp, reflecting the average genome length of the included assemblies. These 155 public genomes were then integrated with four newly sequenced South African *E.*

coli ST10 isolates, resulting in a final dataset of 159 genomes used for all downstream comparative analyses (Table S8). The dataset included in this study are *E. coli* genomes originating from multiple sources comprising of food-producing animals, animal-derived food products, environmental samples and human isolates. This dataset reflects a clear One Health framework of isolates that are found to be reservoirs of pathogenic *E. coli.*

5.3.2 DNA extraction and whole genome sequence

Genomic DNA was isolated from overnight bacterial cultures using the High Pure PCR Template Preparation Kit (Roche, Germany), following the manufacturers protocol. Whole-genome sequencing of the resulting isolates was performed at the Biotechnology Platform of the Agricultural Research Council (ARC) in Onderstepoort, South Africa. For sequencing, DNA libraries were constructed using the TruSeq DNA Library Preparation Kit (Illumina, San Diego, CA, USA) and subsequently sequenced on the Illumina HiSeq 2500 platform, adhering to the manufacturer's recommended procedures.

5.3.3 Quality assessment, pre-processing, and De Novo assembly of whole genome sequence data

Whole-genome sequencing data for *E. coli* isolates underwent initial quality assessment and trimming using fastp version 0.23.4, employing default parameters to remove low-quality bases and adapter sequences (Chen et al., 2018). High-quality paired-end reads were subsequently assembled de novo using Shovill version 1.1.0, specifying the --assembler skesa flag while retaining all other default settings (https://github.com/tseemann/shovill). To ensure accuracy and integrity of the draft genome assemblies, multiple validation steps were employed. Assembly quality metrics were evaluated using QUAST version 5.3.0, incorporating the complete *E. coli* reference genome (*E. coli* ATCC 25922, NCBI Assembly: GCA_000486855.2) via the -r option (Gurevich et al., 2013). Additionally, genome completeness and potential contamination were assessed with CheckM version 1.2.2 using the lineage_wf workflow, which estimates taxon-specific completeness and contamination based on a phylogenetic marker gene set (Parks et al., 2015).

5.3.4 *In silico* serotyping

Using MLST, sequence types were identified using the CGE tool, which is accessible at https://cge.food.dtu.dk/services/MLST/. The MLST configuration (*Escherichia coli* #1) was selected and a minimum depth of 5X was used for alleles (Larsen et al., 2012). Data was retrieved and analysed based on the sequence types identified, confidence scores as well the allele profiles.

5.3.5 Core Single Nucleotide Polymorphism and phylogenetic analysis

To investigate genomic relatedness and population structure, pairwise Average Nucleotide Identity (ANI) was calculated for all 159 genomes in the *E. coli* ST10 dataset using skani v0.2.0 with --min-af 0, under otherwise default settings. The genome exhibiting the highest mean ANI across all comparisons was selected as the reference for SNP-based analyses. A core genome alignment was generated using Snippy v4.6.0 (https://github.com/tseemann/snippy), which performs reference-based SNP calling under default parameters. Putative recombinant regions were identified and masked from the core alignment using Gubbins v3.3.1, with rapid maximum likelihood tree inference specified using the --tree-builder iqtree-fast option. Variable sites from the recombination-filtered alignment were then extracted using snp-sites v2.5.1 with the -c flag, producing a recombination-free core SNP alignment suitable for high-resolution phylogenetic inference. Pairwise SNP distances were computed from the final alignment using snp-dists v0.8.2 and visualised as a heatmap with average-linkage hierarchical clustering using the pheatmap v1.0.12 package in R (Yu et al., 2017). This workflow was replicated for the 36 newly sequenced isolates from chapter 3 (excluding public genomes), enabling intra-study SNP distance comparisons.

5.3.6 Antimicrobial resistance determinant and plasmid replicon detection

To characterise the AMR profiles of the all *E. coli* ST10 dataset (n =159 genomes), each genome was screened using AMRFinderPlus v3.12.8 (Bortolaia et al., 2020). The AMR gene detection was performed using the latest available database version under default parameters unless otherwise specified. To detect plasmid replicons, ABRicate v1.0.1 (https://github.com/tseemann/abricate) was employed with the PlasmidFinder database (--db plasmidfinder), again using default parameters. The plasmid replicon screen was based on the November 2023 release of the PlasmidFinder database (4 Nov 2023) (Carattoli et al., 2014).

5.3.7 Pan-genome analysis

Functional annotation of all 159 genomes in the ST10 dataset was performed using Prokka v1.10 (Seemann, 2014). The resulting GFF3 files were used as input for pan-genome analysis with PPanGGOLiN v2.2.0 (https://github.com/labgem/PPanGGOLiN, (Gautreau et al., 2020), executed with the default parameters, the analysis included clustering genes based on their sequence similarity. Genes were partitioned into persistent (core), shell and cloud (accessory) genomes. The output files included modelling of the gene absence/presence graph. Genes present in 15% or fewer than 95% of genomes were classified as “shell”. The panRGP method (https://github.com/axbazin/panrgp_supdata, (Bazin et al., 2020) was used to predict regions of genomic plasticity, it uses the pangenome graph from PPanGGOLiN as an input to predict these regions. Using Heaps Law, the pangenome accumulation curve was further generated. The resulted graphics were generated using visualised with ggplot2 v3.5.1.

5.4. Results

5.4.1 Genome Dataset Composition and Geographic Distribution of *E. coli* ST10 isolates across Africa

Four *E. coli* ST10 isolates generated in the present study, originating from porcine faecal material and meat samples collected between 2016 and 2018, underwent WGS for comparative genomic analysis. These were combined with 155 high-quality publicly available ST10 genomes retrieved from EnteroBase, yielding a final dataset of 159 genomes. All isolates from public databases were sampled between 1994 and 2022 and originated from diverse ecological and host sources across 19 African countries.

Geographically, the public genomes were unevenly distributed across the continent. Tanzania contributed the largest proportion (17%), followed by South Africa (6%), Nigeria (15%), Kenya (14%), and Ethiopia (12%). Intermediate contributions were observed from Niger, Madagascar, and Ghana (each 4%), and Djibouti and Uganda (3% each). Smaller numbers were obtained from Zambia, Zimbabwe, Burkina Faso, and Gambia (2% each), while Benin, the Democratic Republic of Congo, Egypt, Rwanda, and Tunisia each contributed 0.6% of the dataset. Metadata detailing the source, sample type, and collection year for each isolate are provided in Table S8.

5.4.2 Genomic relatedness of *E. coli* ST10 based on Average Nucleotide Identity and Mash similarity

The genomic relatedness of the *E. coli* ST10 isolates was assessed using pairwise Average Nucleotide Identity (ANI) comparisons. A total of 10,948 genome pairs were evaluated using FastANI, all of which exceeded the 95% ANI species threshold, with values ranging from 98.72% to 100%, confirming the high taxonomic coherence of the dataset. The mean ANI value across all comparisons was 99.44% (SD = 0.17), indicating strong conservation at the nucleotide level across genomes. The interquartile range (IQR) spanned from 99.34% (25th percentile) to 99.54% (75th percentile), suggesting that at least 50% of genome pairs exhibited minimal divergence. The ANI distribution was right-skewed, with a concentration of values near the upper end of the scale (Figure 11), highlighting the close genetic relationships among the ST10 isolates despite their varied geographic and host origins.

Nucleotide composition similarity among genomes was evaluated using the Mash algorithm, which estimates genomic distances based on shared k-mer content. A distance matrix comprising 10,948 genomes was generated and subjected to hierarchical clustering to produce a heatmap illustrating the patterns of genomic relatedness (Figure 12). In the heatmap, varying shades of blue represent different levels of genetic similarity, with darker shades indicating greater identity. Genomes with Mash distances approaching zero are highly similar or identical, corresponding to the darkest shades in the heatmap.Additionally, a Mash distance-based phylogenetic tree was constructed to assess the genomic diversity and to identify any potential outliers among the 331 curated isolates included in this study. No distinct outliers were detected, indicating a consistent clustering pattern within the dataset (Figure S5).

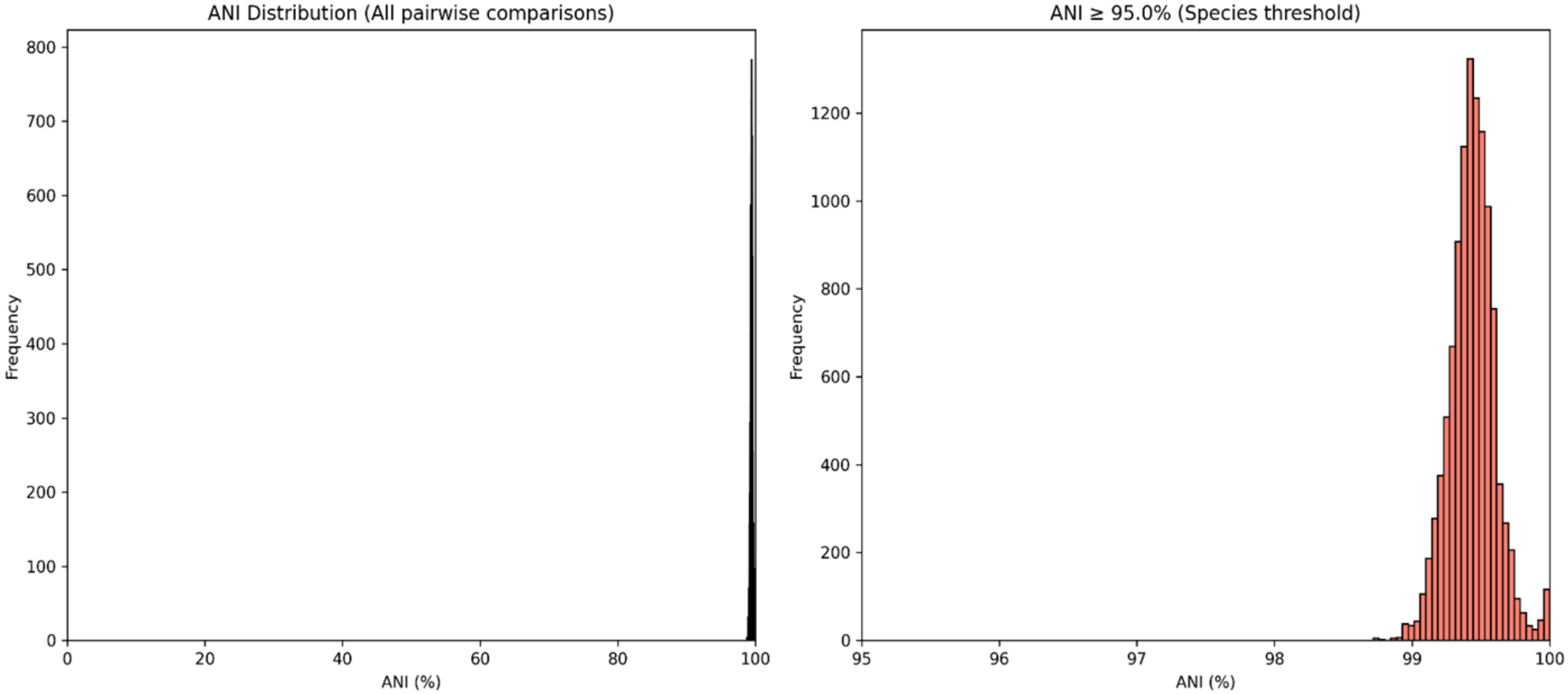


Figure 11: Histogram showing the pairwise ANI distribution among *E. coli* ST10 isolates. The dominant peak above 99% reflects high genomic similarity across the dataset, consistent with a single lineage

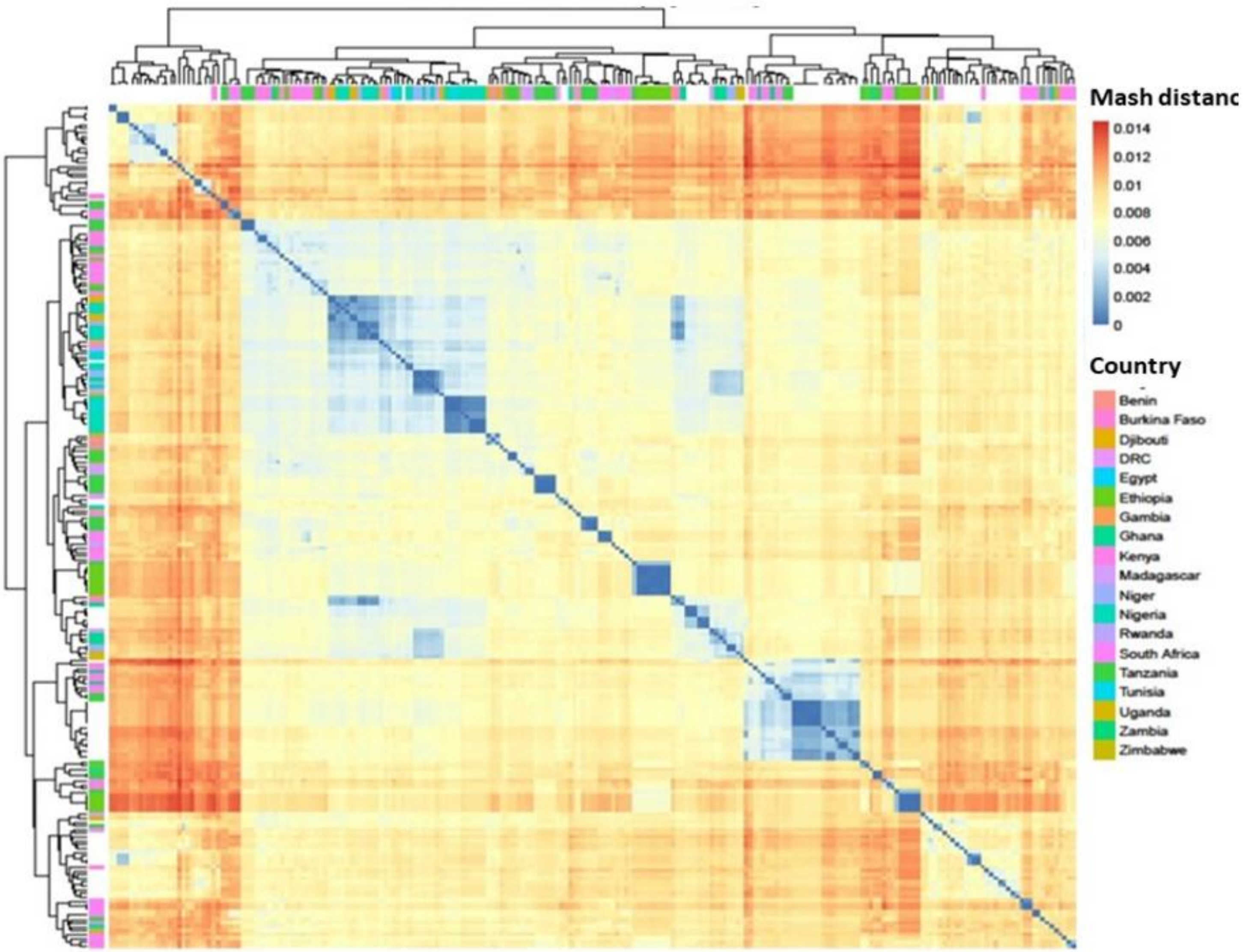

Mash distanc
0.014
0.012
0.01
0.008
0.006
0.004
0.002
0
Country
Benin
Burkina Faso
Djibouti
DRC
Egypt
Ethiopia
Gambia
Ghana
Kenya
Madagascar
Niger
Nigeria
Rwanda
South Africa
Tanzania
Tunisia
Uganda
Zambia
Zimbabwe

Figure 12: The heatmap illustrates a pairwise genomic distances calculated using Mash, where smaller values indicate greater genomic similarity between isolates. Dark blue shading represents near-identical genomes (Mash distance = 0), while increasing genomic divergence is shown by lighter colours, with the darkest brown corresponding to the highest observed Mash distance (0.014). The dendrogram reflects hierarchical clustering of the isolates based on genomic similarity, with branch heights indicating relative levels of divergence. Colour bars displayed above and to the left of the heatmap represent the countries of origin of each isolate, enabling visual assessment of geographic patterns within the ST10 dataset.

5.4.3 Phylogenetic Structure of *E. coli* ST10 Based on Core Genome single nucleotide polymorphisms

High-resolution phylogenetic analysis based on core genome single nucleotide polymorphisms (SNPs) revealed substantial genetic diversity within the *E. coli* ST10 lineage. Single nucleotide polymorphisms calling and phylogeny construction identified 122 distinct SNP clusters among the 157 isolates, two isolates from our study, SRR29469971 and SRR29469969 were excluded since they failed QC on SNP analysis. Majority of clusters consisted of single isolates. This pattern suggests significant microevolutionary divergence across the lineage (Figure 13). Only a small subset of clusters specifically clusters 5, 17, 20, 24, 30, 35, 42, 51, 55, 57, 68, 72, 73, 75, 107, 108, 112, and 114 contained between two and five closely related genomes (Figure 13). These low-SNP-distance clusters accounted for 33% of the total isolates. Most of the closely related isolates were derived from human sources (56%), while the remainder were distributed across poultry, swine, ovine, dairy, fish-associated and environmental samples (44%).

Among the South African *E. coli* ST10 isolates analysed in this study, two isolates SRR29469960 (from a meat sample collected in 2016) and SRR29469978 (from porcine faeces collected in 2018) formed distinct phylogenetic clusters (Cluster 106 and Cluster 113, respectively). These isolates exhibited significant genetic divergence from all other isolates in the dataset and did not cluster with any representatives from other African countries, suggesting the existence of potential region-specific or niche-adapted lineages. The average single nucleotide polymorphism distances between these two isolates and the remaining dataset were 10,276.5 and 9,206.7, respectively, underscoring their substantial genomic distinctiveness.

Country-level clustering patterns revealed that Tanzania (21%) and South Africa (20%) accounted for the largest proportion of SNP clusters, followed by Nigeria and Kenya (18% each), and Ethiopia (16%). Countries contributing the fewest SNP clusters included Niger (6%), Madagascar and Ghana (5%), Uganda (4%), Djibouti (3%), and Burkina Faso, Gambia, Zambia, and Zimbabwe (1.6%). The highest intra-country SNP diversity was observed in Kenya, Tanzania, South Africa, Ethiopia, and Nigeria together representing over 70% of the total SNP clusters. In terms of sample origin, human-derived isolates constituted the majority of clusters (56%), followed by those from poultry (17%), swine (9%), environmental sources including water (6%), and a mixture of other sample types (8%).

Hierarchical clustering of the Mash distance data revealed five distinct phylogenetic clusters among the *E. coli* ST10 isolates (Figure 14). All the isolates from our study were exclusively located within the brown cluster, which also contained many other isolates, suggesting a common lineage or conserved genetic background. The red and orange clusters consisted of isolates originating from South Africa, Kenya, and Tanzania, collected over a span of ten years (2009–2019), indicating long-term regional persistence of certain ST10 lineages. These clusters included isolates from various hosts: the orange cluster (Clusters 24, 99, 6, and 26) was predominantly composed of porcine isolates, with one avian isolate, whereas the red cluster (Clusters 112, 47, 62, and 88) primarily comprised human isolates and one environmental isolate. The consistent genetic relatedness within each cluster, despite host and temporal variation, points to the potential for long-term maintenance and circulation of specific ST10 strains across ecological niches.

The purple cluster (Clusters 100, 78, 13, and 85) was primarily composed of porcine isolates collected in South Africa in 2019, accompanied by a human isolate (Cluster 29) from the same year, suggesting possible interspecies transmission. The blue cluster (Cluster 5) also included porcine isolates from South Africa collected in 2019, further supporting the notion of host-associated clustering. Collectively, these patterns highlight potential host and region-specific structuring of *E. coli* ST10 populations, as well as evidence of persistence and possible cross-species transmission events within and across geographic regions.

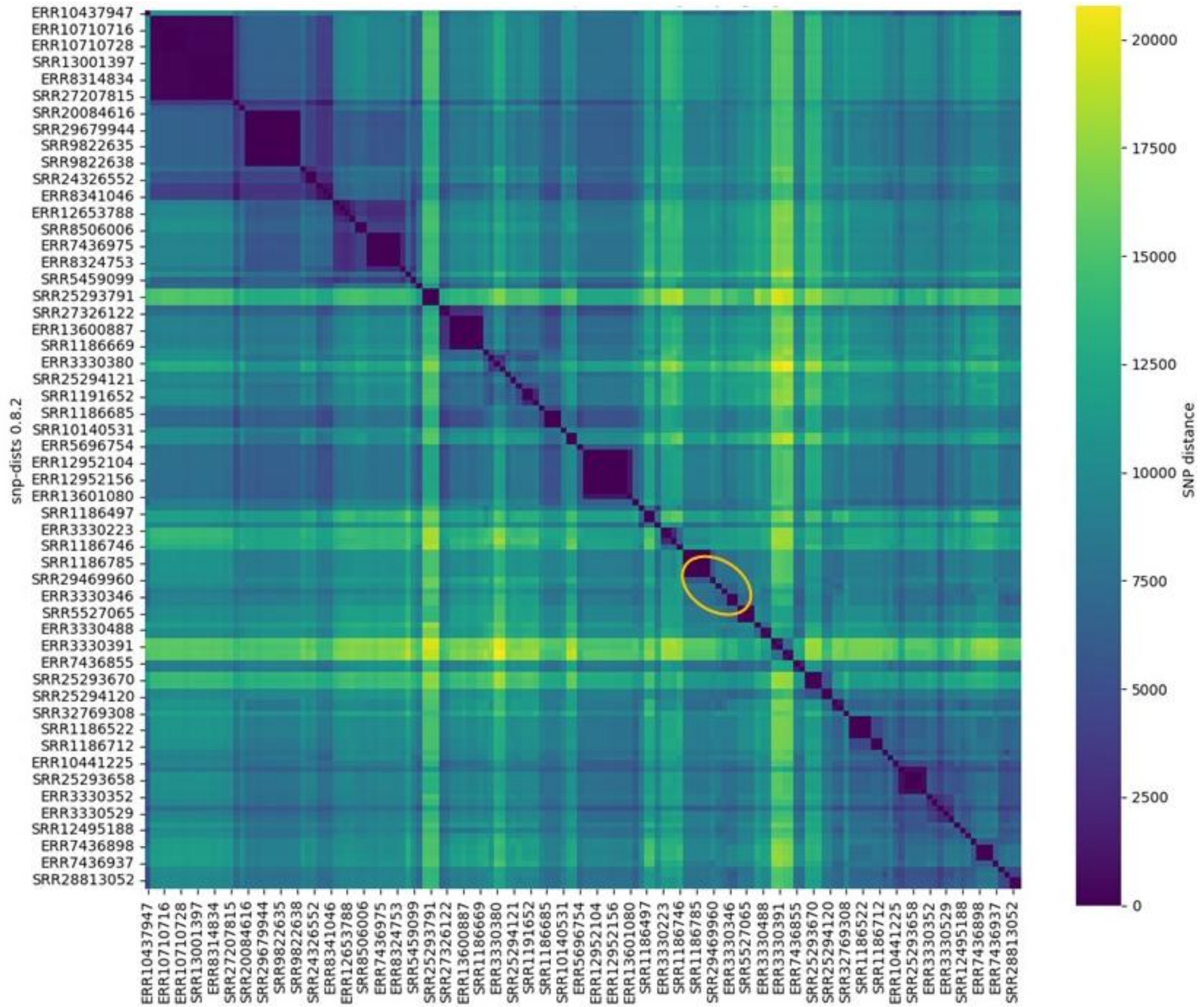


Figure 13: A pairwise SNP-based heatmap of *E. coli* ST10 isolates, each row represents an isolate in this study, the SNP distance is reflected by the intensity of shading, with clusters indicated along the axes. Though there are differences in these isolates, a close genetic relationship was observed. A cluster of isolates from our previous study is indicated with yellow circle on the heatmap.

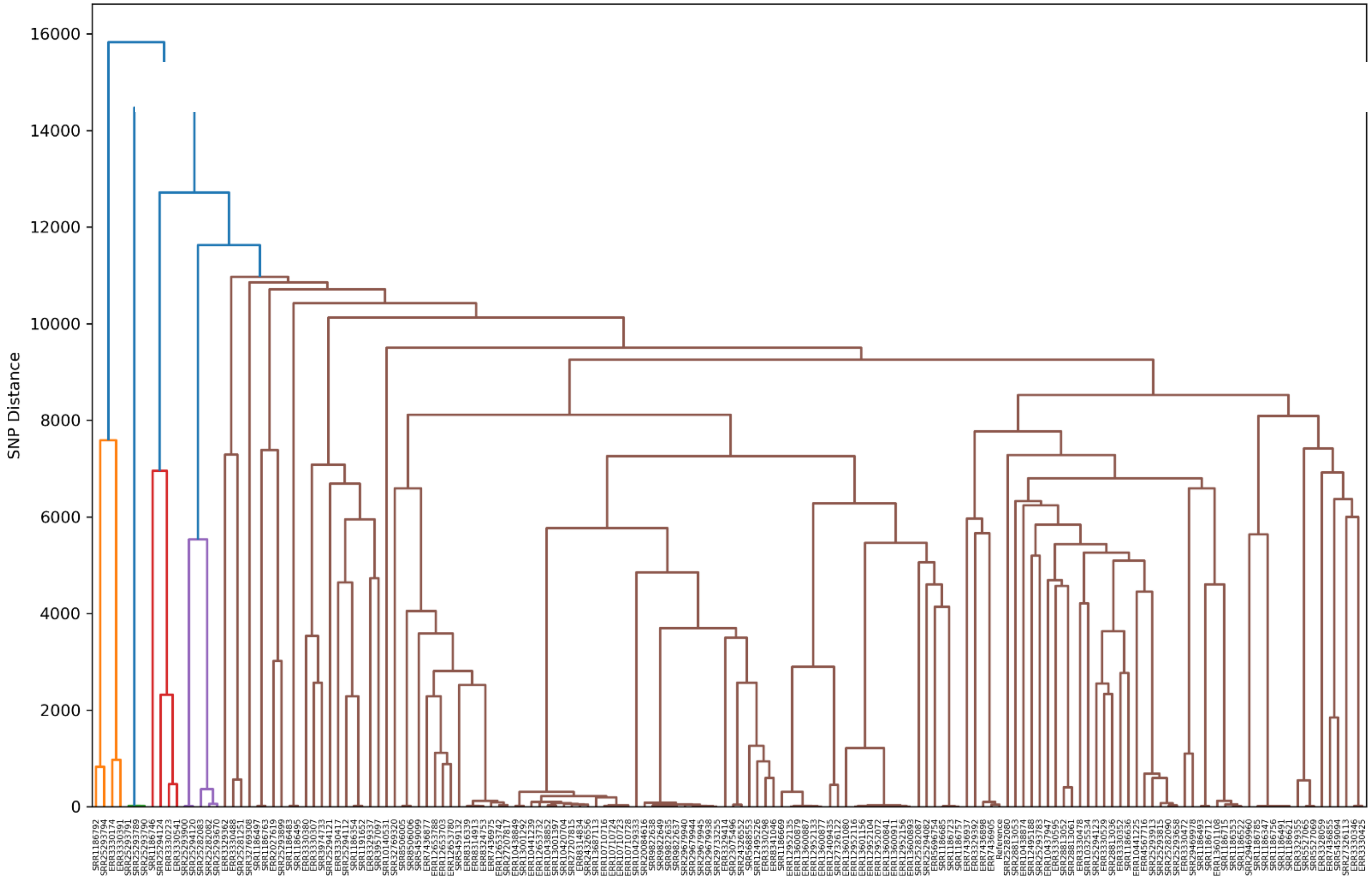


Figure 14: Hierarchical clustering dendogram of *E. coli* ST10 isolates produced using SNPs produced in this study. The genetic clustering relationship between 157 isolates was determined. Different colours on the branches represent the differences observed between these clusters, five clusters are observed. The brown cluster is the most dominant, while the blue colour indicates isolates that are genetically distant from the rest, the orange, red and purple indicate isolates that are closely related but still separated from each other.

5.4.4 Intra and Inter-SNP distance comparisons

To assess the phylogenetic relatedness of *E. coli* ST10 isolates within and between African countries, pairwise SNP distances were stratified into intra- and inter-country comparisons. A total of 117 intra-country genome pairs exhibited markedly low SNP divergence, with a mean pairwise distance of 23.3 SNPs (range: 0–75), indicating the presence of closely related strains within national boundaries. In contrast, the 12,286 inter-country genome pairs demonstrated significantly higher genomic divergence, with a mean SNP distance of 9,610.2 and a range spanning from 53 to 21,056 SNPs (Figure 15). The results also showed underscore the strong geographic signal in the population structure of ST10 across the continent. Most intra-country comparisons with low SNP distances were associated with isolates from Nigeria, while relatively few closely related pairs originated from countries such as Uganda and Niger.

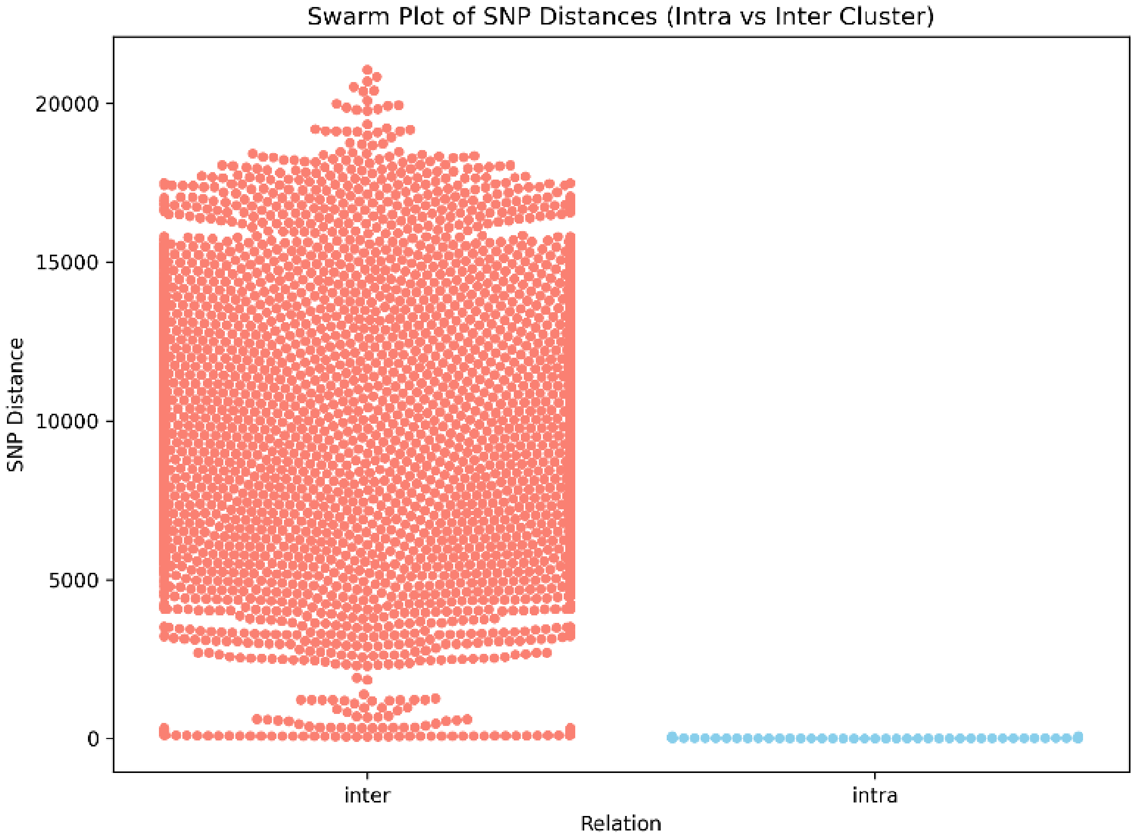


Figure 15: Distribution of pairwise SNP distances comparing intra- and inter-country *E. coli* ST10 isolate relationships. The clear divergence in SNP range reflects both localized clonal spread and broader continental-scale genetic diversity.

5.4.5 Pan genome analysis

5.4.5.1 Genome composition

A pangenome was computed using the PPanGGOLiN tool was used to determine the gene content of the genomes included in this study. This tool uses a graph model to present the pangenome content such as gene families. Furthermore, this pangenome is divided into different segments based on the presence or absence of genes. A total of 16,311 gene families were detected, where 3572 (21%) persistent (core) genes which are found in almost all genomes, 2898 (17%) conserved genes (shell) were detected in of our genomes and 9841 (60%) low frequency genes (cloud) were detected in our genomes (Figure 16).

5.4.5.2 Regions of genome plasticity

Regions of genomic plasticity (RGPs) were identified using the PanRGP method, a computational tool designed to detect genomic islands within bacterial genomes. Across all genomes analysed in this study, a total of 9,446 RGPs were detected. The number of RGPs per genome varied substantially, ranging from 4 to 194 regions, with an average of 13.6 genes per region. The length of these regions ranged from 3,010 bp to 183,755 bp, with a mean length of approximately 13 kb. This wide variability in both the size and gene content of RGPs highlights the remarkable genomic plasticity of the *E. coli* ST10 lineage and underscores the significant role of HGT in shaping its genomic architecture. The pangenome accumulation curve demonstrated that, although the number of genes continues to increase with the inclusion of additional genomes, the curve begins to plateau, suggesting a relatively stable core genome. Meanwhile, the distribution of RGP lengths revealed that many of these regions are relatively short (Figure 17), consistent with the presence of numerous small genomic islands that may contribute to phenotypic diversity and adaptation.

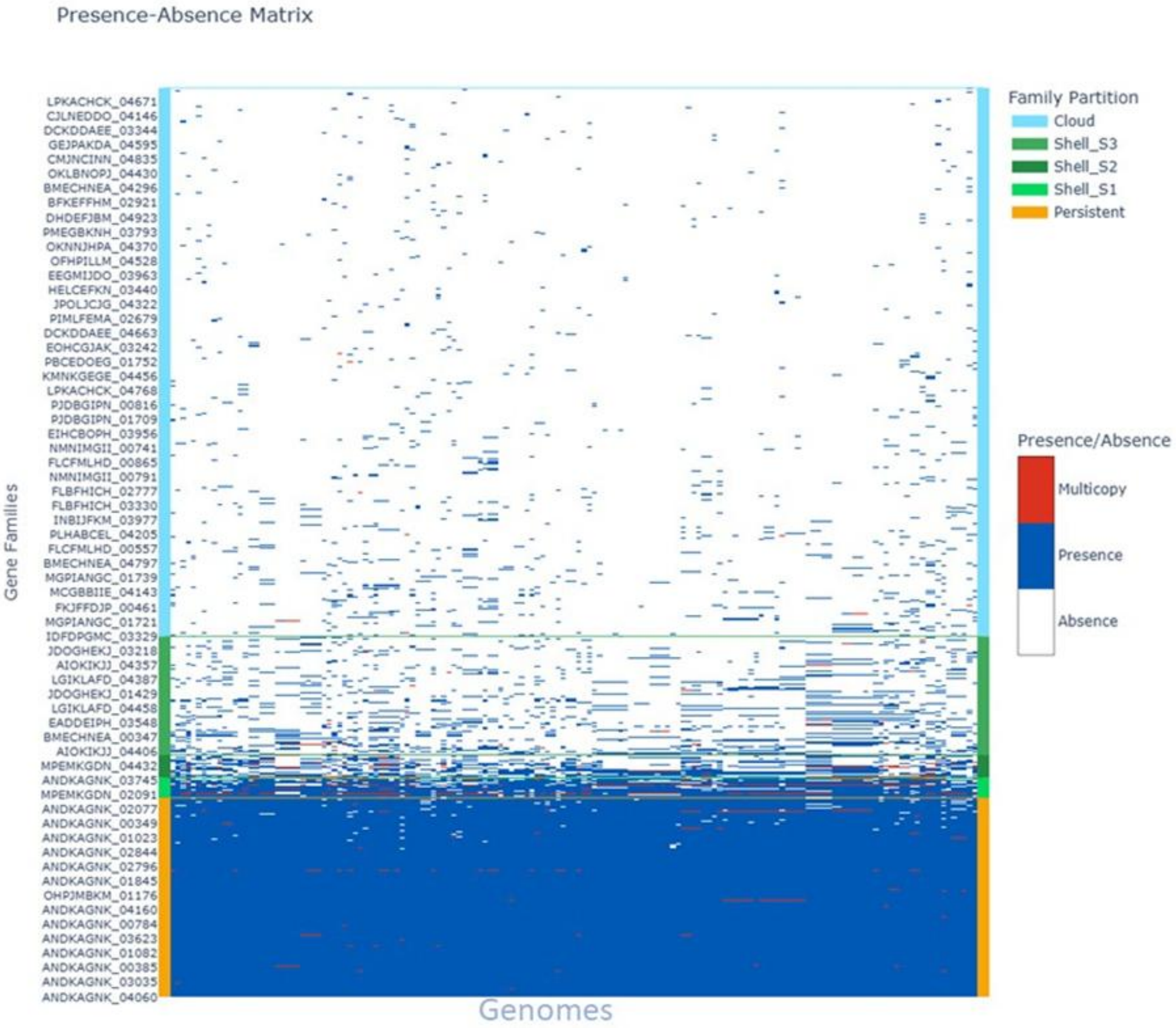


Figure 16: Graphical representation of the pangenome analysis of 159 *E. coli* ST10 isolates. The U-shaped absence/presence heatmap illustrates the presence of genes by a shade of blue, while absent genes are shaded white; the red shade in between the blue/white shades indicates multicopy genes.

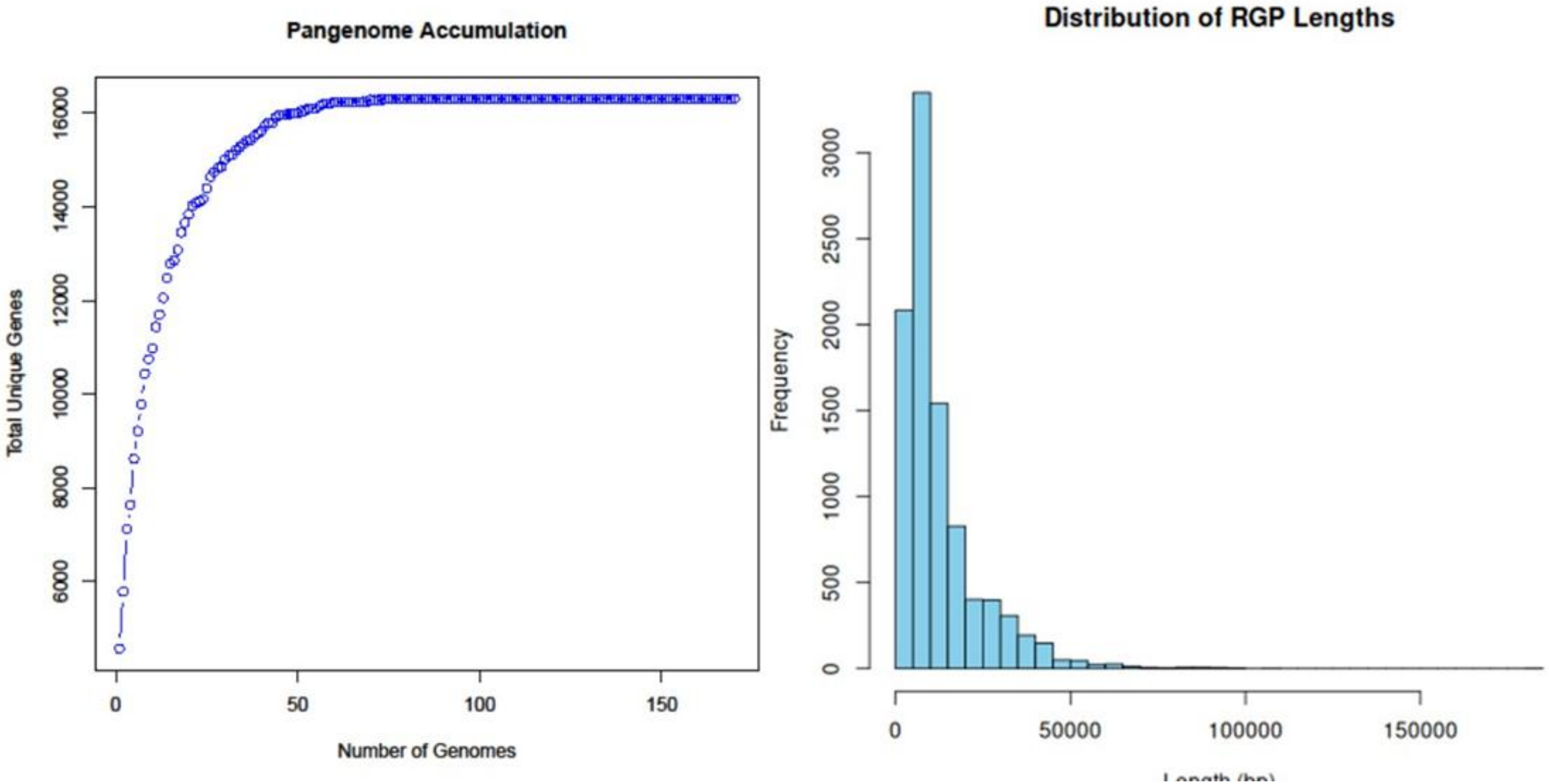


Figure 17: Regions of genomic plasticity figure, on the left side, the pangenome accumulation curve indicating how the number of genes are increasing within the 159 *E. coli* ST10 genomes, on the right-side, a histogram showing the distribution of the Regions of genomic plasticity lengths indicated by their frequency.

5.4.6 Prevalence and distribution of antimicrobial resistance genes in *E. coli* ST10

The occurrence and distribution of AMR genes among *E. coli* ST10 isolates were analysed across all 159 genomes, with comparisons made by country of origin and isolate source. A total of 1,024 AMR genes were identified across the dataset. Nigeria exhibited the highest burden of AMR genes (20%), followed by Tanzania (17%), South Africa (16%), Kenya (9%), Ghana (6%), Niger (6%), Madagascar (5%), Djibouti (4%), Uganda (4%), and Ethiopia (3%. Countries with the lowest observed AMR gene counts ($n < 10$) included Egypt, Gambia, and Tunisia. Notably, the four isolates generated in this study each carried only a single resistance gene, *mdf(A)_1* indicating low resistance.

An analysis of AMR gene classes by country revealed notable geographic patterns in the distribution of resistance determinants among *E. coli* ST10 isolates. In Nigeria, the most frequently detected resistance genes conferred resistance to aminoglycosides, β-lactams, extended-spectrum β-lactamases (ESBLs), sulphonamides, and tetracyclines. The dominant genes included *aph(6)-Id_1*, *aph(3'')-Ib_5*, *blaTEM-1B_1*, *blaCTX-M-15_1*, *sul2_2*, *sul1_5*, and *tet(A)_6*. Ethiopian isolates primarily harboured *blaTEM-1B_1*, indicating resistance predominantly to β-lactams.

South African isolates exhibited a broader resistance profile, encompassing β-lactams, aminoglycosides, tetracyclines, sulphonamides, trimethoprim, and efflux pump-associated genes, including *aadA2_1*, *aph(3'')-Ib_5*, *aph(6)-Id_1*, *blaTEM-1B_1*, *dfrA12_8*, *oqxA_1*, *oqxB_1*, *tet(A)_6*, *tet(B)_2*, and *sul2_3*. In Tanzania, resistance was most prominent against aminoglycosides, followed by β-lactams, sulphonamides, trimethoprim, and tetracyclines. The key genes included *aph(3'')-Ib_5*, *aph(6)-Id_1*, *blaTEM-1B_1*, *dfrA8_1*, *sul2_2*, and *tet(A)_6*. Kenyan isolates showed similar resistance patterns to aminoglycosides, β-lactams, sulphonamides, and tetracyclines, marked by the presence of *aph(3'')-Ib_5*, *aph(6)-Id_1*, *blaTEM-1B_1*, *sul1_5*, *sul2_2*, and *tet(A)_6*. In Ghana, resistance genes associated with aminoglycosides, fluoroquinolones, β-lactams (including ESBLs), tetracyclines, sulphonamides, and trimethoprim were frequently identified. Notable genes included *aac(6')-Ib-cr_1*, *aadA5_1*, *blaCTX-M-15_1*, *blaOXA-1_1*, *dfrA17_1*, *mph(A)_2*, *mdf(A)_1*, *sul1_5*, and *tet(B)_2*.

In Niger, isolates harboured genes conferring resistance to aminoglycosides, ESBLs, β-lactams, and tetracyclines, notably *aac(6')-Ib-cr_1*, *blaCTX-M-15_1*, *blaOXA-1_1*, and *tet(B)_2*. Isolates from Madagascar predicted resistanceto a broad range of antimicrobial classes, including aminoglycosides, ESBLs, β-lactams, sulphonamides, trimethoprim, fluoroquinolones, and tetracyclines. Frequently detected genes were *aph(3'')-Ib_5*, *aph(6)-Id_1*, *blaCTX-M-15_1*, *blaTEM-1B_1*, *dfrA14_5*, *qnrS1_1*, and *tet(A)_6*. All isolates from Djibouti predicted resistance to aminoglycosides, ESBLs, β-lactams, sulphonamides, trimethoprim, fluoroquinolones, and tetracyclines, with the following genes commonly present: *aph(3'')-Ib_5*, *aph(6)-Id_1*, *blaCTX-M-15_1*, *blaOXA-181_1*, *blaTEM-1B_1*, *dfrA14_5*, *mdf(A)_1*, *qnrS1_1*, *sul2_2*, and *tet(A)_6*. Ugandan isolates were also multidrug-resistant, carrying genes conferring resistance to aminoglycosides, β-lactams, macrolides, fluoroquinolones, sulphonamides, and tetracyclines, including *aph(3'')-Ib_5*, *aph(6)-Id_1*, *blaTEM-1B_1*, *mph(A)_2*, *sul2_2*, and *tet(A)_6*.

Importantly, the extended-spectrum β-lactamase (ESBL) gene *blaCTX-M-15_1*—a critical marker of resistance to third-generation cephalosporins—was detected in 28% of isolates (Figure 18). Several other countries, including Benin, the Democratic Republic of Congo (DRC), Egypt, Burkina Faso, Gambia, Rwanda, Zimbabwe, Tunisia, and Zambia, were represented by one or two isolates each. Despite limited sample sizes, these isolates commonly harboured resistance genes against β-lactams, sulphonamides, aminoglycosides, and tetracyclines, indicating widespread distribution of key AMR determinants across the continent.

Stratification by host/source type revealed that human-derived isolates accounted for the majority of resistance genes (51%), followed by those from poultry (19%), swine (14%), environmental sources (7%), and food products (<2%). Multidrug resistance (MDR)—defined as resistance to three or more antibiotic classes was observed in 71% of isolates, indicating widespread co-resistance and highlighting the zoonotic and public health risks associated with the ST10 lineage in Africa. The remaining 29% of isolates did not meet the MDR threshold.

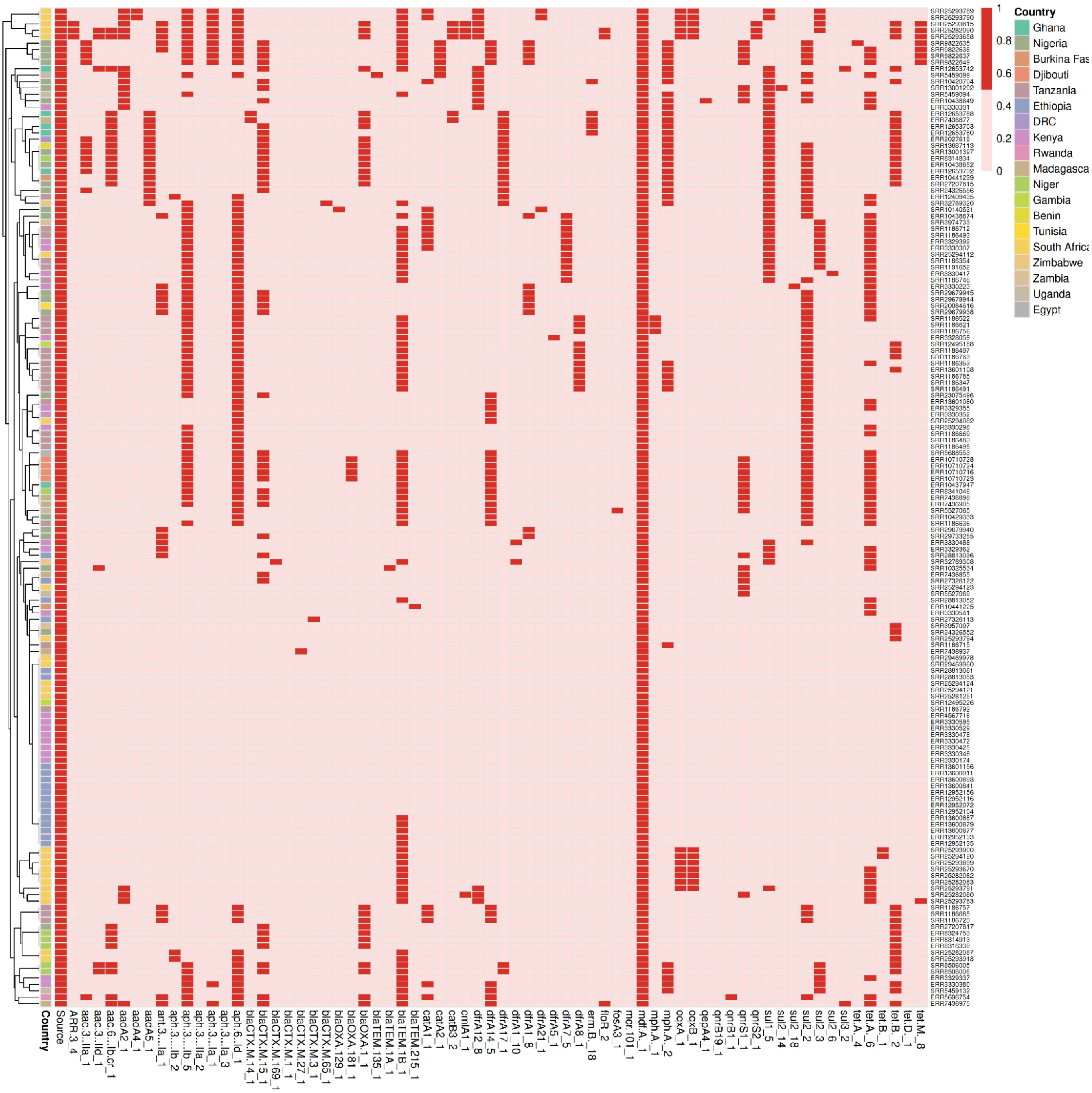


Figure 18: Phylogeny-Linked Heatmap of Antimicrobial Resistance Gene Profiles in African *E. coli* ST10 Isolates. Heatmap depicting the distribution of AMR gene profiles among 159 *E. coli* ST10 isolates from diverse African sources. The left panel displays a SNP-based maximum likelihood phylogenetic tree, constructed from core genome SNP alignments, highlighting the evolutionary relationships between isolates. Isolates are annotated and colour-coded by country of origin, as indicated in the legend on the right. The heatmap matrix shows the presence (red shading) or absence (white) of AMR genes across isolates, with darker red tones denoting gene clustering or frequent occurrence within phylogenetic subgroups. This integrated view illustrates the geographic and evolutionary distribution of resistance determinants across the ST10 lineage in Africa.

5.4.7 Distribution of plasmid replicons in *E. coli* ST10 isolates

Analysis of plasmid replicon profiles within the *E. coli* ST10 dataset revealed the widespread presence of plasmids associated with horizontal gene transfer of both AMR and virulence determinants (Figure 19). The most frequently identified plasmid was the colicin-associated replicon ColRNAI_1, present in 38% of isolates, followed by IncFIB(AP001918)_1 (32%), and IncFIC(FII)_1 (18%). Other notable replicons included Col(MG828)_1, Col156_1, and Col440I_1, each detected in 17% of isolates. Additional plasmid types observed at lower frequencies included Col(BS512)_1 and IncB/O/K/Z_1 (12% each), as well as Col(MP18)_1 (11%). Other less prevalent but relevant replicons detected included IncFIA_1, IncFIB(K)_1_Kpn3, IncB/O/K/Z_2, IncX1_1, and p0111_1. Among the four newly sequenced South African isolates, only a limited set of plasmids was observed, including Col(MG828)_1, Col156_1, ColRNAI_1, and IncI1_1_Alpha.

Geographically, the highest plasmid diversity and abundance were observed in isolates from Tanzania, South Africa, Nigeria, Kenya, and Ethiopia. Source-specific analysis showed that human-derived isolates harboured the greatest number of plasmid types, followed by poultry, porcine, and environmental samples. In Nigeria, South Africa, Kenya, Ethiopia and Tanzania the most predominant plasmid replicon were the Col-and F-plasmids (ColRNAI_1, Col440I_1, Col(MG828)_1, IncFIA_1, IncFIB(AP001918)_1, IncFIB(K)_1_Kpn3, IncFII(29)_1_pUTI89, IncB/O/K/Z_1).

In Ghana and Niger, the Col-and F-plasmids were IncFIA_1, IncFIB(AP001918)_1 and Col(BS512)_1 were predominant. In Djibouti, IncFIA_1, IncFIB(AP001918)_1 and ColKP3_1 were predominant. In Madagascar, only the IncFIB(AP001918)_1 plasmid was predominat. The ColRNAI_1 and IncFIB(AP001918)_1 plasmid were found in Uganda. Countries such as Benin, DRC, Egypt, Burkina Faso, Gambia, Rwanda, Zimbabwe, Tunisia and Zambia each had one or two isolates in this study, majority of these isolates had the IncFIB(AP001918)_1, few Col-plasmids were also observed.

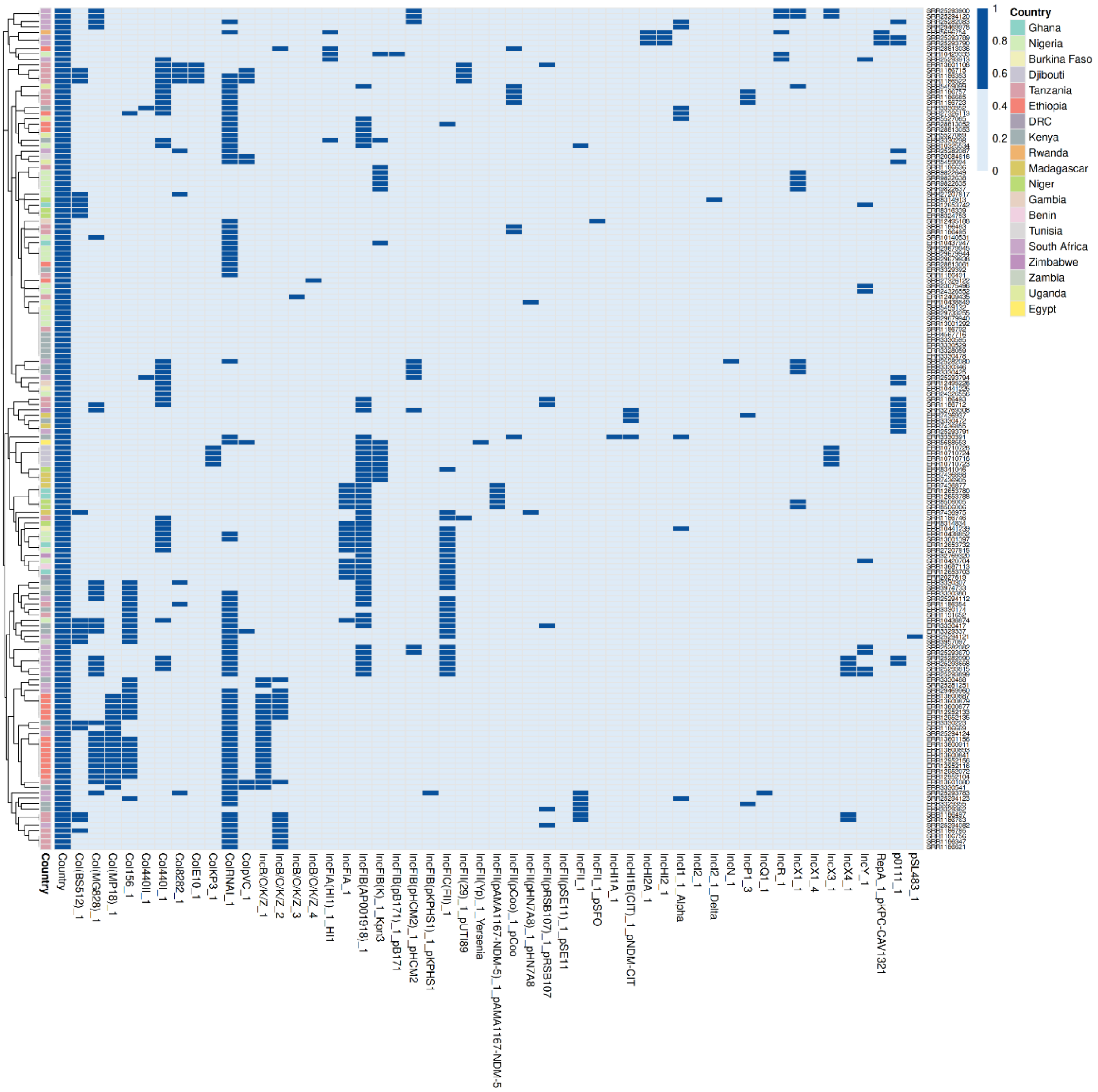


Figure 19: Phylogeny-Linked Heatmap of Plasmid Replicon Profiles in African *E. coli* ST10 Isolates. Heatmap showing the distribution of plasmid replicon profiles among 159 *Escherichia coli* ST10 isolates from various African sources. The left panel presents a core genome SNP-based maximum likelihood phylogenetic tree, depicting the evolutionary relationships between isolates. Isolates are annotated by country of origin, with colour codes indicated in the legend on the right. The heatmap matrix illustrates the presence (dark blue shading) or absence (white) of specific plasmid replicons, clustered in alignment with the phylogeny. This figure highlights the diversity, distribution, and phylogenetic clustering of plasmid types, which are often associated with antimicrobial resistance and horizontal gene transfer across diverse ecological niches.

Functionally, strong associations were observed between specific plasmid replicons and resistance phenotypes. Isolates carrying ColRNAI_1 frequently predicted resistance to tetracyclines, β-lactams, and aminoglycosides. More strikingly, isolates harbouring IncFIB(AP001918)_1 were commonly multidrug-resistant, with resistance profiles spanning most major antimicrobial classes, including β-lactams, tetracyclines, sulphonamides, macrolides, quinolones, trimethoprim, aminoglycosides, and fluoroquinolones.

5.4.8 Time scale phylogeny

A timescale phylogenetic analysis was performed to assess the temporal and geographic evolution of the *E. coli* ST10 lineage among African isolates. The analysis revealed that isolates originating from the same country and collected within the same year tended to cluster together phylogenetically. This pattern was particularly evident in isolates from Nigeria (2019), South Africa (2019), and Ethiopia (2022), suggesting localized transmission events and potential outbreak scenarios within those countries during the respective years. These temporal clusters indicate sustained intra-country transmission and possible clonal expansion of specific ST10 lineages. Interestingly, instances of inter-regional clustering were also observed. For example, a Tanzanian isolate from 2009 was found to cluster closely with Ethiopian isolates from 2022. Such patterns may reflect historical dissemination events or the long-range movement of strains across different regions, potentially facilitated by trade, travel, or other cross-border activities (Figure 20).

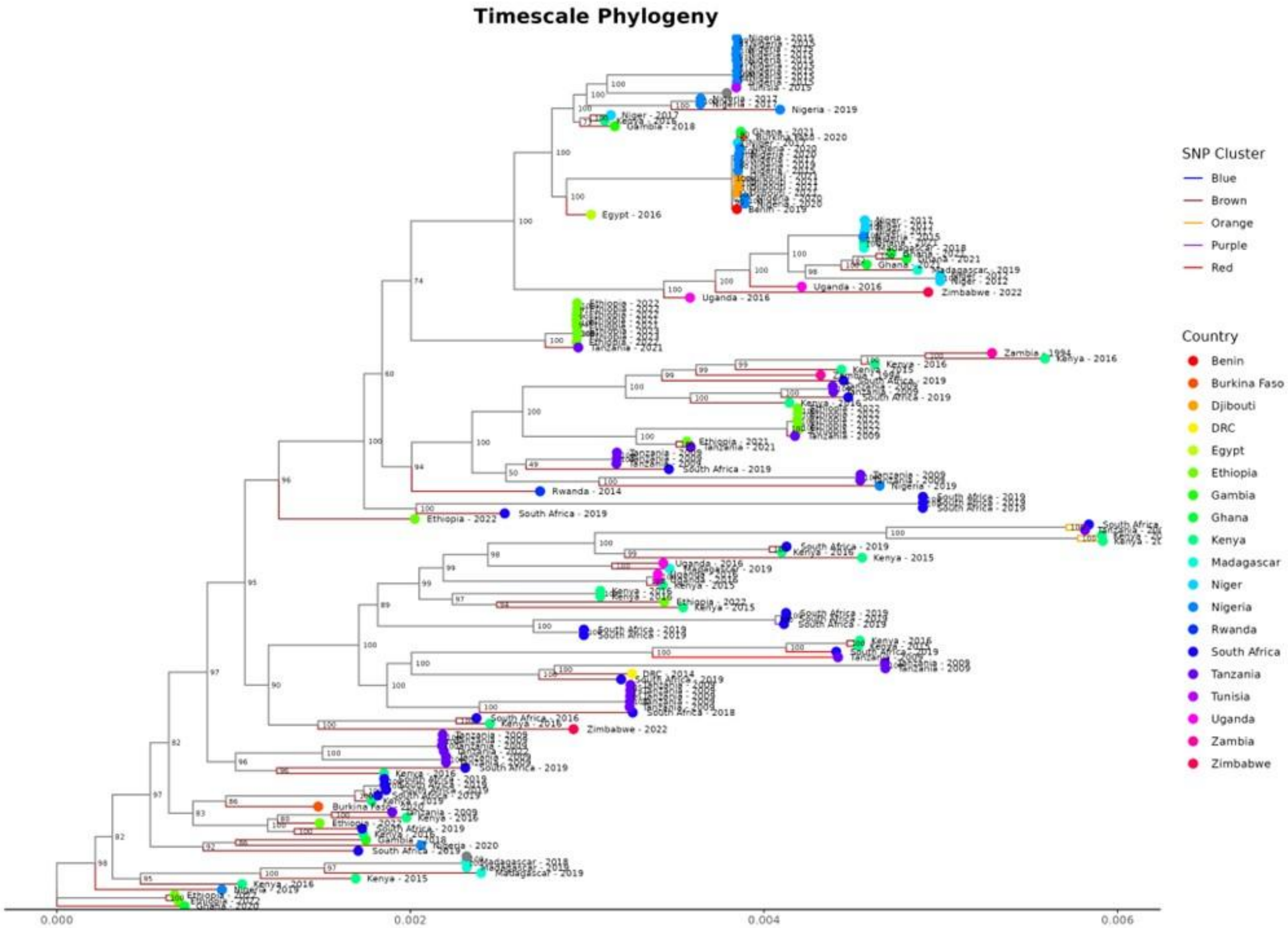


Figure 20: Timescale phylogenetic tree of *E. coli* ST10 isolates from various African countries. Various colours on the nodes indicates the country that is currently occurring there. On the nodes, countries are labelled with the year of sampling. Bootstrap values are indicated on the tree branches to show the maximum likelihood analysis.

5.5 Discussion

This study provides the first comprehensive continent-wide genomic investigation of *E. coli* ST10, incorporating 159 isolates collected from 19 African countries across human, animal, environmental, and food sources. By employing WGS, SNP-based phylogeny, pangenome and AMR profiling, and mobile genetic element analysis, this presents an integrated view of the genetic diversity, evolutionary plasticity, and public health implications of this emerging lineage. Despite high species-level nucleotide conservation (mean ANI: 99.44%, SD = 0.17), the lineage displayed remarkable intra-lineage diversity, reflecting extensive microevolutionary divergence and adaptability to diverse ecological niches (Touchon et al., 2020; Machado et al., 2024).

The high ANI values confirm that the isolates belong to the same lineage, however, their genetic divergence evident in the 122 distinct SNP clusters indicates that *E. coli* ST10 is a highly adaptable clonal complex capable of host-specific adaptation and geographic segregation (Kocsis et al., 2022; Santos-Neto et al., 2022). FastANI comparisons across 10,948 genome pairs confirmed strong within-lineage relatedness, but the right-skewed ANI distribution suggests recent evolutionary changes among some subpopulations. These findings align with previous studies that describe *E. coli* ST10 as a generalist lineage, commonly found in asymptomatic human carriers, livestock, poultry, food products, and aquatic environments (Schaufler et al., 2019; Ogundare et al., 2024).

Notably, localized SNP divergence in isolates from countries such as South Africa, Nigeria, and Ethiopia likely reflects endemic transmission and evolutionary drift within national borders. However, inter-country clustering, for example, between Tanzanian and Ethiopian isolates suggests historical dissemination or long-distance translocation of lineages via trade, travel, or environmental spread (Decano et al., 2021; Takawira et al., 2022).

The predominance of *E. coli* ST10 in several African countries corroborates previous findings. In South Africa, *E. coli* ST10 was first identified in 2019 in a meat sample from Gauteng Province (Bolukaoto et al., 2019) and was later linked to both clinical and environmental isolates (Mbelle et al., 2019; Richter et al., 2021). The lineage is frequently associated with

EHEC and ExPEC pathotypes, often found in wastewater, livestock, and food chains (Ogundare et al., 2024). Similar trends have been reported in Nigeria, where *E. coli* ST10 is a leading cause of childhood diarrhea linked to the EAEC pathotype and is commonly detected across multiple reservoirs, including humans, poultry, and water sources (Okeke et al., 2010; Afolayan et al., 2022; Rabiu et al., 2022;). In Ethiopia, *E. coli* ST10 has emerged as an epidemiologically important lineage in both humans and animals, often carrying the ESBL gene *blaCTX-M-15* and displaying high multidrug resistance (Wolde et al., 2024; Chekole et al., 2025). It's repeated detection in healthy individuals further highlights it's zoonotic and public health significance (Aworh et al., 2021; Sewunet et al., 2022). In East African countries like Kenya, Tanzania, and Zimbabwe, *E. coli* ST10 is a known cause of urinary tract infections and frequently harbours ESBLs, particularly in human isolates (Decano et al., 2021; Takawira et al., 2021).

The extensive AMR gene diversity observed in isolates from Nigeria, Tanzania, and South Africa raises serious public health concerns. Studies show that these countries report the highest prevalence of DEC pathotypes in Africa, often reflecting disparities in antibiotic usage, diagnostics, and surveillance (Ng'etichi et al., 2021; Kariuki et al., 2022; Kalule et al., 2024). Importantly, all isolates harbored the efflux pump gene *mdf(A)_1*, which confers resistance to a broad range of antibiotics, including quinolones and macrolides (Li & Ge, 2023). This supports evidence that intrinsic resistance mechanisms are deeply embedded in the *E. coli* ST10 genome (Nasrollahian et al., 2024). The presence of *blaCTX-M-15* in 28% of isolates, particularly from humans, is a critical finding, as this gene is a global marker of third-generation cephalosporin resistance and frequently occurs in clinical ExPEC strains (Müller et al., 2016; Wang et al., 2021). Notably, ST10, as a member of clonal complex 10 (CC10), has been implicated in AMR gene dissemination in both healthy and diseased individuals and across animal products (Kocsis et al., 2022).

High multidrug resistance detected in our isolates highlights the widespread dissemination of resistance genes within the ST10 lineage across African countries. Of importance, is that majority of isolates presenting with MDR are human-derived. This pattern suggests that there are human-associated reservoirs that play a significant role in dissemination and persistence of AMR within the ST10 lineage. Country level comparisons indicated that there is significant variation in terms of occurrence of MDR in African countries which reflects that there are differences in terms of antibiotic usage.

The high abundance of resistance genes across different antibiotic classes including β-lactams, sulphonamides, tetracyclines, and aminoglycosides mirrors findings in South African veterinary and human sectors, where the overuse of antibiotics, particularly tetracyclines, is poorly regulated (Ramatla et al., 2021; Mokgophi et al., 2021; Tucker et al., 2022; Smith et al., 2024). Furthermore, AMR prevalence is often more strongly correlated with geographic patterns than lineage, as treatment practices vary by country. For example, in Africa, trimethoprim-sulfamethoxazole remains a commonly used antibiotic, while in Asia, ciprofloxacin or azithromycin are more frequently used for treating *E. coli* infections (Ingle et al., 2018).

The detection of IncF and Col-type plasmids in nearly half of all isolates emphasizes their role in the acquisition and dissemination of resistance genes (Mone et al., 2023). The dominance of IncFIB(AP001918)_1 and ColRNAI_1 replicons particularly in MDR isolates mirrors patterns seen in clinical *E. coli* ST10 strains across Europe and Asia (Wu et al., 2023; Fang et al., 2024). These plasmids often co-harbour multiple AMR determinants and virulence genes, enhancing survival under antibiotic pressure and enabling host adaptability (Tello et al., 2022; Shaib et al., 2023).

The presence of multiple replicon types within individual isolates suggests a high level of recombination and potential for cross-species gene flow. The detection of other replicons like Col(MG828)_1, Col156_1, and Col440I_1 previously associated with extensively drug-resistant (XDR) strains further strengthens the hypothesis that *E. coli* ST10 serves as a key reservoir of mobile genetic elements (Ramsamy et al., 2020; Wareth et al., 2022).

Our pangenome analysis, comprising 16,311 genes, revealed a highly plastic genome with 21% persistent (core), 17% shell, and 60% cloud genes. The high proportion of accessory genome content is indicative of frequent HGT, which is typically mediated by plasmids, genomic islands, and integrons (Cortimiglia et al., 2021; Kushwaha et al., 2024). This plasticity may explain *E. coli* ST10's ability to transition from commensal to pathogenic lifestyles depending on host and environmental contexts (Dobrindt et al., 2010).

The time-scaled phylogenetic reconstruction suggests localized clonal expansion of specific lineages within Nigeria, South Africa, and Ethiopia. Remarkably, the close clustering of a Tanzanian isolate from 2009 with Ethiopian isolates from 2022 points to either prolonged environmental persistence or historical introduction of strains that have since evolved regionally (Aibinu et al., 2012; Figueras et al., 2022; Malesa et al., 2024). These patterns highlight the urgent need for harmonized One Health surveillance systems to track the spatiotemporal dynamics of high-risk lineages like *E. coli* ST10.

5.6 Conclusion

Overall, this study provides compelling evidence that the *E. coli* ST10 lineage in Africa maintains a high level of genomic similarity at the species level, while exhibiting significant intra-lineage diversity, as revealed by SNP-based phylogenetic analysis.This apparent paradox underscores the lineage's remarkable evolutionary plasticity and ecological versatility. The findings of our study affirm that humans remain the principal reservoirs for *E. coli* ST10 on the African continent, a pattern probably driven by high AMR selection pressure and unregulated antibiotic use. The widespread of co-occurrence of AMR genes and plasmid replicons especially those linked with HGT futher highlights the capacity of this lineage to persist and disseminate within and between ecological niche including livestock, food products and the environment sources.

The high prevalence of MDR, coupled with the detection of ESBL genes such as *blaCTX-M-15*, highlights the potential for therapeutic failure and underscores the urgent need for coordinated AMR mitigation strategies. These findings emphasize the critical role of genomic surveillance in understanding the epidemiology of high-risk lineages and call for the implementation of robust, integrated One Health surveillance programmes across Africa. Such efforts are especially imperative in low- and middle-income countries, where limited diagnostic infrastructure and weak antimicrobial stewardship exacerbate the risk of widespread resistance gene dissemination.

CHAPTER 6

GENERAL DISCUSSION, RECOMMENDATIONS, FUTURE STUDIES AND CONCLUSION

6.1 General discussion

Chapter 3 focused on the characterization of typical *E. coli* isolates from food-producing animals and animal-derived products in South Africa with the aim of assessing their genetic diversity, pathogenic potential, and AMR profiles. Employing WGS, the study provided important insights into the genomic epidemiological and public health risk posed by generic *E. coli* strains from common protein, sources including poultry, porcine, and bovine products.

The findings of Chapter 3 revealed substantial genetic diversity among isolates with the identification of five different phylogroups (A, B1, B2, C, and E). Phylogroup A is the most common, with the predominance of ExPEC, particularly avian and uropathogenic pathotypes. The high occurrence of ExPEC, primarily in poultry, underscores the zoonotic potential of *E. coli* in the food chain. Intestinal pathotypes, such as ETEC, EPEC, and STEC, were also detected, indicating their relevance to gastrointestinal illnesses.

The chapter reported 23 different serotypes and 19 sequence types, with O101:H37, O2:H4, O6:H5 and O64:H19 being the most predominant serotypes and ST1858, ST975, and ST10 being the most common sequence types. The majority of these sequence types have global distribution and are associated with human and animal infections, while some may represent novel, locally adapted strains. Virulence-associated genes further underscored the pathogenic potential of these isolates. Genes encoding toxins, adhesins, and secretion systems were identified, indicating their ability to cause both intestinal and extraintestinal infections. Notably, the isolates had over a 90% probability of being human pathogens, suggesting a significant risk to public health. The detection of the type I CRISPR system, predominantly subtype I-E, highlights the bacterium's adaptive immune mechanisms, providing insights into the evolutionary dynamics of *E. coli* pathogenicity.

The AMR analysis identified 12 resistance genes, including those conferring resistance to tetracyclines, sulphonamides, β-lactams, and aminoglycosides. Resistance to disinfectants and the detection of mobile genetic elements (e.g., IS26, Tn6196, and plasmid replicons) further suggest a capacity for horizontal gene transfer and persistence in the food production environment. Overall, the chapter highlights the significant zoonotic and AMR risks posed by *E. coli* in South African food systems. Despite a limited sample size, the study contributes critical baseline data for national surveillance and calls for expanded molecular monitoring. Integrating genomic tools into public health strategies will be essential for mitigating foodborne risks and curbing the spread of antimicrobial resistance in both local and global contexts.

Chapter 4 investigated the structural and functional characteristics of the Type III Secretion System (T3SS) in EPEC strains isolated from food-producing animals and animal-derived products in South Africa. Using WGS, the chapter focused on the locus of enterocyte effacement (LEE) a pathogenicity island that encodes the T3SS. This secretion system is a major virulence mechanism in EPEC, enabling the bacterium to inject effector proteins into host cells, thereby subverting cellular processes, evading immune responses, and establishing infection.

Key LEE-encoded proteins essential for T3SS function were identified, including EscV, EscN, EscD, EscU, and EspA. These proteins are central to the assembly and operation of the injectisome a syringe-like structure that mediates effector translocation. EscV and EscD regulate the export gate, while EspA forms the translocation filament. Computational modeling and validation, including Ramachandran plot analyses, confirmed the structural stability of these proteins under physiological conditions. The detection of transmembrane domains in EscV, EscD, EscU, and EspA further affirmed their integration into bacterial membranes, which is critical for the injectisome's structural integrity and function.

The chapter also highlighted the role of both LEE-encoded and non-LEE-encoded effectors, such as NleE and NleB, which contribute to immune modulation and inflammation in the host. These effector proteins act synergistically with structural components to enhance EPEC's infectivity, underscoring the complexity and coordinated functionality of the T3SS machinery.

Physicochemical analyses revealed variations in protein properties consistent with their roles in secretion and interaction with host cells, with EscV and EscD emerging as central to effector export.

This chapter presents the first comprehensive structural characterization of the T3SS in EPEC strains from South African animal-derived food sources. The presence of functional T3SS components in isolates from poultry, pork, and beef staple protein sources for the population raises important food safety and public health concerns. These findings highlight the need for active surveillance programs to monitor the spread of virulent *E. coli* strains in the food supply chain.

Globally, EPEC remains a leading cause of diarrheal disease, particularly in low-resource settings with limited access to sanitation. The T3SS characterization in this study aligns with international research aimed at elucidating the molecular basis of EPEC pathogenicity. By contributing to the global virulence factor database, this work enhances understanding of T3SS biology and its epidemiological implications.

The application of WGS in identifying and characterizing T3SS components offers a powerful tool for outbreak tracking, virulence monitoring, and targeted intervention development. These findings also support the WHO’s goal of reducing foodborne diarrheal diseases by 40% by 2030. The study underscores the urgent need for expanded research to explore the genetic diversity of EPEC strains harboring T3SS and reinforces the importance of integrating genomic technologies into public health surveillance to mitigate the risks associated with foodborne EPEC infections.

Chapter 5 explored the genomic characteristics, evolutionary plasticity, and AMR profiles of the *E. coli* ST10 lineage circulating across multiple African countries. The study aimed to uncover the public health implications posed by this emerging zoonotic lineage, recognized globally for its broad host range and notable AMR patterns. A total of 159 ST10 isolates were analyzed, including both newly sequenced South African strains and publicly available genomes representing 19 African countries: South Africa, Tanzania, Ethiopia, Nigeria, Kenya, Ghana, Uganda, Zimbabwe, Niger, the Democratic Republic of Congo, Burkina Faso, Madagascar, Djibouti, Gambia, Benin, Rwanda, Zambia, Tunisia, and Egypt.

Using WGS, the study revealed that *E. coli* ST10 isolates shared high species-level similarity, with ANI values ranging from 98% to 100%. However, core genome SNP-based analysis revealed significant genetic diversity, with more than 100 distinct SNP clusters many of which consisted of singletons highlighting high intra-lineage variation. A small number of closely related clusters (comprising two to five isolates) were observed, primarily from human-derived isolates within the same geographic region, suggesting localized clonal expansion.

Notably, while literature suggests that *E. coli* ST10 was first identified in South African human isolates in 2019, two isolates from Chapter 3 collected from porcine sources in 2016 and 2018 demonstrated that *E. coli* ST10 has been circulating undetected in the country for several years. Interestingly, despite their similar origins, these South African isolates displayed considerable genomic divergence, underscoring the lineage's evolutionary flexibility.

Region-specific clustering revealed that the majority of SNP diversity was observed in isolates from Tanzania, South Africa, Nigeria, Kenya, and Ethiopia. The elevated number of SNP clusters in these countries is likely linked to higher isolate availability in public databases. Human-derived isolates showed the greatest genetic diversity and accounted for most resistance gene carriage, suggesting that humans are primary reservoirs of *E. coli* ST10, followed by livestock and environmental sources.

Pangenome analysis demonstrated a large number of gene families and extensive genomic plasticity across isolates, with numerous regions of genomic variability linked to the acquisition of mobile genetic elements. This supports the role of *E. coli* ST10 as a hub for HGT. The AMR profiling revealed a high prevalence of resistance genes, particularly in isolates from Nigeria, South Africa, and Kenya. These isolates frequently predicted resistance to multiple antibiotic classes, including β-lactams (notably extended-spectrum β-lactamases [ESBLs]), tetracyclines, sulphonamides, aminoglycosides, trimethoprim, and fluoroquinolones.

Even countries with fewer isolates displayed resistance to these major antimicrobial classes, further indicating widespread dissemination of AMR determinants across the continent. Host-specific analysis confirmed that human-derived isolates carried the highest AMR burden, and 71% of all isolates were classified as MDR, conferring resistance to three or more antibiotic classes.

Time-scaled phylogenetic analysis showed that the earliest *E. coli* ST10 isolates from Africa date back to 1994, with South African isolates traceable to at least 2016. Isolates from the same country and year often clustered together, suggesting local transmission events. However, the detection of clusters linking isolates from different countries and years points to transboundary dissemination, potentially facilitated by trade, animal movement, or environmental factors.

Among the globally significant *E. coli* sequence types, ST10 is now increasingly reported alongside ST131, ST95, and ST69. The observed 28% ESBL production rate among *E. coli* ST10 isolates in this study is particularly concerning given its rising similarity in distribution to the pandemic ST131 clone. Collectively, these findings underscore the zoonotic and public health risks posed by the widespread and adaptable *E. coli* ST10 lineage.

Chapter 5 highlights the urgent need for harmonized genomic surveillance and AMR monitoring protocols across African countries. As *E. coli* ST10 continues to evolve and expand, particularly in human reservoirs, coordinated One Health strategies are critical for mitigating its spread and public health impact.

Additionally, it is important to consider that, although the study used different methods and pipelines for analysis across chapters, these assembly strategies did not influence the overall genomic interpretations of our datasets. In Chapter 3, we employed SPAdes for de novo assembly for historic samples, whereas in chapter 5 we used Shovill for assembly, Shovill was used as it is also a SPAdes assembler, but a faster tool to use considering the number of isolates that we assembled. The differences in tools also reflects how bioinformatics tools are evolving over time. Downstream analysis such as AMR, plasmid replicons, virulence detection, we employed are standardised *in silico* tools.

Although this study provides a baseline comparative genomic analysis of *E. coli* isolates from food producing animals and animal derived food products in South Africa, a number of limitations encountered in this study must be considered. The dataset consisted of isolates from 1976 to 2018 was limited and the animal sources listed introduced bias, the publicly available genomes included in the study had an uneven distribution which introduced a bias in terms of the countries sampled. Overall, methods in this study were entirely *in silico* and were not supported by any wetlab analysis. Consequently, our findings should be interpreted based on the scope of available data rather than a full comprehensive representation globally.

6.2 General conclusion

This study provided a data showing that *E. coli* remains a serious threat to food safety and public health in South Africa. Through the utilization of WGS and *in silico* analyses, the study revealed significant genomic diversity among E. coli strains, including isolates carrying the LEE pathogenic island and a functional T3SS, hallmark features ofEPEC.The presence of these virulence markers indicates a high pathogenic potential in the strains and highlights the zoonotic threat associated with faecal contamination throughout the food production and processing continuum.

In addition to virulence profiling, the study identified *E. coli* ST10 as an important carrier of AMR genes. The detection of these resistance determinants in *E. coli* ST10 strains raises major public health concerns, particularly given the ability of such genes to disseminate across animal, food, and human reservoirs. Although the study's sample size was relatively limited, its findings expose critical gaps in national surveillance systems and highlight the lack of comprehensive genomic data, particularly for isolates originating from livestock and animal-derived food products. This knowledge gap impedes efforts to assess the full extent of AMR circulation and pathogen transmission within the country's agri-food systems.

Taken together, these findings underscore the urgent need for South Africa to strengthen its national foodborne pathogen surveillance systems by integrating genomic technologies such as WGS into routine diagnostic workflows. There is a pressing requirement for stricter food safety controls at all stages of the value chain, as well as a coordinated One Health approach to monitor and respond to the emergence and spread of virulent and antimicrobial-resistant *E. coli* strains. Addressing these challenges is essential to protecting public health, safeguarding trade, and reducing the burden of foodborne illnesses in the country.

6.3 Recommendations

Based on findings in this study, it is recommended that laboratories move towards the use of WGS for diagnosing *E. coli*. Though conversional methods of identifying this bacterium remain efficient, laboratories should align with the industry and employ methods that offer variety of analysis using sequence data to explore these organisms more thoroughly. The advances in WGS have proved to be less time consuming and inexpensive over time. It is also recommended that South Africa work towards strengthening the rules and regulations that govern food safety. Researchers in South Africa are also encouraged to deposit their data pertaining to isolations done on food and animal-based food products in the NCBI SRA to build a database of sequences that can give a good representation of *E. coli* isolates circulationg in this country.

To translate the genomic insights observed in this study requires active public health surveillance and this can only be done through the establishment of a national antimicrobial resistance repository for non-clinical isolates which will further allow a better intergration of genomic data between food, animal-sources, the environment and human-sources, this will further result in a harmonised One Health surveillance system. A repository such as this one will constantly monitor emergence of resistance determinats and also monitor high-risk lineages such as ST10 within different sources. Although our study was purely computational when predicting the T3SS in EPEC strains, future work should incorporate wetlab analysis that includes confirmation of T3SS using electron microscopy to confirm functionality. Lastly, studies that look into genotype-phenotype links should be done to properly validate the correlation of these results. These approaches, should strength our surveillance systems in South Africa and other countries.

6.4 Future perspectives

Part of continued research relating to this study includes several research activities which will be formulated based on some of the gaps that were not addressed on this thesis which includes the following.

6.4.1 Further WGS analysis

Considering the small subset of isolates which were part of this study, further analysis using WGS methods is required. To accomplish this, expansion of sampling is required which will include a wider range of animal species, food products and geographical locations which will giver a better representation of *E. coli* population and its genetic diversity.

6.4.2 Impact on public health

A risk assessment of the impact of both ExPEC pathotypes in foodborne diseases in South Africa needs to be undertaken to better understand isimpact on public *E. coli* infections has on humans and animals.

6.4.3 The role of CRISPR systems

Many *E. coli* strains are often found with the CRISPR cas systems that are sometimes inactive, therefore it is important that further studies explore the role of active CRISPR systems in both pathogenic and non-pathogenic strains and understand the role they play in HGT, AMR, and virulence.

6.4.4 In vitro studies of the T3SS

In order for us to understand how effectors invade and attack the host cells in EPEC strains, it is important that future work evaluates the T3SS system in vitro using cell culture techniques.

6.4.5 Study the unknown sequence types

Several ST’s were identified in this study, and majority of them were unknown, future analysis involves studying these ST’s genetic makeup and to publish them on the public domain.

REFERENCES

Abdalla, S. E., Abia, A. L. K., Amoako, D. G., Perrett, K., Bester, L. A., & Essack, S. Y. (2021). From farm-to-fork: *E. coli* from an intensive pig production system in South Africa shows high resistance to critically important antibiotics for human and animal use. *Antibiotics*, *10*(2). https://doi.org/10.3390/antibiotics10020178

Abdalla, S. E., Abia, A. L. K., Amoako, D. G., Perrett, K., Bester, L. A., & Essack, S. Y. (2022). Food animals as reservoirs and potential sources of multidrug-resistant diarrheagenic *E. coli* pathotypes: Focus on intensive pig farming in South Africa. *Onderstepoort Journal of Veterinary Research*, *89*(1). https://hdl.handle.net/10520/ejc-opvet_v89_i1_a1963

Abdalla, S. E., Bester, L. A., Abia, A. L. K., Allam, M., Ismail, A., Essack, S. Y., & Amoako, D. G. (2025). Genomic Insights of Antibiotic-Resistant *Escherichia coli* Isolated from Intensive Pig Farming in South Africa Using 'Farm-to-Fork' Approach. *Antibiotics*, *14*(5), 446. https://doi.org/10.3390/antibiotics14050446

Abdelrahim, S. S., Zaki, S., Abdelhamid, H., & Kotb, D. N. (2024). Virulence characterization of *Escherichia coli* sequence type 131 clone isolated from hospital-acquired and community-acquired urinary tract infections, Egypt. *Microbes and Infectious Diseases*, *5*(3), 1113–1126. https://doi.org/10.21608/mid.2024.275654.1837

Abebe, E., Gugsa, G., & Ahmed, M. (2020). Review on major food-borne zoonotic bacterial pathogens. *Journal of tropical medicine*, 2020(1), 4674235. https://doi.org/10.1155/2020/4674235

Abraham, S., Trott, D. J., Jordan, D., Gordon, D. M., Groves, M. D., Fairbrother, J. M., Smith, M. G., Zhang, R., & Chapman, T. A. (2014). Phylogenetic and molecular insights into the evolution of multidrug-resistant porcine enterotoxigenic *Escherichia coli* in Australia. *International Journal of Antimicrobial Agents*, *44*(2), 105–111. https://doi.org/10.1016/j.ijantimicag.2014.04.011

Afolayan, A. O., Aboderin, A. O., Oaikhena, A. O., Odih, E. E., Ogunleye, V. O., Adeyemo, A. T., Adeyemo, A. T., Bejide, O. S., Underwood, A., Argimón, S., Abrudan, M., Egwuenu, A., Ihekweazu, C., Aanensen, D. M., & Okeke, I. N. (2022). An ST131 clade and a phylogroup A clade bearing an O101-like O-antigen cluster predominate among bloodstream *Escherichia coli* isolates from South-West Nigeria hospitals. *Microbial Genomics*, *8*(12), 000863. https://doi.org/10.1099/mgen.0.000863

Aguirre-Sánchez, J. R., Valdez-Torres, J. B., del Campo, N. C., Martínez-Urtaza, J., del Campo, N. C., Lee, B. G., Quiñones, B., & Chaidez-Quiroz, C. (2022). Phylogenetic group and virulence profile classification in *Escherichia coli* from distinct isolation sources in Mexico. *Infection, Genetics and Evolution, 106*, 105380. https://doi.org/10.1016/j.meegid.2022.105380

Ahmad, N., Joji, R. M., & Shahid, M. (2023). Evolution and implementation of One Health to control the dissemination of antibiotic-resistant bacteria and resistance genes: A review. *Frontiers in Cellular and Infection Microbiology*, 12, 1065796. https://doi.org/10.3389/fcimb.2022.1065796

Aibinu, I., Odugbemi, T., Koenig, W., & Ghebremedhin, B. (2012). Sequence Type ST131 and ST10 Complex (ST617) predominant among CTX-M-15-producing *Escherichia coli* isolates from Nigeria. *Clinical Microbiology and Infection*, *18*(3), E49–E51. https://doi.org/10.1111/j.1469-0691.2011.03730.x

Aijuka, M., Santiago, A. E., Girón, J. A., Nataro, J. P., & Buys, E. M. (2019). *Escherichia coli* isolated from food sources and irrigation water: A potential risk for causing intestinal dysfunction?. *Food Control*, 102, 139-148. https://doi.org/10.1016/j.foodcont.2019.03.015

Akinlabi, O.C., Dada, R.A., Olayinka, A.A., Oginni-Falajiki, I.O., Bejide, O.S., Adewole, P.D., Thomson, N.R., Aboderin, A.O., Okeke, I.N (2025). Genomic epidemiology of enteropathogenic *Escherichia coli* in southwestern Nigeria. *PLoS neglected tropical diseases*. 19(8), p.e0013442. https://doi.org/10.1371/journal.pntd.0013442

Alara, J. A., & Alara, O. R. (2024). An Overview of the Global Alarming Increase of Multiple Drug Resistant: A Major Challenge in Clinical Diagnosis. *Infectious Disorders - Drug Targets*, *24*(3), 26–42. https://doi.org/10.2174/1871526523666230725103902

Aldea, I., Gibello, A., Hernández, M., Leekitcharoenphon, P., Bortolaia, V., & Moreno, M. A. (2022). Clonal and plasmid-mediated flow of ESBL/AmpC genes in *Escherichia coli* in a commercial laying hen farm. *Veterinary Microbiology*, 270, 109453. https://doi.org/10.1016/j.vetmic.2022.109453

Alfinete, N. W., Bolukaoto, J. Y., Heine, L., Potgieter, N., & Barnard, T. G. (2022). Virulence and phylogenetic analysis of enteric pathogenic *Escherichia coli* isolated from children with diarrhoea in South Africa. *International Journal of Infectious Diseases*, *114*, 226–232. https://doi.org/10.1016/j.ijid.2021.11.017

Alhadlaq, M. A., Aljurayyad, O. I., Almansour, A., Al-Akeel, S. I., Alzahrani, K. O., Alsalman, S. A., Yahya, R., Al-Hindi, R. R., Hakami, M. A., Alshahrani, S. D., Alhumeed, N. A., Al Moneea, A. M., Al-Seghayer, M. S., AlHarbi, A. L., AL-Reshoodi, F. M., & Alajel, S. (2024). Overview of pathogenic *Escherichia coli*, with a focus on Shiga toxin-producing serotypes, global outbreaks (1982–2024) and food safety criteria. *Gut Pathogens*, *16*(1), 1–18. https://doi.org/10.1186/s13099-024-00641-9

Alhassan, M. Y., Kabara, M. K., Ahmad, A. A., Abdulsalam, J., & Habib, H. I. (2025). Revisiting antibiotic stewardship: veterinary contributions to combating antimicrobial resistance globally. *Bulletin of the National Research Centre 2025 49:1*, *49*(1), 1–9. https://doi.org/10.1186/s42269-025-01317-3

Ali, S., & Alsayeqh, A. F. (2022). Review of major meat-borne zoonotic bacterial pathogens. *Frontiers in Public Health*, 10, 1045599. https://doi.org/10.3389/fpubh.2022.1045599

Allard, M. W., Bell, R., Ferreira, C. M., Gonzalez-Escalona, N., Hoffmann, M., Muruvanda, T., Ottesen, A., Ramachandran, P., Reed, E., Sharma, S., Stevens, E., Timme, R., Zheng, J., & Brown, E. W. (2018). Genomics of foodborne pathogens for microbial food safety. *Current opinion in biotechnology*, 49, 224-229. https://doi.org/10.1016/j.copbio.2017.11.002

Almaary, K. S. (2023). Food-Borne Diseases and their Impact on Health. *Biosciences Biotechnology Research Asia*, *20*(3), 745–755. http://dx.doi.org/10.13005/bbra/3129

Almendros, C., Guzmán, N. M., Díez-Villaseñor, C., García-Martínez, J., & Mojica, F. J. M. (2012). Target Motifs Affecting Natural Immunity by a Constitutive CRISPR-Cas System in *Escherichia coli*. *PLOS ONE*, *7*(11), e50797. https://doi.org/10.1371/journal.pone.0050797

Al-Mustapha, A. I., Alada, S. A., Raufu, I. A., Lawal, A. N., Eskola, K., Brouwer, M. S., Adetunji, V., & Heikinheimo, A. (2022). Co-occurrence of antibiotic and disinfectant resistance genes in extensively drug-resistant *Escherichia coli* isolated from broilers in Ilorin, North Central Nigeria. *Journal of Global Antimicrobial Resistance*, 31, 337–344. https://doi.org/10.1016/j.jgar.2022.11.002

Alonso, C. A., Zarazaga, M., Ben Sallem, R., Jouini, A., Ben Slama, K., & Torres, C. (2017). Antibiotic resistance in *Escherichia coli* in husbandry animals: the African perspective. In *Letters in Applied Microbiology* (Vol. 64, Issue 5). https://doi.org/10.1111/lam.12724

Alqasim, A. (2021). Colistin-resistant Gram-negative bacteria in Saudi Arabia: A literature review. In *Journal of King Saud University – Science*, 33(8), 101610. https://doi.org/10.1016/j.jksus.2021.101610

Angamuthu, K., & Piramanayagam, S. (2017). Evaluation of in silico protein secondary structure prediction methods by employing statistical techniques. *Biomedical and Biotechnology Research Journal*, *1*(1), 29–36. https://doi.org/10.4103/BBRJ.BBRJ_28_17

Anueyiagu, K. N., Agu, C. G., Umar, U., & Lopes, B. S. (2024). Antimicrobial Resistance in Diverse *Escherichia coli* Pathotypes from Nigeria. *Antibiotics*, *13*(10), 922. https://doi.org/10.3390/antibiotics13100922

Anyanwu, M. U., Jaja, I. F., Okpala, C. O. R., Njoga, E. O., Okafor, N. A., & Oguttu, J. W. (2023). Mobile Colistin Resistance (mcr) Gene-Containing Organisms in Poultry Sector in Low- and Middle-Income Countries: Epidemiology, Characteristics, and One Health Control Strategies. *Antibiotics*, *12*(7), 1117. https://doi.org/10.3390/ANTIBIOTICS12071117

Argimón, S., David, S., Underwood, A., Abrudan, M., Wheeler, N. E., Kekre, M., Abudahab, K., Yeats, C. A., Goater, R., Taylor, B., Harste, H., Muddyman, D., Feil, E. J., Brisse, S., Holt, K., Donado-Godoy, P., Ravikumar, K. L., Okeke, I. N., Carlos, C., … Stelling, J. (2021). Rapid Genomic Characterization and Global Surveillance of Klebsiella Using Pathogenwatch. *Clinical Infectious Diseases*, *73*(Supplement_4), S325–S335. https://doi.org/10.1093/cid/ciab784

Arnold, K., Bordoli, L., Kopp, J., & Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics*, *22*(2), 195–201. https://doi.org/10.1093/BIOINFORMATICS/BTI770

Arnold, M., Echtermann, T., & Nathues, H. (2024). Infectious Enteric Diseasses in Pigs. *Production Diseases in Farm Animals*, 223–269. https://doi.org/10.1007/978-3-031-51788-4_11

Arredondo, A., Àlvarez, G., Nart, J., Mor, C., Blanc, V., & León, R. (2019). Detection and expression analysis of tet(B) in Streptococcus oralis. *Journal of Oral Microbiology*, *11*(1), 1643204. https://doi.org/10.1080/20002297.2019.1643204

Ateba, C. N., & Mbewe, M. (2014). Genotypic characterization of *Escherichia coli* O157:H7 isolates from different sources in the north-west province, South Africa, using enterobacterial repetitive intergenic consensus PCR analysis. *International Journal of Molecular Sciences*, *15*(6), 9735-9747. https://doi.org/10.3390/ijms15069735

Aydin, S., Personne, Y., Newire, E., Laverick, R., Russell, O., Roberts, A. P., & Enne, V. I. (2017). Presence of Type I-F CRISPR/Cas systems is associated with antimicrobial susceptibility in *Escherichia coli*. *Journal of Antimicrobial Chemotherapy*, *72*(8), 2213–2218. https://doi.org/10.1093/jac/dkx137

Badi, S., Cremonesi, P., Abbassi, M. S., Ibrahim, C., Snoussi, M., Bignoli, G., Luini, M., Castiglioni, B., & Hassen, A. (2018). Antibiotic resistance phenotypes and virulence-associated genes in *Escherichia coli* isolated from animals and animal food products in Tunisia. *FEMS Microbiology Letters*, *365*(10), 088. https://doi.org/10.1093/femsle/fny088

Bahgat, O. T., Rizk, D. E., Kenawy, H. I., & Barwa, R. (2024). Characterization of non-O157 enterohemorrhagic *Escherichia coli* isolated from different sources in Egypt. *BMC Microbiology*, *24*(1), 1–17. https://doi.org/10.1186/s12866-024-03636-3

Baloyi, T., Duvenage, S., Du Plessis, E., Villamizar-Rodríguez, G., & Korsten, L. (2022). Multidrug resistant *Escherichia coli* from fresh produce sold by street vendors in South African informal settlements. *International Journal of Environmental Health Research*, *32*(7), 1513-1528. https://doi.org/10.1080/09603123.2021.1896681

Barker, C. S., Inoue, T., Meshcheryakova, I. V., Kitanobo, S., & Samatey, F. A. (2016). Function of the conserved FHIPEP domain of the flagellar type III export apparatus, protein FlhA. *Molecular Microbiology*, *100*(2), 278–288. https://doi.org/10.1111/MMI.13315;PAGE:STRING:ARTICLE/CHAPTER

Baruah, S. M. (2024). The Versatile and Pathogenic Faces of *Escherichia coli*. *Assam Journal of Internal Medicine*, *14*(1), 1–2. https://doi.org/10.4103/AJOIM.AJOIM_15_24

Bazin, A., Gautreau, G., Médigue, C., Vallenet, D., & Calteau, A. (2020). panRGP: a pangenome-based method to predict genomic islands and explore their diversity. *Bioinformatics*, *36*(Supplement_2), i651–i658. https://doi.org/10.1093/bioinformatics/btaa792

Behera, M., Parmanand, Roshan, M., Rajput, S., Gautam, D., Vats, A., Ghorai, S. M., & De, S. (2023). Novel aadA5 and dfrA17 variants of class 1 integron in multidrug-resistant *Escherichia coli* causing bovine mastitis. *Applied Microbiology and Biotechnology*, *107*(1), 433–446. https://doi.org/10.1007/S00253-022-12304-3/FIGURES/6

Bernreiter-hofer, T., Schwarz, L., Müller, E., Cabal-rosel, A., Korus, M., Misic, D., Frankenfeld, K., Abraham, K., Grünzweil, O., Weiss, A., Feßler, A. T., Allerberger, F., Schwarz, S., Szostak, M. P., Ruppitsch, W., Ladinig, A., Spergser, J., Braun, S. D., Monecke, S., … Loncaric, I. (2021). The pheno- and genotypic characterization of porcine *Escherichia coli* isolates. *Microorganisms*, *9*(8), 1676. https://doi.org/10.3390/MICROORGANISMS9081676/S1

Bevan, E. R., Jones, A. M., & Hawkey, P. M. (2017). Global epidemiology of CTX-M β-lactamases: temporal and geographical shifts in genotype. *Journal of Antimicrobial Chemotherapy*, *72*(8), 2145–2155. https://doi.org/10.1093/JAC/DKX146

Bhandari, R. K., Pandey, A. K., Malhotra, S., Kakkar, A. K., Singh, S., Cohn, J., Holmes, A., Charani, E., & Shafiq, N. (2024). Addressing Challenges in Antibiotic Access: Barriers, Implications and Strategies for Solution: Challenges and Solutions in Antibiotic Access: R. K. Bhandari et al. *Pharmaceutical Medicine*, *38*(6), 387–397. https://doi.org/10.1007/S40290-024-00538-7/FIGURES/2

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T. G., Bertoni, M., Bordoli, L., & Schwede, T. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research*, *42*(W1), W252–W258. https://doi.org/10.1093/NAR/GKU340

Blankenship, H. M., Dietrich, S. E., Burgess, E., Wholehan, J., Soehnlen, M., & Manning, S. D. (2023). Whole-Genome Sequencing of Shiga Toxin-Producing *Escherichia coli* for Characterization and Outbreak Investigation. *Microorganisms*, *11*(5), 1298. https://doi.org/10.3390/MICROORGANISMS11051298/S1

Blasey, N., Rehrmann, D., Riebisch, A.K., Mühlen, S. (2023). Targeting bacterial pathogenesis by inhibiting viru-lence-associated Type III and Type IV secretion systems. *Frontiers in Cellular and Infection Microbiology*. 12, p.1065561. https://doi.org/10.3389/fcimb.2022.1065561

Bloomfield, S. F., Aiello, A. E., Cookson, B., O'Boyle, C., & Larson, E. L. (2007). The effectiveness of hand hygiene procedures in reducing the risks of infections in home and community settings including handwashing and alcohol-based hand sanitizers. *American Journal of Infection Control*, *35*(10), S27–S64. https://doi.org/10.1016/J.AJIC.2007.07.001

Blount, Z. D. (2015). The unexhausted potential of *E. coli*. *ELife*, 4, e05826 https://doi.org/10.7554/eLife.05826

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/BIOINFORMATICS/BTU170

Bolukaoto, J. Y., Kock, M. M., Strydom, K. A., Mbelle, N. M., & Ehlers, M. M. (2019). Molecular characteristics and genotypic diversity of enterohaemorrhagic *Escherichia coli* O157:H7 isolates in Gauteng region, South Africa. *Science of the Total Environment*, 692, 297-304. https://doi.org/10.1016/j.scitotenv.2019.07.119

Bolukaoto, J. Y., Singh, A., Alfinete, N., & Barnard, T. G. (2021). Occurrence of hybrid diarrhoeagenic *Escherichia coli* associated with multidrug resistance in environmental water, johannesburg, south africa. *Microorganisms*, *9*(10), 2163. https://doi.org/10.3390/MICROORGANISMS9102163/S1

Bortolaia, V., Kaas, R. S., Ruppe, E., Roberts, M. C., Schwarz, S., Cattoir, V., Philippon, A., Allesoe, R. L., Rebelo, A. R., Florensa, A. F., Fagelhauer, L., Chakraborty, T., Neumann, B., Werner, G., Bender, J. K., Stingl, K., Nguyen, M., Coppens, J., Xavier, B. B., … Aarestrup, F. M. (2020). ResFinder 4.0 for predictions of phenotypes from genotypes.

Journal of Antimicrobial Chemotherapy, *75*(12), 3491–3500. https://doi.org/10.1093/JAC/DKAA345

Brockhurst, M. A., Harrison, E., Hall, J. P. J., Richards, T., McNally, A., & MacLean, C. (2019). The Ecology and Evolution of Pangenomes. *Current Biology*, *29*(20), PR1094–R1103. https://doi.org/10.1016/j.cub.2019.08.012

Brown, E., Dessai, U., Mcgarry, S., & Gerner-Smidt, P. (2019). Use of Whole-Genome Sequencing for Food Safety and Public Health in the United States. In *Foodborne Pathogens and Disease* 16(7), 441-450. https://doi.org/10.1089/fpd.2019.2662

Bumunang, E. W., McAllister, T. A., Zaheer, R., Polo, R. O., Stanford, K., King, R., Niu, Y. D., & Ateba, C. N. (2019). Characterization of non-o157 *Escherichia coli* from cattle faecal samples in the North-West province of South Africa. *Microorganisms*, *7*(8), 272. https://doi.org/10.3390/microorganisms7080272

Caniça, M., Manageiro, V., Abriouel, H., Moran-Gilad, J., & Franz, C. M. A. P. (2019). Antibiotic resistance in foodborne bacteria. In *Trends in Food Science and Technology* 84, 41-44. https://doi.org/10.1016/j.tifs.2018.08.001

Capasso, C., & Supuran, C. T. (2014). Sulfa and trimethoprim-like drugs-antimetabolites acting as carbonic anhydrase, dihydropteroate synthase and dihydrofolate reductase inhibitors. *Journal of Enzyme Inhibition and Medicinal Chemistry*, *29*(3), 379–387. https://doi.org/10.3109/14756366.2013.787422

Carattoli, A., Zankari, E., Garciá-Fernández, A., Larsen, M. V., Lund, O., Villa, L., Aarestrup, F. M., & Hasman, H. (2014). In Silico detection and typing of plasmids using plasmidfinder and plasmid multilocus sequence typing. *Antimicrobial Agents and Chemotherapy*, *58*(7), 3895–3903. https://doi.org/10.1128/aac.02412-14

Cardoso, M. H., Cândido, E. S., Chan, L. Y., Der Torossian Torres, M., Oshiro, K. G. N., Rezende, S. B., Porto, W. F., Lu, T. K., De La Fuente-Nunez, C., Craik, D. J., & Franco, O. L. (2018). A Computationally Designed Peptide Derived from *Escherichia coli* as a Potential Drug Template for Antibacterial and Antibiofilm Therapies. *ACS Infectious Diseases*, *4*(12), 1727–1736. https://doi.org/10.1021/acsinfecdis.8b00219

Carroll, L. M., Pierneef, R., Mathole, A., Atanda, A., & Matle, I. (2022). Genomic Sequencing of Bacillus cereus Sensu Lato Strains Isolated from Meat and Poultry Products in South Africa Enables Inter- and Intranational Surveillance and Source Tracking. *Microbiology Spectrum*, *10*(3), e00700-22. https://doi.org/10.1128/spectrum.00700-22

Castellanos, L. R., Donado-Godoy, P., León, M., Clavijo, V., Arevalo, A., Bernal, J. F., Timmerman, A. J., Mevius, D. J., Wagenaar, J. A., & Hordijk, J. (2017). High heterogeneity of *Escherichia coli* sequence types harbouring ESBL/AmpC genes on IncI1 plasmids in the Colombian poultry chain. *PLoS ONE*, *12*(1), e0170777. https://doi.org/10.1371/journal.pone.0170777

Cepeda-Molero, M., Schüller, S., Frankel, G., Fernández, L. Á., Cepeda-Molero, M., Schüller, S., Frankel, G., & Fernández, L. Á. (2020). Systematic deletion of type III secretion system effectors in Enteropathogenic *E. coli* unveils the role of non-LEE effectors in a/E lesion formation. In *E. Coli Infections-Importance of Early Diagnosis and Efficient Treatment*. IntechOpen. https://www.intechopen.com/chapters/71365

Chantziaras, I., Boyen, F., Callens, B., & Dewulf, J. (2014). Correlation between veterinary antimicrobial use and antimicrobial resistance in food-producing animals: a report on seven countries. *Journal of Antimicrobial Chemotherapy*, *69*(3), 827–834. https://doi.org/10.1093/JAC/DKT443

Chapman, T. A., Wu, X. Y., Barchia, I., Bettelheim, K. A., Driesen, S., Trott, D., Wilson, M., & Chin, J. J. C. (2006). Comparison of virulence gene profiles of *Escherichia coli* strains isolated from healthy and diarrheic swine. *Applied and Environmental Microbiology*, *72*(7), 4782–4795. https://doi.org/10.1128/AEM.02885-05

Chekole, W. S., Potgieter, L., Adamu, H., Sternberg-Lewerin, S., Tessema, T. S., & Magnusson, U. (2025). Genomic insights into antimicrobial resistance and virulence of *E. coli* in central Ethiopia: a one health approach. *Frontiers in Microbiology*, *16*, 1597580. https://doi.org/10.3389/fmicb.2025.1597580

Chen, L., Song, N., Liu, B., Zhang, N., Alikhan, N. F., Zhou, Z., Zhou, Y., Zhou, S., Zheng, D., Chen, M., Hapeshi, A., Healey, J., Waterfield, N. R., Yang, J., & Yang, G. (2019). Genome-wide Identification and Characterization of a Superfamily of Bacterial Extracellular Contractile Injection Systems. *Cell Reports*, *29*(2), 511-521. https://doi.org/10.1016/j.celrep.2019.08.096

Chen, L., Zheng, D., Liu, B., Yang, J., & Jin, Q. (2016). VFDB 2016: hierarchical and refined dataset for big data analysis—10 years on. *Nucleic Acids Research*, *44*(D1), D694–D697. https://doi.org/10.1093/NAR/GKV1239

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884–i890. https://doi.org/10.1093/BIOINFORMATICS/BTY560

Clamp, M., Cuff, J., Searle, S.M., Barton, G.J. (2004). The jalview java alignment editor. *Bioinformatics*. 20(3), pp.426-427. https://doi.org/10.1093/bioinformatics/btg430

Clermont, O., Christenson, J. K., Denamur, E., & Gordon, D. M. (2013). The Clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. *Environmental Microbiology Reports*, *5*(1), 58–65. https://doi.org/10.1111/1758-2229.12019

Clermont, O., Dixit, O. V. A., Vangchhia, B., Condamine, B., Dion, S., Bridier-Nahmias, A., Denamur, E., & Gordon, D. (2019). Characterization and rapid identification of phylogroup G in *Escherichia coli*, a lineage with high virulence and antibiotic resistance potential. *Environmental Microbiology*, *21*(8), 3107–3117. https://doi.org/10.1111/1462-2920.14713

Cortimiglia, C., Borney, M. F., Bassi, D., & Cocconcelli, P. S. (2021). Genomic Investigation of Virulence Potential in Shiga Toxin *Escherichia coli* (STEC) Strains From a Semi-Hard Raw Milk Cheese. *Frontiers in Microbiology*, *11*, 629189. https://doi.org/10.3389/fmicb.2020.629189

Cosentino, S., Voldby Larsen, M., Møller Aarestrup, F., & Lund, O. (2013). PathogenFinder - Distinguishing Friend from Foe Using Bacterial Whole Genome Sequence Data. *PLOS ONE*, *8*(10), e77302. https://doi.org/10.1371/JOURNAL.PONE.0077302

Cummins, M. L., Reid, C. J., & Djordjevic, S. P. (2022). F Plasmid Lineages in *Escherichia coli* ST95: Implications for Host Range, Antibiotic Resistance, and Zoonoses. *MSystems*, *7*(1), e01212-21. https://doi.org/10.1128/msystems.01212-21

Dadgostar, P. (2019). Antimicrobial resistance: implications and costs. In *Infection and Drug Resistance*, 12, 3903-3910. https://doi.org/10.2147/IDR.S234610

Darphorn, T. S., Bel, K., Koenders-van Sint Anneland, B. B., Brul, S., & Ter Kuile, B. H. (2021). Antibiotic resistance plasmid composition and architecture in *Escherichia coli* isolates from meat. *Scientific Reports*, *11*(1), 2136. https://doi.org/10.1038/s41598-021-81683-w

Das, P., Singh, A. K., Mukherjee, S., Rajendran, K., Saha, D. R., Koley, H., & Basu, S. (2013). Composition of *Escherichia coli* population in the neonatal gut: Phylogroups and virulence determinants. *Journal of Medical Microbiology*, *62*(PART 11), 1680–1687. https://doi.org/10.1099/jmm.0.052225-0

Decano, A. G., Pettigrew, K., Sabiiti, W., Sloan, D. J., Neema, S., Bazira, J., Kiiru, J., Onyango, H., Asiimwe, B., & Holden, M. T. G. (2021). Pan-resistome characterization of uropathogenic *Escherichia coli* and klebsiella pneumoniae strains circulating in uganda and kenya, isolated from 2017–2018. *Antibiotics*, *10*(12), 1547. https://doi.org/10.3390/ANTIBIOTICS10121547/S1

Deekshit, V. K., & Srikumar, S. (2022). 'To be, or not to be'—The dilemma of 'silent' antimicrobial resistance genes in bacteria. *Journal of Applied Microbiology*, *133*(5), 2902–2914. https://doi.org/10.1111/JAM.15738

de Mesquita Souza Saraiva, M., Lim, K., do Monte, D. F. M., Givisiez, P. E. N., Alves, L. B. R., de Freitas Neto, O. C., Kariuki, S., Júnior, A. B., de Oliveira, C. J. B., & Gebreyes, W. A. (2021). Antimicrobial resistance in the globalized food chain: a One Health perspective applied to the poultry industry. *Brazilian Journal of Microbiology 2021 53:1*, *53*(1), 465–486. https://doi.org/10.1007/S42770-021-00635-8

Denamur, E., Clermont, O., Bonacorsi, S., & Gordon, D. (2021). The population genetics of pathogenic *Escherichia coli*. In *Nature Reviews Microbiology*, 19(1), 37-54. https://doi.org/10.1038/s41579-020-0416-x

Deng, W., Puente, J. L., Gruenheid, S., Li, Y., Vallance, B. A., Vázquez, A., Barba, J., Ibarra, J. A., O'Donnell, P., Metalnikov, P., Ashman, K., Lee, S., Goode, D., Pawson, T., & Finlay, B. B. (2004). Dissecting virulence: Systematic and functional analyses of a pathogenicity island. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(10), 3597–3602. https://doi.org/10.1073/PNAS.0400326101/SUPPL_FILE/00326FIG7.PDF

Deng, W., Marshall, N.C., Rowland, J.L., McCoy, J.M., Worrall, L.J., Santos, A.S., Strynadka, N.C., Finlay, B.B. (2017). Assembly, structure, function and regulation of type III secretion systems. *Nature Reviews Microbiology*, 15(6), pp.323-337. https://doi.org/10.1038/nrmicro.2017.20

Desmasures, N., Opportune, W., & Guéguen, M. (1997). Lactococcus spp., yeasts and Pseudomonas spp. on teats and udders of milking cows as potential sources of milk contamination. *International Dairy Journal*, *7*(10), 643–646. https://doi.org/10.1016/S0958-6946(97)00042-3

Dion, M. B., Shah, S. A., Deng, L., Thorsen, J., Stokholm, J., Krogfelt, K. A., Schjørring, S., Horvath, P., Allard, A., Nielsen, D. S., Petit, M. A., & Moineau, S. (2024). *Escherichia coli* CRISPR arrays from early life fecal samples preferentially target prophages. *The ISME Journal*, *18*(1), 5. https://doi.org/10.1093/ISMEJO/WRAE005

Djordjevic, S.P., Jarocki, V.M., Seemann, T., Cummins, M.L., Watt, A.E., Drigo, B., Wyrsch, E.R., Reid, C.J., Donner, E. & Howden, B.P. (2024). Genomic surveillance for antimicrobial resistance—a One Health perspective. *Nature Reviews Genetics*, 25(2), pp.142-157. https://doi.org/10.1038/s41576-023-00649-y

Dobrindt, U., Zdziarski, J., Salvador, E., & Hacker, J. (2010). Bacterial genome plasticity and its impact on adaptation during persistent infection. *International Journal of Medical Microbiology*, *300*(6), 363–366. https://doi.org/10.1016/J.IJMM.2010.04.010

Donkor, E. S., Newman, M. J., & Yeboah-Manu, D. (2012). Epidemiological aspects of non-human antibiotic usage and resistance: implications for the control of antibiotic resistance in Ghana. *Tropical Medicine & International Health*, *17*(4), 462–468. https://doi.org/10.1111/j.1365-3156.2012.02955.x

Doudna, J. A., & Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. *Science*, *346*(6213), 2258096. https://doi.org/10.1126/SCIENCE.1258096

Eagar, H., & Naidoo, V. (2017). Veterinary antimicrobial stewardship in South Africa. *International Biology Review*, *1*(2), 1-14. https://doi.org/10.18103/ibr.v1i2.1367

Ejaz, H., Younas, S., Qamar, M. U., Junaid, K., Abdalla, A. E., Abosalif, K. O. A., Alameen, A. A. M., Elamir, M. Y. M., Ahmad, N., Hamam, S. S. M., Salem, E. H. M., & Bukhari, S. N. A. (2021). Molecular Epidemiology of Extensively Drug-Resistant mcr Encoded Colistin-Resistant Bacterial Strains Co-Expressing Multifarious β-Lactamases. *Antibiotics 2021, Vol. 10, Page 467*, *10*(4), 467. https://doi.org/10.3390/ANTIBIOTICS10040467

Ekici, G. and Dümen, E. (2019). *Escherichia coli* and food safety. In *The universe of Escherichia coli*. IntechOpen.

Elliott, S. J., Sperandio, V., Giron, J. A., Shin, S., Mellies, J. L., Wainwright, L., Hutcheson, S. W., McDaniel, T. K., & Kaper, J. B. (2000). The locus of enterocyte effacement (LEE)-encoded regulator controls expression of both LEE- and non-LEE-encoded virulence factors in enteropathogenic and enterohemorrhagic *Escherichia coli*. *Infection and Immunity*, *68*(11), 6115–6126. https://doi.org/10.1128/iai.68.11.6115-6126.2000

Enciso-Martínez, Y., González-Aguilar, G. A., Martínez-Téllez, M. A., González-Pérez, C. J., Valencia-Rivera, D. E., Barrios-Villa, E., & Ayala-Zavala, J. F. (2022). Relevance of tracking the diversity of *Escherichia coli* pathotypes to reinforce food safety. *International Journal of Food Microbiology*, *374*, 109736. https://doi.org/10.1016/j.ijfoodmicro.2022.109736

Endale, H., Mathewos, M. & Abdeta, D. (2023). Potential causes of spread of antimicrobial resistance and preventive measures in one health perspective-a review. *Infection and drug resistance*, pp.7515-7545. https://doi.org/10.2147/IDR.S428837

Enne, V. I., Delsol, A. A., Roe, J. M., & Bennett, P. M. (2006). Evidence of antibiotic resistance gene silencing in *Escherichia coli*. *Antimicrobial Agents and Chemotherapy*, *50*(9), 3003–3010. https://doi.org/10.1128/aac.00137-06

Fang, L., Chen, R., Li, C., Sun, J., Liu, R., Shen, Y., & Guo, X. (2024). The association between the genetic structures of commonly incompatible plasmids in Gram-negative bacteria, their distribution and the resistance genes. *Frontiers in Cellular and Infection Microbiology*, 14, 1472876. https://doi.org/10.3389/fcimb.2024.1472876

Fatoba, D. O., Amoako, D. G., Akebe, A. L. K., Ismail, A., & Essack, S. Y. (2022). Genomic analysis of antibiotic-resistant *Enterococcus spp*. reveals novel enterococci strains and the spread of plasmid-borne *Tet(M)*, *Tet(L)* and *Erm(B)* genes from chicken litter to agricultural soil in South Africa. *Journal of Environmental Management*, 302, 114101. https://doi.org/10.1016/j.jenvman.2021.114101.

Felise, H.B., Nguyen, H.V., Pfuetzner, R.A., Barry, K.C., Jackson, S.R., Blanc, M.P., Bronstein, P.A., Kline, T., Miller, S.I. (2008). An inhibitor of gram-negative bacterial virulence protein secretion. *Cell host & microbe*, 4(4), pp.325-336. https://doi.org/10.1016/j.chom.2008.08.001

Feng, S., & Wang, Y. T. (2024). Pathogenicity islands: origins, structure, and roles in bacterial pathogenesis. *Molecular Medical Microbiology,* 3, 325–336. https://doi.org/10.1016/B978-0-12-818619-0.00035-6.

Fernandes, M. R., Sellera, F. P., Moura, Q., Esposito, F., Sabino, C. P., & Lincopan, N. (2020). Identification and genomic features of halotolerant extended-spectrum-β-lactamase (CTX-M)-producing *Escherichia coli* in urban-impacted coastal waters, Southeast Brazil. *Marine Pollution Bulletin*, 150, 110689. https://doi.org/10.1016/j.marpolbul.2019.110689

Figueras, J., Bravo, A. F., Ramaite, K., Deogratias Ekwanzala, M., Ndombo, M., & Momba, B. (2022). Prevalence and Molecular Characterisation of Extended-Spectrum Beta-Lactamase-Producing Shiga Toxin-Producing *Escherichia coli*, from Cattle Farm to Aquatic Environments. *Pathogens 2022*, *11*(6), 674. https://doi.org/10.3390/PATHOGENS11060674.

Flament-Simon, S. C., Nicolas-Chanoine, M. H., García, V., Duprilot, M., Mayer, N., Alonso, M. P., García-Meniño, I., Blanco, J. E., Blanco, M., & Blanco, J. (2020). Clonal Structure, Virulence Factor-encoding Genes and Antibiotic Resistance of *Escherichia coli*, Causing Urinary Tract Infections and Other Extraintestinal Infections in Humans in Spain and France during 2016. *Antibiotics 2020*, *9*(4), 161. https://doi.org/10.3390/ANTIBIOTICS9040161.

Foster-Nyarko, E., & Pallen, M. J. (2022). The microbial ecology of *Escherichia coli* in the vertebrate gut. *FEMS Microbiology Reviews*, *46*(3), fuac008. https://doi.org/10.1093/femsre/fuac008.

Founou, L. L., Amoako, D. G., Founou, R. C., & Essack, S. Y. (2018). Antibiotic Resistance in Food Animals in Africa: A Systematic Review and Meta-Analysis. In *Microbial Drug Resistance* (24)5, 648-665. https://doi.org/10.1089/mdr.2017.0383.

Franzin, F. M., & Sircili, M. P. (2015). Locus of Enterocyte Effacement: A Pathogenicity Island Involved in the Virulence of Enteropathogenic and Enterohemorragic *Escherichia coli* Subjected to a Complex Network of Gene Regulation. *BioMed Research International*, *2015*(1), 534738. https://doi.org/10.1155/2015/534738

Frost, L. S., Leplae, R., Summers, A. O., & Toussaint, A. (2005). Mobile genetic elements: The agents of open source evolution. *Nature Reviews Microbiology*, *3*(9), 722–732. https://doi.org/10.1038/NRMICRO1235;KWRD=LIFE+SCIENCES

Fu, Y., Nawrocki, E. M., M'ikanatha, N. M., & Dudley, E. G. (2024). Host species shapes genotype, antimicrobial resistance, and virulence profiles of enterotoxigenic *Escherichia coli* (ETEC) from livestock in the United States. *Applied and Environmental Microbiology*, 90(8), 00749-24. https://doi.org/10.1128/aem.00749-24.

Galindo-Méndez, M. (2020). Antimicrobial Resistance in *Escherichia coli*. *E. coli Infections - Importance of Early Diagnosis and Efficient Treatment*. https://doi.org/10.5772/INTECHOPEN.93115

Gambushe, S. M., Zishiri, O. T., & El Zowalaty, M. E. (2022). Review of *Escherichia coli* O157:H7 Prevalence, Pathogenicity, Heavy Metal and Antimicrobial Resistance, African Perspective. *Infection and Drug Resistance*, 15, 4645–4673. https://doi.org/10.2147/IDR.S365269

García-Gutiérrez, E., Almendros, C., Mojica, F. J. M., Guzmán, N. M., & García-Martínez, J. (2015). CRISPR Content Correlates with the Pathogenic Potential of *Escherichia coli*. *PLOS ONE*, 10(7), e0131935. https://doi.org/10.1371/JOURNAL.PONE.0131935

García, V., García-Meniño, I., Mora, A., Flament-Simon, S. C., Díaz-Jiménez, D., Blanco, J. E., Alonso, M. P., & Blanco, J. (2018). Co-occurrence of mcr-1, mcr-4 and mcr-5 genes in multidrug-resistant ST10 Enterotoxigenic and Shiga toxin-producing *Escherichia coli* in Spain (2006-2017). *International Journal of Antimicrobial Agents*, *52*(1), 104–108. https://doi.org/10.1016/J.IJANTIMICAG.2018.03.022

Garrity, G. M., Bell, J. A., & Lilburn, T. (2005). The Epsilonproteobacteria. *In Bergey's Manual of Systematic Bacteriology*, 2, 1145–1194. https://doi.org/10.1007/978-0-387-29298-4_4

Gautreau, G., Bazin, A., Gachet, M., Planel, R., Burlot, L., Dubois, M., Perrin, A., Médigue, C., Calteau, A., Cruveiller, S., Matias, C., Ambroise, C., Rocha, E. P. C., & Vallenet, D. (2020). PPanGGOLiN: Depicting microbial diversity via a partitioned pangenome graph. *PLOS Computational Biology*, *16*(3), e1007732. https://doi.org/10.1371/journal.pcbi.1007732

Gekenidis, M. T., Qi, W., Hummerjohann, J., Zbinden, R., Walsh, F., & Drissner, D. (2018). Antibiotic-resistant indicator bacteria in irrigation water: High prevalence of extended-spectrum beta-lactamase (ESBL)-producing *Escherichia coli*. *PLOS ONE*, *13*(11), e0207857. https://doi.org/10.1371/JOURNAL.PONE.0207857

Gemeda, B. A., Wieland, B., Alemayehu, G., Knight-Jones, T. J. D., Wodajo, H. D., Tefera, M., Kumbe, A., Olani, A., Abera, S., & Amenu, K. (2023). Antimicrobial Resistance of *Escherichia coli* Isolates from Livestock and the Environment in Extensive Smallholder Livestock Production Systems in Ethiopia. *Antibiotics 2023, Vol. 12, Page 941*, *12*(5), 941. https://doi.org/10.3390/ANTIBIOTICS12050941

Geurtsen, J., de Been, M., Weerdenburg, E., Zomer, A., McNally, A., & Poolman, J. (2022). Genomics and pathotypes of the many faces of *Escherichia coli*. *FEMS Microbiology Reviews*, *46*(6), fuac031. https://doi.org/10.1093/femsre/fuac031

Ghorbani, A., Khoshbakht, R., Kaboosi, H., Shirzad-Aski, H., & Ghadikolaii, F. P. (2022). Study on antibiotic resistance and phylogenetic comparison of avian-pathogenic *Escherichia coli* (APEC) and uropathogenic *Escherichia coli* (UPEC) isolates. *Veterinary Research Forum*, *13*(4), 569. https://doi.org/10.30466/VRF.2021.527563.3160

Gieryńska, M., Szulc-Dąbrowska, L., Struzik, J., Mielcarska, M. B., & Gregorczyk-Zboroch, K. P. (2022). Integrity of the Intestinal Barrier: The Involvement of Epithelial Cells and Microbiota—A Mutual Relationship. *Animals 2022, Vol. 12, Page 145*, *12*(2), 145. https://doi.org/10.3390/ANI12020145

Gong, T., Hayes, V. M., & Chan, E. K. F. (2021). Detection of somatic structural variants from short-read next-generation sequencing data. *Briefings in Bioinformatics*, *22*(3), bbaa056. https://doi.org/10.1093/bib/bbaa056

Green, E. R., & Mecsas, J. (2016). Bacterial Secretion Systems: An Overview. *Microbiology Spectrum*, *4*(1), 213-239. https://doi.org/10.1128/9781555819286.ch8

Grossman, A. S., Mauer, T. J., Forest, K. T., & Goodrich-Blair, H. (2021). A widespread bacterial secretion system with diverse substrates. *MBio*, *12*(4), 10-1128. https://doi.org/10.1128/mbio.01956-21

Gupta, R., & Gupta, N. (2021). Fundamentals of Bacterial Physiology and Metabolism. *Fundamentals of Bacterial Physiology and Metabolism*, 1(670), 289-305. https://doi.org/10.1007/978-981-16-0723-3/

Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics*, *29*(8), 1072–1075. https://doi.org/10.1093/BIOINFORMATICS/BTT086

Guruprasad, K., Reddy, B. V. B., & Pandit, M. W. (1990). Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. *Protein Engineering*, *4*(2), 155–161. https://doi.org/10.1093/PROTEIN/4.2.155

Halaji, M., Fayyazi, A., Rajabnia, M., Zare, D., Pournajaf, A., & Ranjbar, R. (2022). Phylogenetic Group Distribution of Uropathogenic *Escherichia coli* and Related Antimicrobial Resistance Pattern: A Meta-Analysis and Systematic Review. *Frontiers in Cellular and Infection Microbiology*, *12*, 790184. https://doi.org/10.3389/fcimb.2022.790184

Hazen, T. H., Daugherty, S. C., Shetty, A., Mahurkar, A. A., White, O., Kaper, J. B., & Rasko, D. A. (2015). RNA-Seq analysis of isolate- and growth phase-specific differences in the global transcriptomes of enteropathogenic *Escherichia coli* prototype isolates. *Frontiers in Microbiology*, *6*(JUN), 145098. https://doi.org/10.3389/FMICB.2015.00569/BIBTEX

Higginson, E. E., Abu Sayeed, M., Dias, J. P., Shetty, V., Ballal, M., Srivastava, S. K., Willis, I., Qadri, F., Dougan, G., & Mutreja, A. (2022). Microbiome Profiling of Enterotoxigenic *Escherichia coli* (ETEC) Carriers Highlights Signature Differences between Symptomatic and Asymptomatic Individuals. *MBio*, *13*(3), e00157-22. https://doi.org/10.1128/mbio.00157-22

Hotinger, J. A., & May, A. E. (2020). Antibodies inhibiting the type iii secretion system of gram-negative pathogenic bacteria. In *Antibodies* (Vol. 9, Issue 3). https://doi.org/10.3390/antib9030035

Hotinger, J. A., Pendergrass, H. A., & May, A. E. (2021). Molecular Targets and Strategies for Inhibition of the Bacterial Type III Secretion System (T3SS); Inhibitors Directly Binding to T3SS Components. *Biomolecules 2021, Vol. 11, Page 316*, *11*(2), 316. https://doi.org/10.3390/BIOM11020316

Hounkpe, E.C., Sessou, P., Farougou, S., Daube, G., Delcenserie, V., Azokpota, P., Korsak, N. (2023). Prevalence, antibiotic resistance, and virulence gene profile of *Escherichia coli* strains shared between food and other sources in Africa: A systematic review. *Veterinary World*, 16(10), p.2016. https://doi.org/10.14202/vetworld.2023.2016-2028

Hueck, C. J. (1998). Type III Protein Secretion Systems in Bacterial Pathogens of Animals and Plants. *Microbiology and Molecular Biology Reviews*, *62*(2), 379-433. https://doi.org/10.1128/mmbr.62.2.379-433.1998

Hu, J., Afayibo, D. J. A., Zhang, B., Zhu, H., Yao, L., Guo, W., Wang, X., Wang, Z., Wang, D., Peng, H., Tian, M., Qi, J., & Wang, S. (2022). Characteristics, pathogenic mechanism, zoonotic potential, drug resistance, and prevention of avian pathogenic *Escherichia coli* (APEC). *Frontiers in Microbiology*, *13*, 1049391. https://doi.org/10.3389/fmicb.2022.1049391

Hu, T., Chitnis, N., Monos, D., & Dinh, A. (2021). Next-generation sequencing technologies: An overview. *Human Immunology*, *82*(11), 801-811. https://doi.org/10.1016/j.humimm.2021.02.012

Igwaran, A., Iweriebor, B. C., & Okoh, A. I. (2018). Molecular characterization and antimicrobial resistance pattern of *Escherichia coli* recovered from wastewater treatment plants in Eastern Cape South Africa. *International Journal of Environmental Research and Public Health*, *15*(6), 1237. https://doi.org/10.3390/ijerph15061237

Im, S. B., Gupta, S., Jain, M., Chande, A. T., Carleton, H. A., Jordan, I. K., & Rishishwar, L. (2021). Genome-Enabled Molecular Subtyping and Serotyping for Shiga Toxin-Producing *Escherichia coli*. *Frontiers in Sustainable Food Systems*, *5,* 752873. https://doi.org/10.3389/fsufs.2021.752873

Ingle, D. J., Levine, M. M., Kotloff, K. L., Holt, K. E., & Robins-Browne, R. M. (2018). Dynamics of antimicrobial resistance in intestinal *Escherichia coli* from children in community settings in South Asia and sub-Saharan Africa. *Nature Microbiology, 3*(9), 1063–1073. https://doi.org/10.1038/s41564-018-0217-4

Islam, M. A., Bose, P., Rahman, M. Z., Muktaruzzaman, M., Sultana, P., Ahamed, T., & Khatun, M. M. (2024). A review of antimicrobial usage practice in livestock and poultry production and its consequences on human and animal health. *Journal of Advanced Veterinary and Animal Research*, *11*(3), 675. https://doi.org/10.5455/JAVAR.2024.K817

Iwan, E., Zając, M., Bomba, A., Olejnik, M., Skarżyńska, M., Wasiński, B., Wieczorek, K., Tłuścik, K., & Wasyl, D. (2024). Phylogenetics and Mobilization of Genomic Traits of Cephalosporin-Resistant *Escherichia coli* Originated from Retail Meat. *Pathogens*, *13*(8), 700. https://doi.org/10.3390/PATHOGENS13080700/S1

Jaja, I. F., Jaja, C. J. I., Chigor, N. V., Anyanwu, M. U., Maduabuchi, E. K., Oguttu, J. W., & Green, E. (2020). Antimicrobial Resistance Phenotype of Staphylococcus aureus and *Escherichia coli* Isolates Obtained from Meat in the Formal and Informal Sectors in South Africa. *BioMed Research International*, *2020*, 1, 3979482. https://doi.org/10.1155/2020/3979482

Jaja, I. F., Oguttu, J., Jaja, C. J. I., & Green, E. (2020). Prevalence and distribution of antimicrobial resistance determinants of *Escherichia coli* isolates obtained from meat in South Africa. *PLoS ONE*, *15*(5), 0216914. https://doi.org/10.1371/journal.pone.0216914

Janda, J. M., & Abbott, S. L. (2021). The changing face of the family enterobacteriaceae (Order: Enterobacterales): New members, taxonomic issues, geographic expansion, and new diseases and disease syndromes. *Clinical Microbiology Reviews*, *34*(2), 1–45. https://doi.org/10.1128/cmr.00174-20

Jang, J., Hur, H. G., Sadowsky, M. J., Byappanahalli, M. N., Yan, T., & Ishii, S. (2017). Environmental *Escherichia coli*: ecology and public health implications—a review. In *Journal of Applied Microbiology*, 123(3), 570-581. https://doi.org/10.1111/jam.13468

Jenia, G., Dor, B., & Neta, S. M. (2021). Transmembrane domains of type III-secreted proteins affect bacterial-host interactions in enteropathogenic *E. coli*. *Virulence*, *12*(1), 902–917. https://doi.org/10.1080/21505594.2021.1898777;PAGE:STRING:ARTICLE/CHAPTER

Jesser, K. J., & Levy, K. (2020). Updates on defining and detecting diarrheagenic *Escherichia coli* pathotypes. *Current Opinion in Infectious Diseases*, *33*(5), 372-380. https://doi.org/10.1097/QCO.0000000000000665

Jiang, H., Cheng, H., Liang, Y., Yu, S., Yu, T., Fang, J., & Zhu, C. (2019). Diverse Mobile Genetic Elements and Conjugal Transferability of Sulfonamide Resistance Genes (sul1, sul2, and sul3) in *Escherichia coli* Isolates From Penaeus vannamei and Pork From Large Markets in Zhejiang, China. *Frontiers in Microbiology*, *10*, 464511. https://doi.org/10.3389/fmicb.2019.01787

Jian, Z., Zeng, L., Xu, T., Sun, S., Yan, S., Yang, L., Huang, Y., Jia, J., & Dou, T. (2021). Antibiotic resistance genes in bacteria: Occurrence, spread, and control. *Journal of Basic Microbiology*, *61*(12), 1049–1070. https://doi.org/10.1002/jobm.202100201

Joensen, K. G., Tetzschner, A. M. M., Iguchi, A., Aarestrup, F. M., & Scheutz, F. (2015). Rapid and easy in silico serotyping of *Escherichia coli* isolates by use of whole-genome sequencing data. *Journal of Clinical Microbiology*, *53*(8), 2410–2426. https://doi.org/10.1128/JCM.00008-15/SUPPL_FILE/ZJM999094355SD2.XLSX

Johnson, J. R., & Russo, T. A. (2018). Molecular Epidemiology of Extraintestinal Pathogenic *Escherichia coli*. *EcoSal Plus*, *8*(1), 10-1128. https://doi.org/10.1128/ecosalplus.esp-0004-2017

Kalule, J. B., Bester, L. A., Banda, D. L., Derra, F. A., Msefula, C., Smith, A. M., Ajayi, A., Kumburu, H., Kwenda, G., Yamba, K., Mwaba, J., Fakim, Y. J., Sithole, N., Kanzi, A. M., Njage, P. M. K., Chikuse, F., Tessema, S. K., Smith, S. I., & Foster-Nyarko, E. (2024). Molecular Epidemiology and AMR Perspective of Diarrhoeagenic *Escherichia coli* in Africa: A Systematic Review and Meta-analysis. *Journal of Epidemiology and Global Health*, *14*(4), 1381–1396. https://doi.org/10.1007/S44197-024-00301-W/FIGURES/4

Kang, H. J., & Lee, Y. J. (2022). Distribution of CRISPR in *Escherichia coli* Isolated from Bulk Tank Milk and Its Potential Relationship with Virulence. *Animals*, *12*(4), 503. https://doi.org/10.3390/ANI12040503/S1

Kariuki, S., Kering, K., Wairimu, C., Onsare, R., & Mbae, C. (2022). Antimicrobial Resistance Rates and Surveillance in Sub-Saharan Africa: Where Are We Now? *Infection and Drug Resistance*, *15*, 3589–3609. https://doi.org/10.2147/IDR.S342753

Kathayat, D., Lokesh, D., Ranjit, S., & Rajashekara, G. (2021). Avian Pathogenic *Escherichia coli* (APEC): An Overview of Virulence and Pathogenesis Factors, Zoonotic Potential, and Control Strategies. *Pathogens 2021*, *10*(4), 467. https://doi.org/10.3390/PATHOGENS10040467

Kaur, P., & Dudeja, P. K. (2023). Pathophysiology of Enteropathogenic *Escherichia coli*-induced Diarrhea. *Newborn (Clarksville, Md.)*, *2*(1), 102. https://doi.org/10.5005/JP-JOURNALS-11002-0056

Kazemnia, A., Ahmadi, M., & Dilmaghani, M. (2014). Antibiotic Resistance Pattern of Different *Escherichia coli* Phylogenetic Groups Isolated from Human Urinary Tract Infection and Avian Colibacillosis. *Iranian Biomedical Journal*, *18*(4), 219. https://doi.org/10.6091/IBJ.1394.2014

Keddy, K. H., Hoffmann, S., Founou, L. L., Estrada-Garcia, T., Gobena, T., Havelaar, A. H., Jakobsen, L. S., Kubota, K., Law, C., Lake, R., Minato, Y., Al-Natour, F. N. R., Pires, S. M., Rachmawati, T., Sripa, B., Torgerson, P., & Walter, E. S. (2025). Quantifying national burdens of foodborne disease-Four imperatives for global impact. *PLOS Global Public Health*, *5*, e0004309. https://doi.org/10.1371/journal.pgph.0004309

Khabo-Mmekoa, C. M., Genthe, B., & Momba, M. N. B. (2022). Enteric Pathogens Risk Factors Associated with Household Drinking Water: A Case Study in Ugu District Kwa-Zulu Natal Province, South Africa. *International Journal of Environmental Research and Public Health 2022*, *19*(8), 4431. https://doi.org/10.3390/IJERPH19084431

Khairy, R. M., Mohamed, E. S., Ghany, H. M. A., & Abdelrahim, S. S. (2019). Phylogenic classification and virulence genes profiles of uropathogenic *E. coli* and diarrhegenic *E. coli* strains isolated from community acquired infections. *PLOS ONE*, *14*(9), e0222441. https://doi.org/10.1371/JOURNAL.PONE.0222441

Kiiti, R. W., Komba, E. V., Msoffe, P. L., Mshana, S. E., Rweyemamu, M., & Matee, M. I. N. (2021). Antimicrobial Resistance Profiles of *Escherichia coli* Isolated from Broiler and

Layer Chickens in Arusha and Mwanza, Tanzania. *International Journal of Microbiology*, *2021*(1), 6759046. https://doi.org/10.1155/2021/6759046

Kirk, M. D., Pires, S. M., Black, R. E., Caipo, M., Crump, J. A., Devleesschauwer, B., Döpfer, D., Fazil, A., Fischer-Walker, C. L., Hald, T., Hall, A. J., Keddy, K. H., Lake, R. J., Lanata, C. F., Torgerson, P. R., Havelaar, A. H., & Angulo, F. J. (2015). World Health Organization Estimates of the Global and Regional Disease Burden of 22 Foodborne Bacterial, Protozoal, and Viral Diseases, 2010: A Data Synthesis. *PLoS Medicine*, *12*(12), e1001940. https://doi.org/10.1371/journal.pmed.1001921

Kocsis, B., Gulyás, D., & Szabó, D. (2022). Emergence and Dissemination of Extraintestinal Pathogenic High-Risk International Clones of *Escherichia coli*. *Life 2022*, *12*(12), 2077. https://doi.org/10.3390/LIFE12122077

Koonin, E. V., Makarova, K. S., & Zhang, F. (2017). Diversity, classification and evolution of CRISPR-Cas systems. *Current Opinion in Microbiology*, *37*, 67–78. https://doi.org/10.1016/j.mib.2017.05.008

Koutsoumanis, K., Allende, A., Álvarez-Ordóñez, A., Bolton, D., Bover-Cid, S., Chemaly, M., Davies, R., De Cesare, A., Herman, L. and Hilbert, F., (2022). Transmission of antimicrobial resistance (AMR) during animal transport. *EFSA Journal*, *20*(10), e07586. https://doi.org/10.2903/j.efsa.2022.7586

Ksiezarek, M., Novais, Â., & Peixe, L. (2022). The darkest place is under the candlestick-healthy urogenital tract as a source of worldwide disseminated extraintestinal pathogenic *Escherichia coli* lineages. *Microorganisms*, *10*(1), 27. https://doi.org/10.3390/MICROORGANISMS10010027/S1

Kuan, N. L., Chen, Y. P., Shien, J. H., & Yeh, K. S. (2024). Characteristics of the extended-spectrum-β-lactamase-producing *Escherichia coli* isolated from diseased livestock and poultry in Taiwan. *Scientific Reports*, *14*(1), 1–8. https://doi.org/10.1038/S41598-024-80943-9;SUBJMETA=326,631,692,699;KWRD=DISEASES,MICROBIOLOGY

Kuebutornye, F. K. A., Liao, J., Pang, H., Lu, Y., Ayiku, S., & Sakyi, M. E. (2018). Molecular Cloning and Bioinformatics Analysis of T3SS Inner Membrane Ring HrpQ from *Vibrio harveyi*. *Genomics and Applied Biology*, 9(7), 40-47. https://doi.org/10.5376/GAB.2018.09.0007

Kuhlen, L., Abrusci, P., Johnson, S., Gault, J., Deme, J., Caesar, J., Dietsche, T., Mebrhatu, M. T., Ganief, T., Macek, B., Wagner, S., Robinson, C. V., & Lea, S. M. (2018). Author Correction: Structure of the core of the type III secretion system export apparatus. *Nature Structural & Molecular Biology*, *25*(8), 743–743. https://doi.org/10.1038/S41594-018-0095-8

Kumar, N., Chatterjee, K., Deka, S., Shankar, R., & Kalita, D. (2021). *Increased Isolation of Extended-Spectrum Beta-Lactamase-Producing Escherichia coli From Community-Onset Urinary Tract Infection Cases in Uttarakhand, India*, 13(3), e13837. https://doi.org/10.7759/cureus.13837

Kushwaha, S. K., Wu, Y., Avila, H. L., Anand, A., Sicheritz-Pontén, T., Millard, A., Marathe, S. A., & Nobrega, F. L. (2024). Comprehensive blueprint of Salmonella genomic plasticity identifies hotspots for pathogenicity genes. *PLOS Biology*, *22*(8), e3002746. https://doi.org/10.1371/journal.pbio.3002746

Kyte, J., & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*, *157*(1), 105–132. https://doi.org/10.1016/0022-2836(82)90515-0

Larsen, M. V., Cosentino, S., Rasmussen, S., Friis, C., Hasman, H., Marvig, R. L., Jelsbak, L., Sicheritz-Pontén, T., Ussery, D. W., Aarestrup, F. M., & Lund, O. (2012). Multilocus sequence typing of total-genome-sequenced bacteria. *Journal of Clinical Microbiology*, *50*(4), 1355–1361. https://doi.org/10.1128/JCM.06094-11/SUPPL_FILE/JCM-JCM06094-11-S02.PDF

Lawrence, J. G., & Ochman, H. (1998). Molecular archaeology of the *Escherichia coli* genome. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(16), 9413-9417. https://doi.org/10.1073/pnas.95.16.9413

Lee, W., Ha, J., Choi, J., Jung, Y., Kim, E., An, E. S., Kim, S. H., Shin, H., Ryu, S., Kim, S. H., & Kim, H. Y. (2024). Genetic and virulence characteristics of hybrid Shiga toxin-producing and atypical enteropathogenic *Escherichia coli* strains isolated in South Korea. *Frontiers in Microbiology*, *15*, 1398262. https://doi.org/10.3389/fmicb.2024.1398262

Li, D., Elankumaran, P., Kudinha, T., Kidsley, A. K., Trott, D. J., Jarocki, V. M., & Djordjevic, S. P. (2023). Dominance of *Escherichia coli* sequence types ST73, ST95, ST127 and ST131 in Australian urine isolates: a genomic analysis of antimicrobial resistance and virulence linked to F plasmids. *Microbial Genomics*, *9*(7), 001068. https://doi.org/10.1099/MGEN.0.001068/CITE/REFWORKS

Lima, T., Domingues, S., & Silva, G. J. Da. (2020). Manure as a potential hotspot for antibiotic resistance dissemination by horizontal gene transfer events. In *Veterinary Sciences*, 7, (3), 110. https://doi.org/10.3390/vetsci7030110

Li, M., Li, Z., Zhong, Q., Liu, J., Han, G., Li, Y., & Li, C. (2022). Antibiotic resistance of fecal carriage of *Escherichia coli* from pig farms in China: a meta-analysis. *Environmental Science and Pollution Research*, *29*(16), 22989–23000. https://doi.org/10.1007/s11356-021-17339-z

Ling, Z., Yin, W., Shen, Z., Wang, Y., Shen, J., & Walsh, T. R. (2020). Epidemiology of mobile colistin resistance genes mcr-1 to mcr-9. *Journal of Antimicrobial Chemotherapy*, *75*(11), 3087–3095. https://doi.org/10.1093/JAC/DKAA205

Li, T., Qiang, N., Bao, Y., Li, Y., Zhao, S., Chong, K. C., Deng, X., Zhang, X., Ran, J., & Han, L. (2024). Global burden of enteric infections related foodborne diseases, 1990–2021: findings from the Global Burden of Disease Study 2021. *Science in One Health*, *3*, 100075. https://doi.org/10.1016/j.soh.2024.100075

Liu, B., Conroy, J. M., Morrison, C. D., Odunsi, A. O., Qin, M., Wei, L., Trump, D. L., Johnson, C. S., Liu, S., & Wang, J. (2015). Structural variation discovery in the cancer genome using next generation sequencing: Computational solutions and perspectives. *Oncotarget*, *6*(8), 5477–5489. https://doi.org/10.18632/oncotarget.3491

Liu, B., Furevi, A., Perepelov, A. V., Guo, X., Cao, H., Wang, Q., Reeves, P. R., Knirel, Y. A., Wang, L., & Widmalm, G. (2020). Structure and genetics of *Escherichia coli* O antigens. In *FEMS Microbiology Reviews* , 44(6), 685-683. https://doi.org/10.1093/femsre/fuz028

Li, Y., & Ge, X. (2023). Role of Berberine as a Potential Efflux Pump Inhibitor against MdfA from *Escherichia coli*: In Vitro and In Silico Studies. *Microbiology Spectrum*, *11*(2), e03324-22. https://doi.org/10.1128/spectrum.03324-22

Llorente, M. T., Escudero, R., Ramiro, R., Remacha, M. A., Martínez-Ruiz, R., Galán-Sánchez, F., de Frutos, M., Elía, M., Onrubia, I., & Sánchez, S. (2023). Enteroaggregative *Escherichia coli* as etiological agent of endemic diarrhea in Spain: A prospective multicenter prevalence study with molecular characterization of isolates. *Frontiers in Microbiology*, *14*, 1120285. https://doi.org/10.3389/fmicb.2023.1120285

Louwen, R., Staals, R. H. J., Endtz, H. P., van Baarlen, P., & van der Oost, J. (2014). The Role of CRISPR-Cas Systems in Virulence of Pathogenic Bacteria. *Microbiology and Molecular Biology Reviews*, *78*(1), 74–88. https://doi.org/10.1128/MMBR.00039-13/ASSET/E8B35881-37BA-4D8D-A91E-3E052CB5B22D/ASSETS/GRAPHIC/ZMR9990923520005.JPEG

Luna-Guevara, J. J., Arenas-Hernandez, M. M. P., Martínez De La Peña, C., Silva, J. L., & Luna-Guevara, M. L. (2019). The Role of Pathogenic *E. coli* in Fresh Vegetables: Behavior, Contamination Factors, and Preventive Measures. *International Journal of Microbiology*, *1*, 2894328. https://doi.org/10.1155/2019/2894328

Lupindu, A. M. (2017). Isolation and Characterization of *Escherichia coli* from Animals, Humans, and Environment. *Escherichia coli. Recent Advances on Physiology, Pathogenesis and Biotechnological Applications*, *IntechOpen Limited*, 187-206 . https://doi.org/10.5772/67390

Lyimo, B., & Sonola, V. (2025). Antimicrobial Resistance Gene Distribution and Population Structure of *Escherichia coli* isolated from Humans, Livestock, and the Environment: Insights from a One Health Approach. *MedRxiv*, 04, 25325408. https://doi.org/10.1101/2025.04.07.25325408

Machado, M. A. M., Panzenhagen, P., Lázaro, C., Rojas, M., Figueiredo, E. E. de S., & Conte-Junior, C. A. (2024). Unveiling the High Diversity of Clones and Antimicrobial Resistance Genes in *Escherichia coli* Originating from ST10 across Different Ecological Niches. *Antibiotics*, *13*(8), 737. https://doi.org/10.3390/ANTIBIOTICS13080737/S1

Madoroba, E., Malokotsa, K. P., Ngwane, C., Lebelo, S., & Magwedere, K. (2022). Presence and Virulence Characteristics of Shiga Toxin *Escherichia coli* and Non-Shiga Toxin-Producing *Escherichia coli* O157 in Products from Animal Protein Supply Chain Enterprises in South Africa. *Foodborne Pathogens and Disease*, *19*(6), 386–393. https://doi.org/10.1089/fpd.2021.0062

Mahfouz, N., Caucci, S., Achatz, E., Semmler, T., Guenther, S., Berendonk, T. U., & Schroeder, M. (2018). High genomic diversity of multi-drug resistant wastewater *Escherichia coli*. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-27292-6

Mainil, J. (2013). *Escherichia coli* virulence factors. *Veterinary Immunology and Immunopathology*, *152*(1–2), 2–12. https://doi.org/10.1016/J.VETIMM.2012.09.032

Makarova, K. S., Koonin, E. V., & Albers, S. V. (2016). Diversity and evolution of type IV pili systems in Archaea. *Frontiers in Microbiology*, *7*(MAY), 195821. https://doi.org/10.3389/FMICB.2016.00667/BIBTEX

Malahlela, M. N., Cenci-Goga, B. T., Marufu, M. C., Fonkui, T. Y., Grispoldi, L., Etter, E., Kalake, A., & Karama, M. (2022). Occurrence, Serotypes and Virulence Characteristics of Shiga-Toxin-Producing *Escherichia coli* Isolates from Goats on Communal Rangeland in South Africa. *Toxins*, *14*(5), 353. https://doi.org/10.3390/TOXINS14050353/S1

Malema, M. S., Abia, A. L. K., Tandlich, R., Zuma, B., Kahinda, J. M. M., & Ubomba-Jaswa, E. (2018). Antibiotic-Resistant Pathogenic *Escherichia coli* Isolated from Rooftop Rainwater-Harvesting Tanks in the Eastern Cape, South Africa. *International Journal of Environmental Research and Public Health 2018* , *15*(5), 892. https://doi.org/10.3390/IJERPH15050892

Malesa, R., Pierneef, R., Magwedere, K., Mafuna, T., Matle, I., (2024). Genomic characterisation of generic *Escherichia coli* from food-producing animals and products of animal origin in South Africa. *Frontiers in Bacteriology*, *3*, 1432292. https://doi.org/10.3389/fbrio.2024.1432292

Maluta, R. P., Logue, C. M., Casas, M. R. T., Meng, T., Guastalli, E. A. L., Rojas, T. C. G., Montelli, A. C., Sadatsune, T., Ramos, M. D. C., Nolan, L. K., & Silveira, W. D. Da. (2014). Overlapped Sequence Types (STs) and Serogroups of Avian Pathogenic (APEC) and Human Extra-Intestinal Pathogenic (ExPEC) *Escherichia coli* Isolated in Brazil. *PLOS ONE*, *9*(8), e105016. https://doi.org/10.1371/JOURNAL.PONE.0105016

Manges, A.R. (2016). *Escherichia coli* and urinary tract infections: the role of poultry-meat. *Clinical Microbiology and Infection*, 22(2), pp.122-129. https://doi.org/10.1016/j.cmi.2015.11.010

Manyi-Loh, C. E., & Lues, R. (2023). A South African Perspective on the Microbiological and Chemical Quality of Meat: Plausible Public Health Implications. *Microorganisms 2023*, *11*(10), 2484. https://doi.org/10.3390/MICROORGANISMS11102484

Mardis, E. R. (2013). Next-generation sequencing platforms. Annual Review of Analytical Chemistry, 6, 287-303. https://doi.org/10.1146/annurev-anchem-062012-092628

Mare, A. D., Ciurea, C. N., Man, A., Tudor, B., Moldovan, V., Decean, L., & Toma, F. (2021). Enteropathogenic *Escherichia coli*—A Summary of the Literature. *Gastroenterology Insights 2021*, *12*(1), 28–40. https://doi.org/10.3390/GASTROENT12010004

Martak, D., Henriot, C.P. & Hocquet, D. (2024). Environment, animals, and food as reservoirs of antibiotic-resistant bacteria for humans: One health or more?. *Infectious Diseases Now*, 54(4), p.104895. https://doi.org/10.1016/j.idnow.2024.104895

Martinez-Medina, M. (2021). Special issue: Pathogenic *Escherichia coli*: Infections and therapies. In *Antibiotics*,10(2), 112. https://doi.org/10.3390/antibiotics10020112

Massella, E., Giacometti, F., Bonilauri, P., Reid, C. J., Djordjevic, S. P., Merialdi, G., Bacci, C., Fiorentini, L., Massi, P., Bardasi, L., Rubini, S., Savini, F., Serraino, A., & Piva, S. (2021). Antimicrobial resistance profile and expec virulence potential in commensal *Escherichia coli* of multiple sources. *Antibiotics*, *10*(4), 351. https://doi.org/10.3390/ANTIBIOTICS10040351/S1

Massella, E., Reid, C. J., Cummins, M. L., Anantanawat, K., Zingali, T., Serraino, A., Piva, S., Giacometti, F., & Djordjevic, S. P. (2020). Snapshot Study of Whole Genome Sequences of *Escherichia coli* from Healthy Companion Animals, Livestock, Wildlife, Humans and Food in Italy. *Antibiotics 2020*, *9*(11), 782. https://doi.org/10.3390/ANTIBIOTICS9110782

Masters, N., Wiegand, A., Ahmed, W., & Katouli, M. (2011). *Escherichia coli* virulence genes profile of surface waters as an indicator of water quality. *Water Research*, *45*(19), 6321–6333. https://doi.org/10.1016/J.WATRES.2011.09.018

Mbanga, J., Amoako, D. G., Abia, A. L. K., Allam, M., Ismail, A., & Essack, S. Y. (2021). Genomic Insights of Multidrug-Resistant *Escherichia coli* From Wastewater Sources and Their Association With Clinical Pathogens in South Africa. *Frontiers in Veterinary Science*, *8,* 636715. https://doi.org/10.3389/fvets.2021.636715

Mbelle, N. M., Feldman, C., Osei Sekyere, J., Maningi, N. E., Modipane, L., & Essack, S. Y. (2019). The Resistome, Mobilome, Virulome and Phylogenomics of Multidrug-Resistant *Escherichia coli* Clinical Isolates from Pretoria, South Africa. *Scientific Reports*, *9*(1), 16457. https://doi.org/10.1038/s41598-019-52859-2

Meena, P. R., Priyanka, P., & Singh, A. P. (2023). Extraintestinal pathogenic *Escherichia coli* (ExPEC) reservoirs, and antibiotics resistance trends: a one-health surveillance for risk analysis from "farm-to-fork". *Letters in Applied Microbiology*, *76*(1), 1–12. https://doi.org/10.1093/LAMBIO/OVAC016

Meena, P. R., Yadav, P., Hemlata, H., Tejavath, K. K., & Singh, A. P. (2021). Poultry-origin extraintestinal *Escherichia coli* strains carrying the traits associated with urinary tract infection, sepsis, meningitis and avian colibacillosis in India. *Journal of Applied Microbiology*, *130*(6), 2087–2101. https://doi.org/10.1111/JAM.14905

Mengistu, D. Y., & Mengesha, Y. (2023). New approaches for severity intervention and rapid diagnosis of enterohemorrhagic *Escherichia coli*: a review. *All Life*, *16*(1), 2218582. https://doi.org/10.1080/26895293.2023.2218582;WGROUP:STRING:PUBLICATION

Messele, Y. E., Trott, D. J., Hasoon, M. F., Veltman, T., McMeniman, J. P., Kidd, S. P., Djordjevic, S. P., Petrovski, K. R., & Low, W. Y. (2023). Phylogenetic Analysis of *Escherichia coli* Isolated from Australian Feedlot Cattle in Comparison to Pig Faecal and Poultry/Human Extraintestinal Isolates. *Antibiotics*, *12*(5), 895. https://doi.org/10.3390/ANTIBIOTICS12050895/S1

Meyers, B. C., & McLellan, S. L. (2022). Influence of Nutrients and the Native Community on *E. coli* Survival in the Beach Environment. *Applied and Environmental Microbiology*, *88*(21), e01043-22. https://doi.org/10.1128/aem.01043-22

Michaelis, C., & Grohmann, E. (2023). Horizontal Gene Transfer of Antibiotic Resistance Genes in Biofilms. *Antibiotics 2023*, *12*(2), 328. https://doi.org/10.3390/ANTIBIOTICS12020328

Milne-Davies, B., Wimmi, S., & Diepold, A. (2021). Adaptivity and dynamics in type III secretion systems. In *Molecular Microbiology* 115(3), 395-411. https://doi.org/10.1111/mmi.14658

Mitrović, B., Lezerovich, S., & Sal-Man, N. (2021). The Role of the Membrane-Associated Domain of the Export Apparatus Protein, EscV (SctV), in the Activity of the Type III Secretion System. *Frontiers in Microbiology*, *12*, 719469. https://doi.org/10.3389/fmicb.2021.719469

Mohamed, M. Y. I., & Habib, I. (2023). Pathogenic *E. coli* in the Food Chain across the Arab Countries: A Descriptive Review. *Foods 2023*, , *12*(20), 3726. https://doi.org/10.3390/FOODS12203726

Moi, I. M., Ibrahim, Z., Abubakar, B. M., Katagum, Y. M., Abdullahi, A., Yiga, G. A., Abdullahi, B., Mustapha, I., Ali, J., Mahmud, Z., Maimusa, H., Katagum, H. O., Malami, A. M., Mustapha, A., Ayuba, I., Moi, I. M., Ibrahim, Z., Abubakar, B. M., Katagum, Y. M., … Ayuba, I. (2022). *Properties of foodborne pathogens and their diseases.* IntechOpen. https://doi.org/10.5772/INTECHOPEN.105694

Mokgophi, T. M., Gcebe, N., Fasina, F., & Adesiyun, A. A. (2021). Antimicrobial Resistance Profiles of Salmonella Isolates on Chickens Processed and Retailed at Outlets of the Informal Market in Gauteng Province, South Africa. *Pathogens 2021*, *10*(3), 273. https://doi.org/10.3390/PATHOGENS10030273

Mone, N. S., Syed, S., Ravichandiran, P., Satpute, S. K., Kim, A. R., & Yoo, D. J. (2023). How Structure–Function Relationships of 1,4-Naphthoquinones Combat Antimicrobial Resistance in Multidrug-Resistant (MDR) Pathogens. *ChemMedChem*, *18*(2), e202200471. https://doi.org/10.1002/CMDC.202200471;REQUESTEDJOURNAL:JOURNAL:18607187;PAGE:STRING:ARTICLE/CHAPTER

Montso, P. K., Bezuidenhout, C. C., Mienie, C., Somorin, Y. M., Odeyemi, O. A., Mlambo, V., & Ateba, C. N. (2022). Genetic diversity and whole genome sequence analysis data of multidrug resistant atypical enteropathogenic *Escherichia coli* O177 strains: An assessment of food safety and public health implications. *International Journal of Food Microbiology*, *365*, 109555. https://doi.org/10.1016/j.ijfoodmicro.2022.109555

Muhummed, A., Alemu, A., Hosch, S., Osman, Y., Tschopp, R., Yersin, S., Schindler, T., Hattendorf, J., Zinsstag, J., Cissé, G., & Vonaesch, P. (2024). Fecal carriage of ESBL-producing *E. coli* and genetic characterization in rural children and livestock in the Somali region, Ethiopia: a one health approach. *Antimicrobial Resistance and Infection Control*, *13*(1), 1–14. https://doi.org/10.1186/S13756-024-01502-5/FIGURES/4

Müller, A., Stephan, R., & Nüesch-Inderbinen, M. (2016). Distribution of virulence factors in ESBL-producing *Escherichia coli* isolated from the environment, livestock, food and humans. *Science of The Total Environment*, *541*, 667–672. https://doi.org/10.1016/j.scitotenv.2015.09.135

Muloi, D.M., Wee, B.A., McClean, D.M., Ward, M.J., Pankhurst, L., Phan, H., Ivens, A.C., Kivali, V., Kiyong’a, A., Ndinda, C. and Gitahi, N., (2022). Population genomics of *Escherichia coli* in livestock-keeping households across a rapidly developing urban landscape. *Nature Microbiology*, *7*(4), 581-589.

Muloi, D., Ward, M. J., Pedersen, A. B., Fèvre, E. M., Woolhouse, M. E. J., & Van Bunnik, B. A. D. (2018). Are Food Animals Responsible for Transfer of Antimicrobial-Resistant *Escherichia coli* or Their Resistance Determinants to Human Populations? A Systematic

Review. *Foodborne Pathogens and Disease*, *15*(8), 467-474. https://doi.org/10.1089/fpd.2017.2411

Murphy, K. N., & Brinkworth, A. J. (2021). Manipulation of focal adhesion signaling by pathogenic microbes. *International Journal of Molecular Sciences*, *22*(3), 1–48. https://doi.org/10.3390/ijms22031358

Murugan, K., Babu, K., Sundaresan, R., Rajan, R., & Sashital, D. G. (2017). The Revolution Continues: Newly Discovered Systems Expand the CRISPR-Cas Toolkit. *Molecular Cell*, *68*(1), 15–25. https://doi.org/10.1016/J.MOLCEL.2017.09.007/ASSET/E15399B1-4B16-413E-818A-FB74CA4F67ED/MAIN.ASSETS/GR5.JPG

Mwakyoma, A. A., Kidenya, B. R., Minja, C. A., Mushi, M. F., Sandeman, A., Sabiti, W., Holden, M. T. G., & Mshana, S. E. (2023). Allele distribution and phenotypic resistance to ciprofloxacin and gentamicin among extended-spectrum β-lactamase-producing *Escherichia coli* isolated from the urine, stool, animals, and environments of patients with presumptive urinary tract infection in Tanzania. *Frontiers in Antibiotics*, *2*, 1164016. https://doi.org/10.3389/frabi.2023.1164016

Mwape, K., Bosomprah, S., Chibesa, K., Silwamba, S., Luchen, C.C., Sukwa, N., Mubanga, C., Phiri, B., Chibuye, M., Liswaniso, F., Somwe, P. (2023). Prevalence of diarrhoeagenic *Escherichia coli* among children aged between 0–36 months in peri-urban areas of Lusaka. *Microorganisms*, 11(11), p.2790. https://doi.org/10.3390/microorganisms11112790

Nadalian, B., Yadegar, A., Houri, H., Olfatifar, M., Shahrokh, S., Asadzadeh Aghdaei, H., Suzuki, H., & Zali, M. R. (2021). Prevalence of the pathobiont adherent-invasive *Escherichia coli* and inflammatory bowel disease: a systematic review and meta-analysis. *Journal of Gastroenterology and Hepatology*, *36*(4), 852–863. https://doi.org/10.1111/JGH.15260

Naidoo, N., & Zishiri, O. T. (2025). Presence, Pathogenicity, Antibiotic Resistance, and Virulence Factors of *Escherichia coli*: A Review. *Bacteria*, *4*(1), 16. https://doi.org/10.3390/BACTERIA4010016

Nasrollahian, S., Graham, J. P., & Halaji, M. (2024). A review of the mechanisms that confer antibiotic resistance in pathotypes of *E. coli*. *Frontiers in Cellular and Infection Microbiology*, *14*, 1387497. https://doi.org/10.3389/fcimb.2024.1387497

Nastasijevic, I., Proscia, F., Jurica, K., & Veskovic-Moracanin, S. (2023). Tracking Antimicrobial Resistance Along the Meat Chain: One Health Context. *Food Reviews International*, *40*(9), 2775–2809. https://doi.org/10.1080/87559129.2023.2279590;SUBPAGE:STRING:FULL

Navarro-García, F., Serapio-Palacios, A., González-Pedrajo, B., Larzábal, M., Molina, N., & Vidal, R. (2023). New Molecular Mechanisms of Virulence and Pathogenesis in *E. coli*. *Trending Topics in Escherichia coli Research: The Latin American Perspective*, 79–106. https://doi.org/10.1007/978-3-031-29882-0_4

Nazarian, S., Mousavi Gargari, S. L., Rasooli, I., Amani, J., Bagheri, S., & Alerasool, M. (2012). An in silico chimeric multi subunit vaccine targeting virulence factors of enterotoxigenic *Escherichia coli* (ETEC) with its bacterial inbuilt adjuvant. *Journal of Microbiological Methods*, *90*(1), 36–45. https://doi.org/10.1016/J.MIMET.2012.04.001

Nelson, D. W., Moore, J. E., & Rao, J. R. (2019). Antimicrobial resistance (AMR): Significance to food quality and safety. In *Food Quality and Safety* 3(1). https://doi.org/10.1093/fqsafe/fyz003

Ng'etichi, A. K. S., Voyi, K., Kirinyet, R. C., & Mutero, C. M. (2021). A systematic review on improving implementation of the revitalised integrated disease surveillance and response system in the African region: A health workers' perspective. *PLOS ONE*, *16*(3), e0248998. https://doi.org/10.1371/JOURNAL.PONE.0248998

Nielsen, K. L., Stegger, M., Kiil, K., Lilje, B., Ejrnæs, K., Leihof, R. F., Skjøt-Rasmussen, L., Godfrey, P., Monsen, T., Ferry, S., Hammerum, A. M., & Frimodt-Møller, N. (2021). *Escherichia coli* causing recurrent urinary tract infections: Comparison to non-recurrent isolates and genomic adaptation in recurrent infections. *Microorganisms*, *9*(7), 1416. https://doi.org/10.3390/MICROORGANISMS9071416/S1

Ochieng, J.B., Powell, H., Sugerman, C.E., Omore, R., Ogwel, B., Juma, J., Awuor, A.O., Sow, S.O., Sanogo, D., Onwu-chekwa, U., Keita, A.M. (2023). Epidemiology of Enteroaggregative, Enteropathogenic, and Shiga Toxin–Producing *Escherichia coli* among children aged< 5 years in 3 countries in Africa, 2015–2018: vaccine impact on Diarrhea in Africa (VIDA) Study. *Clinical Infectious Diseases*, 76(Supplement_1), pp.S77-S86. https://doi.org/10.1093/cid/ciad035

Ogino, T., Ohno, R., Sekiya, K., Kuwae, A., Matsuzawa, T., Nonaka, T., Fukuda, H., Imajoh-Ohmi, S., & Abe, A. (2006). Assembly of the type III secretion apparatus of enteropathogenic *Escherichia coli*. *Journal of Bacteriology*, *188*(8), 2801–2811. https://doi.org/10.1128/JB.188.8.2801-2811.2006/ASSET/E318CEC2-09CC-4D76-B551-BE0960312833/ASSETS/GRAPHIC/ZJB0080656440006.JPEG

Ogundare, S. T., Fasina, F. O., Makumbi, J. P., van der Zel, G. A., Geertsma, P. F., Kock, M. M., Smith, A. M., & Ehlers, M. M. (2024). Epidemiology and antimicrobial resistance profiles of pathogenic *Escherichia coli* from commercial swine and poultry abattoirs and farms in South Africa: A One Health approach. *Science of The Total Environment*, *951*, 175705. https://doi.org/10.1016/j.scitotenv.2024.175705

Okeke, I. N., Wallace-Gadsden, F., Simons, H. R., Matthews, N., Labar, A. S., Hwang, J., & Wain, J. (2010). Multi-Locus Sequence Typing of Enteroaggregative *Escherichia coli* Isolates from Nigerian Children Uncovers Multiple Lineages. *PLOS ONE*, *5*(11), e14093. https://doi.org/10.1371/JOURNAL.PONE.0014093

Oladokun, S., & Sharif, S. (2024). Exploring the complexities of poultry respiratory microbiota: colonization, composition, and impact on health. *Animal Microbiome* *6*(1), 1–21. https://doi.org/10.1186/S42523-024-00308-5

Olanya, O. M., Hoshide, A. K., Ijabadeniyi, O. A., Ukuku, D. O., Mukhopadhyay, S., Niemira, B. A., & Ayeni, O. (2019). Cost estimation of listeriosis (Listeria monocytogenes) occurrence in South Africa in 2017 and its food safety implications. *Food Control*, *102*. https://doi.org/10.1016/j.foodcont.2019.02.007

Oluwarinde, B. O., Ajose, D. J., Abolarinwa, T. O., Montso, P. K., Du Preez, I., Njom, H. A., & Ateba, C. N. (2023). Safety Properties of *Escherichia coli* O157:H7 Specific Bacteriophages: Recent Advances for Food Safety. *Foods*, *12*(21), 3989. https://doi.org/10.3390/FOODS12213989

Omar, K. B., & Barnard, T. G. (2010). The occurrence of pathogenic *Escherichia coli* in South African wastewater treatment plants as detected by multiplex PCR. *Water SA*, *36*(2), 172-176. https://doi.org/10.4314/wsa.v36i2.183725

Omolajaiye, S. A., Afolabi, K. O., & Iweriebor, B. C. (2020). Pathotyping and Antibiotic Resistance Profiling of *Escherichia coli* Isolates from Children with Acute Diarrhea in Amatole District Municipality of Eastern Cape, South Africa. *BioMed Research International*, (1), 4250165. https://doi.org/10.1155/2020/4250165

Omotoriogun, T. C., Ayodeji, O. H., Fadare, O. S., Iwajomo, S. B., Tende, T., Chaskda, A., & Salam, L. B. (2025). Metagenomic investigations of the gut microbiota of northern and southern Nigerian Village Weaver (Ploceus cucullatus) populations revealed zoonotic pathogens of public health importance. *Biologia*, 14, 1–16. https://doi.org/10.1007/S11756-025-01958-9/METRICS

Onyeka, L. O., Adesiyun, A. A., Keddy, K. H., Madoroba, E., Manqele, A., & Thompson, P. N. (2020). Shiga toxin–producing *Escherichia coli* contamination of raw beef and beef-based ready-to-eat products at retail outlets in Pretoria, South Africa. *Journal of Food Protection*, *83*(3), 476-484. https://doi.org/10.4315/0362-028X.JFP-19-372

Pakbin, B., Allahyari, S., Amani, Z., Brück, W. M., Mahmoudi, R., & Peymani, A. (2021). Prevalence, Phylogroups and Antimicrobial Susceptibility of *Escherichia coli* Isolates from Food Products. *Antibiotics*, *10*(11), 1291. https://doi.org/10.3390/ANTIBIOTICS10111291

Pakbin, B., Brück, W. M., & Rossen, J. W. A. (2021). Virulence factors of enteric pathogenic *Escherichia coli*: A review. In *International Journal of Molecular Sciences*, 22(18), 9922. https://doi.org/10.3390/ijms22189922

Palma, E., Tilocca, B., & Roncada, P. (2020). Antimicrobial Resistance in Veterinary Medicine: An Overview. *International Journal of Molecular Sciences*,*21*(6), 1914. https://doi.org/10.3390/IJMS21061914

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., & Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Research*, *25*(7), 1043–1055. https://doi.org/10.1101/GR.186072.114

Partridge, S. R., & Hall, R. M. (2005). Evolution of transposons containing blaTEM genes [2]. *Antimicrobial Agents and Chemotherapy*, *49*(3), 1267–1268. https://doi.org/10.1128/AAC.49.3.1267-1268.2005/ASSET/71E2967B-6031-4E8C-9A56-32A40C5E26F1/ASSETS/GRAPHIC/ZAC0030547720002.JPEG

Peker, N., Garcia-Croes, S., Dijkhuizen, B., Wiersma, H. H., Van Zanten, E., Wisselink, G., Friedrich, A. W., Kooistra-Smid, M., Sinha, B., Rossen, J. W. A., & Couto, N. (2019). A comparison of three different bioinformatics analyses of the 16S-23S rRNA encoding region for bacterial identification. *Frontiers in Microbiology*, *10*(MAR), 432724. https://doi.org/10.3389/FMICB.2019.00620/BIBTEX

Pitout, J. D. D., Thomson, K. S., Hanson, N. D., Ehrhardt, A. F., Moland, E. S., & Sanders, C. C. (1998). β-lactamases responsible for resistance to expanded-spectrum cephalosporins in Klebsiella pneumoniae, Escherichia colt, and Proteus mirabills isolates recovered in South Africa. *Antimicrobial Agents and Chemotherapy*, *42*(6), 1350-1354. https://doi.org/10.1128/aac.42.6.1350

Poirel, L., Madec, J.-Y., Lupo, A., Schink, A.-K., Kieffer, N., Nordmann, P., & Schwarz, S. (2018). Antimicrobial Resistance in *Escherichia coli*. *Microbiology Spectrum*, *6*(4), 10-1128. https://doi.org/10.1128/microbiolspec.arba-0026-2017

Pokharel, P., Dhakal, S., & Dozois, C. M. (2023). The Diversity of *Escherichia coli* Pathotypes and Vaccination Strategies against This Versatile Bacterial Pathogen. *Microorganisms*, *11*(2), 344. https://doi.org/10.3390/MICROORGANISMS11020344

Portaliou, A. G., Tsolis, K. C., Loos, M. S., Balabanidou, V., Rayo, J., Tsirigotaki, A., Crepin, V. F., Frankel, G., Kalodimos, C. G., Karamanou, S., & Economou, A. (2017). Hierarchical protein targeting and secretion is controlled by an affinity switch in the type III secretion system of enteropathogenic *Escherichia coli*. *The EMBO Journal*, *36*(23), 3517–3531. https://doi.org/10.15252/embj.201797515

Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A., & Korobeynikov, A. (2020). Using SPAdes De Novo Assembler. *Current Protocols in Bioinformatics*, *70*(1), e102. https://doi.org/10.1002/cpbi.102

Qian, X., Gu, J., Sun, W., Wang, X. J., Su, J. Q., & Stedfeld, R. (2018). Diversity, abundance, and persistence of antibiotic resistance genes in various types of animal manure following industrial composting. *Journal of Hazardous Materials*, *344*, 716-722. https://doi.org/10.1016/j.jhazmat.2017.11.020

Rabiu, A. G., Falodun, O. I., Fagade, O. E., Dada, R. A., & Okeke, I. N. (2022). Potentially pathogenic *Escherichia coli* from household water in peri-urban Ibadan, Nigeria. *Journal of Water and Health*, *20*(7), 1137–1149. https://doi.org/10.2166/WH.2022.117

Ramadan, A. A. (2022). Bacterial typing methods from past to present: A comprehensive overview. *Gene Reports*, *29*, 101675. https://doi.org/10.1016/j.genrep.2022.101675

Ramadan, H., Jackson, C. R., Frye, J. G., Hiott, L. M., Samir, M., Awad, A., & Woodley, T. A. (2020). Antimicrobial resistance, genetic diversity and multilocus sequence typing of *Escherichia coli* from humans, retail chicken and ground beef in Egypt. *Pathogens*, *9*(5), 357. https://doi.org/10.3390/pathogens9050357

Ramatla, T., Tawana, M., Onyiche, T. E., Lekota, K. E., & Thekisoe, O. (2021). Prevalence of antibiotic resistance in salmonella serotypes concurrently isolated from the environment, animals, and humans in South Africa: A systematic review and meta-analysis. *Antibiotics*, *10*(12), 1435. https://doi.org/10.3390/antibiotics10121435

Ramos, S., Silva, V., de Lurdes Enes Dapkevicius, M., Caniça, M., Tejedor-Junco, M. T., 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*, *10*(12), 2239. https://doi.org/10.3390/ANI10122239

Ramsamy, Y., Mlisana, K. P., Allam, M., Amoako, D. G., Abia, A. L. K., Ismail, A., Singh, R., Kisten, T., Han, K. S., Muckart, D. J. J., Hardcastle, T., Suleman, M., & Essack, S. Y. (2020). Genomic Analysis of Carbapenemase-Producing Extensively Drug-Resistant Klebsiella pneumoniae Isolates Reveals the Horizontal Spread of p18-43_01 Plasmid Encoding blaNDM-1 in South Africa. *Microorganisms* *8*(1), 137. https://doi.org/10.3390/MICROORGANISMS8010137

Ramtahal, M. A., Somboro, A. M., Amoako, D. G., Abia, A. L. K., Perrett, K., Bester, L. A., & Essack, S. Y. (2022). Molecular Epidemiology of Salmonella enterica in Poultry in South Africa Using the Farm-to-Fork Approach. *International Journal of Microbiology*, 1, 5121273 . https://doi.org/10.1155/2022/5121273

Rapisarda, C., & Fronzes, R. (2018). Secretion Systems Used by Bacteria to Subvert Host Functions. In *Current issues in molecular biology*, 25, 1-42. https://doi.org/10.21775/cimb.025.001

Ribeiro, J., Silva, V., Monteiro, A., Vieira-Pinto, M., Igrejas, G., Reis, F. S., Barros, L., & Poeta, P. (2023). Antibiotic Resistance among Gastrointestinal Bacteria in Broilers: A Review Focused on *Enterococcus spp*. and *Escherichia coli*. *Animals*, *13*(8), 1362. https://doi.org/10.3390/ANI13081362

Richter, L., du Plessis, E. M., Duvenage, S., Allam, M., Ismail, A., & Korsten, L. (2021). Whole Genome Sequencing of Extended-Spectrum- and AmpC- β-Lactamase-Positive Enterobacterales Isolated From Spinach Production in Gauteng Province, South Africa. Frontiers in Microbiology, 12(2), 734649. https://doi.org/10.3389/FMICB.2021.734649/XML/NLM

Richter, L., Duvenage, S., du Plessis, E. M., Msimango, T., Dlangalala, M., Mathavha, M. T., Molelekoa, T., Kgoale, D. M., & Korsten, L. (2024). Genomic Evaluation of Multidrug-Resistant Extended-Spectrum β-Lactamase (ESBL)-Producing *Escherichia coli* from Irrigation Water and Fresh Produce in South Africa: A Cross-Sectional Analysis. *Environmental Science and Technology*, *58*(32), 14421–14438. https://doi.org/10.1021/ACS.EST.4C02431/ASSET/IMAGES/LARGE/ES4C02431_0003.JPEG

Riley, L. W. (2020). Distinguishing Pathovars from Nonpathovars: *Escherichia coli*. *Microbiology Spectrum*, *8*(4), 10-1128. https://doi.org/10.1128/microbiolspec.ame-0014-2020

Robert, E., Grippa, M., Nikiema, D. E., Kergoat, L., Koudougou, H., Auda, Y., & Rochelle-Newall, E. (2021). Environmental determinants of *E. coli*, link with the diarrheal diseases, and indication of vulnerability criteria in tropical area (Kapore, Burkina Faso). *MedRxiv*, 2021.04.21.21255867. https://doi.org/10.1101/2021.04.21.21255867

Robins-Browne, R. M., & Hartland, E. L. (2002). *Escherichia coli* as a cause of diarrhea. *Journal of Gastroenterology and Hepatology (Australia)*, *17*(4), 467-475. https://doi.org/10.1046/j.1440-1746.2002.02769.x

Rogers, A. P., Mileto, S. J., & Lyras, D. (2022). Impact of enteric bacterial infections at and beyond the epithelial barrier. *Nature Reviews Microbiology, 21*(4), 260–274. https://doi.org/10.1038/s41579-022-00794-x

Rojas-Lopez, M., Monterio, R., Pizza, M., Desvaux, M., & Rosini, R. (2018). Intestinal pathogenic *Escherichia coli*: Insights for vaccine development. *Frontiers in Microbiology*, *9*, 333650. https://doi.org/10.3389/fmicb.2018.00440

Roth, N., Käsbohrer, A., Mayrhofer, S., Zitz, U., Hofacre, C., & Domig, K. J. (2019). The application of antibiotics in broiler production and the resulting antibiotic resistance in *Escherichia coli*: A global overview. In *Poultry Science* 98(4), 1791-1804. https://doi.org/10.3382/ps/pey539

Royer, G., Darty, M. M., Clermont, O., Condamine, B., Laouenan, C., Decousser, J. W., Vallenet, D., Lefort, A., de Lastours, V., Denamur, E., Wolff, M., Alavoine, L., Duval, X., Skurnik, D., Woerther, P. L., Andremont, A., Carbonnelle, E., Lortholary, O., Nassif, X., … Basse, B. (2021). Phylogroup stability contrasts with high within sequence type complex dynamics of *Escherichia coli* bloodstream infection isolates over a 12-year period. *Genome Medicine*, *13*(1), 1–17. https://doi.org/10.1186/S13073-021-00892-0/METRICS

Ruano-Gallego, D., Sanchez-Garrido, J., Kozik, Z., Núñez-Berrueco, E., Cepeda-Molero, M., Mullineaux-Sanders, C., Clark, J. N. B., Slater, S. L., Wagner, N., Glegola-Madejska, I., Roumeliotis, T. I., Pupko, T., Fernández, L. Á., Rodríguez-Patón, A., Choudhary, J. S., & Frankel, G. (2021). Type III secretion system effectors form robust and flexible intracellular virulence networks. *Science*, *371*(6534), eabc9531. https://doi.org/10.1126/science.abc9531

Ruiz, N., & Silhavy, T. J. (2022). How *Escherichia coli* Became the Flagship Bacterium of Molecular Biology. *Journal of Bacteriology*, *204*(9), e00230-22. https://doi.org/10.1128/jb.00230-22

Ruwandeepika, H. A. D., Madushanka, D. N., & Jayaweera, T. S. P. (2023). Foodborne Organisms: General Characteristics, Virulence Factors, and Clinical Manifestations. *Food Microbial and Molecular Biology: From Fundamentals to Applications*, 1, 213–285. https://doi.org/10.1201/9781003300601-10/

Sabença, C., Romero-Rivera, M., Barbero-Herranz, R., Sargo, R., Sousa, L., Silva, F., Lopes, F., Abrantes, A. C., Vieira-Pinto, M., Torres, C., Igrejas, G., del Campo, R., & Poeta, P. (2024). Molecular Characterization of Multidrug-Resistant *Escherichia coli* from Fecal Samples of Wild Animals. *Veterinary Sciences*, *11*(10), 469. https://doi.org/10.3390/VETSCI11100469/S1

Saeed, A., Abd, H., Sandstrom, G. (2015). Microbial aetiology of acute diarrhoea in children under five years of age in Khartoum, Sudan. *Journal of medical microbiology*, 64(4), pp.432-437. https://doi.org/10.1099/jmm.0.000043

Saeed, U., Insaf, R. A., Piracha, Z. Z., Tariq, M. N., Sohail, A., Abbasi, U. A., Fida Rana, M. S., Gilani, S. S., Noor, S., Noor, E., Waheed, Y., Wahid, M., Najmi, M. H., & Fazal, I. (2023). Crisis averted: a world united against the menace of multiple drug-resistant superbugs -pioneering anti-AMR vaccines, RNA interference, nanomedicine, CRISPR-based antimicrobials, bacteriophage therapies, and clinical artificial intelligence strategies to safeguard global antimicrobial arsenal. *Frontiers in Microbiology*, *14*, 1270018. https://doi.org/10.3389/fmicb.2023.1270018

Sanchez-Garrido, J., Ruano-Gallego, D., Choudhary, J. S., & Frankel, G. (2022). The type III secretion system effector network hypothesis. In *Trends in Microbiology*, 30(6), 524-533. https://doi.org/10.1016/j.tim.2021.10.007

Sands, K., Carvalho, M. J., Portal, E., Thomson, K., Dyer, C., Akpulu, C., Andrews, R., Ferreira, A., Gillespie, D., Hender, T., Hood, K., Mathias, J., Milton, R., Nieto, M., Taiyari, K., Chan, G. J., Bekele, D., Solomon, S., Basu, S., … Walsh, T. R. (2021). Characterization of antimicrobial-resistant Gram-negative bacteria that cause neonatal sepsis in seven low- and middle-income countries. *Nature Microbiology*, *6*(4), 512–523. https://doi.org/10.1038/S41564-021-00870

Sanger, F., & Coulson, A. R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology*, *94*(3), 441-446. https://doi.org/10.1016/0022-2836(75)90213-2

Santos, A. S., & Finlay, B. B. (2015). Bringing down the host: enteropathogenic and enterohaemorrhagic *Escherichia coli* effector-mediated subversion of host innate immune pathways. *Cellular Microbiology*, *17*(3), 318–332. https://doi.org/10.1111/CMI.12412

Santos-Neto, J. F., Santos, A. C. M., Nascimento, J. A. S., Trovão, L. O., Santos, F. F., Valiatti, T. B., Gales, A. C., Marques, A. L. V. R., Pinaffi, I. C., Vieira, M. A. M., Silva, R. M., Falsetti, I. N., & Gomes, T. A. T. (2022). Virulence Profile, Antibiotic Resistance, and Phylogenetic Relationships among *Escherichia coli* Strains Isolated from the Feces and Urine of Hospitalized Patients. *Pathogens*, *11*(12), 1528. https://doi.org/10.3390/PATHOGENS11121528/S1

Schaufler, K., Semmler, T., Wieler, L. H., Trott, D. J., Pitout, J., Peirano, G., Bonnedahl, J., Dolejska, M., Literak, I., Fuchs, S., Ahmed, N., Grobbel, M., Torres, C., McNally, A., Pickard, D., Ewers, C., Croucher, N. J., Corander, J., & Guenther, S. (2019). Genomic and functional analysis of emerging virulent and multidrug-resistant *Escherichia coli* lineage sequence type 648. *Antimicrobial Agents and Chemotherapy*, *63*(6), 10-1128. https://doi.org/10.1128/aac.00243-19

Schultz, J., Milpetz, F., Bork, P., & Ponting, C. P. (1998). SMART, a simple modular architecture research tool: Identification of signaling domains. *Proceedings of the National Academy of Sciences*, *95*(11), 5857–5864. https://doi.org/10.1073/PNAS.95.11.5857

Seemann, T. (2014). Prokka: Rapid prokaryotic genome annotation. *Bioinformatics*, *30*(14), 2068–2069. https://doi.org/10.1093/bioinformatics/btu153

Sekyere, J. O. (2016). Current state of resistance to antibiotics of last-resort in South Africa: A review from a public health perspective. In *Frontiers in Public Health*, 4, 209. https://doi.org/10.3389/fpubh.2016.00209

Serapio-Palacios, A., & Finlay, B. B. (2020). Dynamics of expression, secretion and translocation of type III effectors during enteropathogenic *Escherichia coli* infection. *Current Opinion in Microbiology*, *54*, 67–76. https://doi.org/10.1016/J.MIB.2019.12.001

Sewunet, T., Asrat, D., Woldeamanuel, Y., Ny, S., Westerlund, F., Aseffa, A., & Giske, C. G. (2022). Polyclonal spread of blaCTX-M-15 through high-risk clones of *Escherichia coli* at a tertiary hospital in Ethiopia. *Journal of Global Antimicrobial Resistance*, *29*, 405–412. https://doi.org/10.1016/j.jgar.2021.09.017

Shah, M. S., Eppinger, M., Ahmed, S., Shah, A. A., Hameed, A., & Hasan, F. (2015). Multidrug-resistant diarrheagenic *E. coli* pathotypes are associated with ready-to-eat salad and vegetables in Pakistan. *Journal of the Korean Society for Applied Biological Chemistry*, *58*(2), 267–273. https://doi.org/10.1007/S13765-015-0019-9/TABLES/4

Shaib, H., Aoun, P., Ghaddar, A., Al Labadi, H., & Obeid, Y. (2023). Multidrug Resistance and Plasmid Profiles of *Escherichia coli* Isolated from Lebanese Broiler Farms. *International Journal of Microbiology*, *2023*(1), 8811675. https://doi.org/10.1155/2023/8811675

Shawa, M., Furuta, Y., Paudel, A., Kabunda, O., Mulenga, E., Mubanga, M., Kamboyi, H., Zorigt, T., Chambaro, H., Simbotwe, M., Hang'ombe, B., & Higashi, H. (2022). Clonal relationship between multidrug-resistant *Escherichia coli* ST69 from poultry and humans in Lusaka, Zambia. *FEMS Microbiology Letters*, *368*(21–24), 1–11. https://doi.org/10.1093/FEMSLE/FNAC004

Shen, J., Zhi, S., Guo, D., Jiang, Y., Xu, X., Zhao, L., & Lv, J. (2022). Prevalence, Antimicrobial Resistance, and Whole Genome Sequencing Analysis of Shiga Toxin-Producing *Escherichia coli* (STEC) and Enteropathogenic *Escherichia coli* (EPEC) from Imported Foods in China during 2015–2021. *Toxins*, *14*(2), 68. https://doi.org/10.3390/TOXINS14020068

Silva, A., Silva, V., Dapkevicius, M. de L. E., Azevedo, M., Cordeiro, R., Pereira, J. E., Valentão, P., Falco, V., Igrejas, G., Caniça, M., & Poeta, P. (2024). Unveiling Antibiotic Resistance, Clonal Diversity, and Biofilm Formation in *E. coli* Isolated from Healthy Swine in Portugal. *Pathogens*, *13*(4), 305. https://doi.org/10.3390/PATHOGENS13040305

Singh, P., Metgud, S. C., Roy, S., & Purwar, S. (2019). Evolution of diarrheagenic *Escherichia coli* pathotypes in India. *Journal of Laboratory Physicians*, *11*(04), 346–351. https://doi.org/10.4103/JLP.JLP_58_19

Sirisaengtaksin, N., O'Donoghue, E. J., Jabbari, S., Roe, A. J., & Krachler, A. M. (2023). Bacterial outer membrane vesicles provide an alternative pathway for trafficking of *Escherichia coli* O157 type III secreted effectors to epithelial cells. *MSphere*, *8*(6), e00520-23. https://doi.org/10.1128/msphere.00520-23

Slater, S. L., Sågfors, A. M., Pollard, D. J., Ruano-Gallego, D., & Frankel, G. (2018). The Type III Secretion System of Pathogenic *Escherichia coli*. *Current Topics in Microbiology and Immunology*, *416*, 51–72. https://doi.org/10.1007/82_2018_116

Slatko, B. E., Gardner, A. F., & Ausubel, F. M. (2018). Overview of Next-Generation Sequencing Technologies. *Current Protocols in Molecular Biology*, *122*(1), e59. https://doi.org/10.1002/cpmb.59

Smith, A. M., Ramudzulu, M., Munk, P., Avot, B. J. P., Esterhuyse, K. C. M., van Blerk, N., Kwenda, S., & Sekwadi, P. (2024). Metagenomics analysis of sewage for surveillance of antimicrobial resistance in South Africa. *PLOS ONE*, *19*(8), e0309409. https://doi.org/10.1371/JOURNAL.PONE.0309409

Soufi, L., Sáenz, Y., Vinué, L., Abbassi, M. S., Ruiz, E., Zarazaga, M., Ben Hassen, A., Hammami, S., & Torres, C. (2011). *Escherichia coli* of poultry food origin as reservoir of sulphonamide resistance genes and integrons. *International Journal of Food Microbiology*, *144*(3), 497–502. https://doi.org/10.1016/J.IJFOODMICRO.2010.11.008

Stephens, C., Arismendi, T., Wright, M., Hartman, A., Gonzalez, A., Gill, M., Pandori, M., & Hess, D. (2020). F Plasmids Are the Major Carriers of Antibiotic Resistance Genes in Human-Associated Commensal *Escherichia coli*. *MSphere*, *5*(4), 10-1128. https://doi.org/10.1128/msphere.00709-20

Stoesser, N., Sheppard, A. E., Pankhurst, L., de Maio, N., Moore, C. E., Sebra, R., Turner, P., Anson, L. W., Kasarskis, A., Batty, E. M., Kos, V., Wilson, D. J., Phetsouvanh, R., Wyllie, D., Sokurenko, E., Manges, A. R., Johnson, T. J., Price, L. B., Peto, T. E. A., … Crook, D. W. (2016). Evolutionary history of the global emergence of the *Escherichia coli* epidemic clone ST131. *MBio*, *7*(2), 10-1128. https://doi.org/10.1128/mbio.02162-15

Stratakos, A. C., Ijaz, U. Z., Ward, P., Linton, M., Kelly, C., Pinkerton, L., Scates, P., McBride, J., Pet, I., Criste, A., Stef, D., Couto, J. M., Sloan, W. T., Dorrell, N., Wren, B. W., Stef, L., Gundogdu, O., & Corcionivoschi, N. (2020). In vitro and in vivo characterisation of Listeria monocytogenes outbreak isolates. *Food Control*, *107*, 106784. https://doi.org/10.1016/j.foodcont.2019.106784

Streicher, L. M. (2021). Exploring the future of infectious disease treatment in a post-antibiotic era: A comparative review of alternative therapeutics. *Journal of Global Antimicrobial Resistance*, *24*, 285–295. https://doi.org/10.1016/J.JGAR.2020.12.025

Swelum, A. A., Elbestawy, A. R., El-Saadony, M. T., Hussein, E. O. S., Alhotan, R., Suliman, G. M., Taha, A. E., Ba-Awadh, H., El-Tarabily, K. A., & Abd El-Hack, M. E. (2021). Ways to minimize bacterial infections, with special reference to *Escherichia coli*, to cope with the first-week mortality in chicks: an updated overview. *Poultry Science*, *100*(5), 101039. https://doi.org/10.1016/J.PSJ.2021.101039

Taati Moghadam, M., Mirzaei, M., Fazel Tehrani Moghaddam, M., Babakhani, S., Yeganeh, O., Asgharzadeh, S., Farahani, H. E., & Shahbazi, S. (2021). The Challenge of Global Emergence of Novel Colistin-Resistant *Escherichia coli* ST131. *Microbial Drug Resistance*, *27*(11), 1513-1524. https://doi.org/10.1089/mdr.2020.0505

Takawira, F. T., Pitout, J. D. D., Thilliez, G., Mashe, T., Gutierrez, A. V., Kingsley, R. A., Peirano, G., Matheu, J., Midzi, S. M., Mwamakamba, L. W., Gally, D. L., Tarupiwa, A., Mukavhi, L., Ehlers, M. M., Mtapuri-Zinyowera, S., & Kock, M. M. (2021). Molecular epidemiology of extended-spectrum beta-lactamase–producing extra-intestinal pathogenic *Escherichia coli* strains over a 2-year period (2017–2019) from Zimbabwe. *European Journal of Clinical Microbiology and Infectious Diseases*, 2021, 1–14. https://doi.org/10.1007/S10096-021-04379-Z/FIGURES/4

Takawira, F. T., Pitout, J. D. D., Thilliez, G., Mashe, T., Gutierrez, A. V., Kingsley, R. A., Peirano, G., Matheu, J., Midzi, S. M., Mwamakamba, L. W., Gally, D. L., Tarupiwa, A., Mukavhi, L., Ehlers, M. M., Mtapuri-Zinyowera, S., & Kock, M. M. (2022). Faecal carriage of ESBL producing and colistin resistant *Escherichia coli* in avian species over a

2-year period (2017-2019) in Zimbabwe. *Frontiers in Cellular and Infection Microbiology*, *12*, 1035145. https://doi.org/10.3389/fcimb.2022.1035145

Tanih, N. F., Sekwadi, E., Ndip, R. N., & Bessong, P. O. (2015). Detection of Pathogenic *Escherichia coli* and Staphylococcus aureus from Cattle and Pigs Slaughtered in Abattoirs in Vhembe District, South Africa. *The Scientific World Journal*, *2015*(1), 195972. https://doi.org/10.1155/2015/195972

Tao, S., Chen, H., Li, N., & Liang, W. (2022). The Application of the CRISPR-Cas System in Antibiotic Resistance. *Infection and Drug Resistance*, *Volume 15*, 4155–4168. https://doi.org/10.2147/IDR.S370869

Tejeda-Dominguez, F., Huerta-Cantillo, J., Chavez-Dueñas, L., & Navarro-Garcia, F. (2017). A novel mechanism for protein delivery by the type 3 secretion system for extracellularly secreted proteins. *MBio*, *8*(2), 10-1128. https://doi.org/10.1128/mbio.00184-17

Tello, M., Ocejo, M., Oporto, B., Lavín, J. L., & Hurtado, A. (2022). Within-farm dynamics of ESBL-producing *Escherichia coli* in dairy cattle: Resistance profiles and molecular characterization by long-read whole-genome sequencing. *Frontiers in Microbiology*, *13*, 936843. https://doi.org/10.3389/fmicb.2022.936843

Tetzschner, A. M. M., Johnson, J. R., Johnston, B. D., Lund, O., & Scheutz, F. (2020). In Silico genotyping of *Escherichia coli* isolates for extraintestinal virulence genes by use of whole-genome sequencing data. *Journal of Clinical Microbiology*, *58*(10), e01269. https://doi.org/10.1128/jcm.01269-20

Tonini Da Rocha, D., De Oliveira Salle, F., Borges, K. A., Furian, T. Q., Pinheiro Do Nascimento, V., Luiz De Souza Moraes, H., Tadeu, C., & Salle, P. (2021). Avian pathogenic Escherichia coli (APEC) and uropathogenic Escherichia coli (UPEC): characterization and comparison. *Journal of Infection in Developing Countries*, 15(7), 962-971. https://doi.org/10.3855/jidc.14217

Toro, M., Cao, G., Ju, W., Allard, M., Barrangou, R., Zhao, S., Brown, E., & Meng, J. (2014). Association of clustered regularly interspaced short palindromic repeat (CRISPR) elements with specific serotypes and virulence potential of Shiga toxin-producing *Escherichia coli*. *Applied and Environmental Microbiology*, *80*(4), 1411–1420. https://doi.org/10.1128/AEM.03018-13/SUPPL_FILE/ZAM999105137SD1.XLSX

Tosi, T., Pflug, A., Discola, K. F., Neves, D., & Dessen, A. (2013). Structural basis of eukaryotic cell targeting by type III secretion system (T3SS) effectors. *Research in Microbiology*, *164*(6), 605–619. https://doi.org/10.1016/j.resmic.2013.03.019

Touchon, M., Charpentier, S., Pognard, D., Picard, B., Arlet, G., Rocha, E. P. C., Denamur, E., & Branger, C. (2012). Antibiotic resistance plasmids spread among natural isolates of *Escherichia coli* in spite of CRISPR elements. *Microbiology (United Kingdom)*, *158*(12), 2997–3004. https://doi.org/10.1099/MIC.0.060814-0/CITE/REFWORKS

Touchon, M., Perrin, A., De Sousa, J. A. M., Vangchhia, B., Burn, S., O'Brien, C. L., Denamur, E., Gordon, D., & Rocha, E. P. C. (2020). Phylogenetic background and habitat drive the genetic diversification of *Escherichia coli*. *PLOS Genetics*, *16*(6), e1008866. https://doi.org/10.1371/JOURNAL.PGEN.1008866

Toval, F., Köhler, C. D., Vogel, U., Wagenlehner, F., Mellmann, A., Fruth, A., Schmidt, M. A., Karch, H., Bielaszewska, M., & Dobrindt, U. (2014). Characterization of *Escherichia coli* isolates from hospital inpatients or outpatients with urinary tract infection. *Journal of Clinical Microbiology*, *52*(2), 407–418. https://doi.org/10.1128/JCM.02069-13/ASSET/60B3CBC4-6168-4A01-9E8A-019613CD85BA/ASSETS/GRAPHIC/ZJM9990931170005.JPEG

Tucker, K., Mageiros, L., Carstens, A., Bröcker, L., Archer, E., Smith, K., Mourkas, E., Pascoe, B., Nel, D., Meric, G., Sheppard, S. K., Kasprzyk-Hordern, B., Botes, M., Feil, E. J., & Wolfaardt, G. (2022). Spatiotemporal Investigation of Antibiotic Resistance in the Urban Water Cycle Influenced by Environmental and Anthropogenic Activity. *Microbiology Spectrum*, *10*(5), e02473-22. https://doi.org/10.1128/spectrum.02473-22

Uelze, L., Grützke, J., Borowiak, M., Hammerl, J. A., Juraschek, K., Deneke, C., Tausch, S. H., & Malorny, B. (2020). Typing methods based on whole genome sequencing data. *One Health Outlook 2020 2:1*, *2*(1), 1–19. https://doi.org/10.1186/S42522-020-0010-1

Urban-Chmiel, R., Marek, A., Stępień-Pyśniak, D., Wieczorek, K., Dec, M., Nowaczek, A., & Osek, J. (2022). Antibiotic Resistance in Bacteria—A Review. *Antibiotics*, *11*(8), 1079. https://doi.org/10.3390/ANTIBIOTICS11081079/S1

Van Boeckel, T. P., Pires, J., Silvester, R., Zhao, C., Song, J., Criscuolo, N. G., Gilbert, M., Bonhoeffer, S., & Laxminarayan, R. (2019). Global trends in antimicrobial resistance in animals in low- And middle-income countries. *Science*, *365*(6459), eaaw1944. https://doi.org/10.1126/science.aaw1944

Van Duijkeren, E., Schwarz, C., Bouchard, D., Catry, B., Pomba, C., Baptiste, K. E., Moreno, M. A., Rantala, M., Ružauskas, M., Sanders, P., Teale, C., Wester, A. L., Ignate, K., Kunsagi, Z., & Jukes, H. (2019). The use of aminoglycosides in animals within the EU: development of resistance in animals and possible impact on human and animal health: a review. *Journal of Antimicrobial Chemotherapy*, *74*(9), 2480–2496. https://doi.org/10.1093/JAC/DKZ161

Wagner, S., & Diepold, A. (2020). A Unified Nomenclature for Injectisome-Type Type III Secretion Systems. *Current Topics in Microbiology and Immunology*, *427*, 1–10. https://doi.org/10.1007/82_2020_210

Wagner, S., Grin, I., Malmsheimer, S., Singh, N., Torres-Vargas, C. E., & Westerhausen, S. (2018). Bacterial type III secretion systems: A complex device for the delivery of bacterial effector proteins into eukaryotic host cells. *FEMS Microbiology Letters*, *365*(19), fny201. https://doi.org/10.1093/femsle/fny201

Wagner, S., Königsmaier, L., Lara-Tejero, M., Lefebre, M., Marlovits, T. C., & Galán, J. E. (2010). Organization and coordinated assembly of the type III secretion export apparatus. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(41), 17745-17750. https://doi.org/10.1073/pnas.1008053107

Wang, D., Ji, X., Jiang, B., Yuan, Y., Liang, B., Sun, S., Zhu, L., Liu, J., Guo, X., Yin, Y., & Sun, Y. (2024). Prevalence of Antibiotic Resistance and Virulence Genes in *Escherichia coli* Carried by Migratory Birds on the Inner Mongolia Plateau of Northern China from 2018 to 2023. *Microorganisms*, *12*(6), 1076. https://doi.org/10.3390/MICROORGANISMS12061076/S1

Wang, J., Chitsaz, F., Derbyshire, M. K., Gonzales, N. R., Gwadz, M., Lu, S., Marchler, G. H., Song, J. S., Thanki, N., Yamashita, R. A., Yang, M., Zhang, D., Zheng, C., Lanczycki, C. J., & Marchler-Bauer, A. (2023). The conserved domain database in 2023. *Nucleic Acids Research*, *51*(D1), D384–D388. https://doi.org/10.1093/NAR/GKAC1096

Wang, L., Wu, Y., Cai, P., & Huang, Q. (2020). The attachment process and physiological properties of *Escherichia coli* O157:H7 on quartz. *BMC Microbiology*, *20*(1), 1-14. https://doi.org/10.1186/s12866-020-02043-8

Wang, W., Yu, L., Hao, W., Zhang, F., Jiang, M., Zhao, S., & Wang, F. (2021). Multi-Locus Sequence Typing and Drug Resistance Analysis of Swine Origin *Escherichia coli* in Shandong of China and Its Potential Risk on Public Health. *Frontiers in Public Health*, *9*, 780700. https://doi.org/10.3389/fpubh.2021.780700

Wareth, G., Linde, J., Hammer, P., Pletz, M. W., Neubauer, H., & Sprague, L. D. (2022). WGS-Based Phenotyping and Molecular Characterization of the Resistome, Virulome and Plasmid Replicons in Klebsiella pneumoniae Isolates from Powdered Milk Produced in Germany. *Microorganisms*, *10*(3), 564. https://doi.org/10.3390/MICROORGANISMS10030564/S1

Webale, M.K., Guyah, B., Wanjala, C., Nyanga, P.L., Webale, S.K., Abonyo, C., Kitungulu, N., Kiboi, N., Bowen, N. (2020). Phenotypic and genotypic antibiotic resistant diarrheagenic *Escherichia coli* pathotypes isolated from children with diarrhea in Nairobi City, Kenya. *Ethiopian Journal of Health Sciences*, 30(6). https://doi.org/10.4314/ejhs.v30i6.5

Wee, B. A., Muloi, D. M., & van Bunnik, B. A. D. (2020). Quantifying the transmission of antimicrobial resistance at the human and livestock interface with genomics. In *Clinical Microbiology and Infection*, 26(12), 1612-1616. https://doi.org/10.1016/j.cmi.2020.09.019

Wei, B., Cha, S. Y., Shang, K., Zhang, J. F., Jang, H. K., & Kang, M. (2021). Genetic diversity of extended-spectrum cephalosporin resistance in Salmonella enterica and *E. coli* isolates in a single broiler chicken. *Veterinary Microbiology*, *254*, *109010*. https://doi.org/10.1016/j.vetmic.2021.109010

Wesonga, S., Muluvi, G., Okemo, P., & Kariuki, S. (2010). Antibiotic resistant Salmonella and *Escherichia coli* isolated from indiginous Gallus domesticus in Nairobi, Kenya. *East African Medical Journal*, *87*(5), 205–210. https://doi.org/10.4314/EAMJ.V87I5.63075

Whelan, S., Lucey, B., & Finn, K. (2023). Uropathogenic *Escherichia coli* (UPEC)-Associated Urinary Tract Infections: The Molecular Basis for Challenges to Effective Treatment. *Microorganisms 2023, Vol. 11, Page 2169*, *11*(9), 2169. https://doi.org/10.3390/MICROORGANISMS11092169

World Health Organization. (2024). WHO Foodborne Disease Burden Epidemiology Reference Group for 2021-2024: second meeting report, 19 October-2 November 2021. World Health Organization.

Wolde, D., Eguale, T., Medhin, G., Haile, A. F., Alemayehu, H., Mihret, A., Pirs, M., Strašek Smrdel, K., Avberšek, J., Kušar, D., Cerar Kišek, T., Janko, T., Steyer, A., & Starčič Erjavec, M. (2024). Genomic Characterization of Extended-Spectrum β-Lactamase-

Producing and Third-Generation Cephalosporin-Resistant *Escherichia coli* Isolated from Stools of Primary Healthcare Patients in Ethiopia. *Antibiotics*, *13*(9), 851. https://doi.org/10.3390/ANTIBIOTICS13090851/S1

Woolhouse, M., Ward, M., Van Bunnik, B., & Farrar, J. (2015). Antimicrobial resistance in humans, livestock and the wider environment. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1670), 20140083. https://doi.org/10.1098/rstb.2014.0083

Wu, S., Feng, Y., Yang, Y., Zhu, K., Wang, S., Wang, X., Li, G. B., & Zong, Z. (2023). Characterization of a novel carbapenem-hydrolysing β-lactamase OXA-1041 in *Escherichia coli*. *Journal of Antimicrobial Chemotherapy*, *78*(5), 1288–1294. https://doi.org/10.1093/JAC/DKAD091

Wyrsch, E. R., Dolejska, M., & Djordjevic, S. P. (2022). Genomic Analysis of an I1 Plasmid Hosting a sul3-Class 1 Integron and blaSHV-12 within an Unusual *Escherichia coli* ST297 from Urban Wildlife. *Microorganisms*, *10*(7), 1387. https://doi.org/10.3390/microorganisms10071387

Xue, C., & Sashital, D. G. (2019). Mechanisms of Type I-E and I-F CRISPR-Cas Systems in Enterobacteriaceae . *EcoSal Plus*, *8*(2), 10-1128. https://doi.org/10.1128/ecosalplus.esp-0008-2018

Xu, L., Zhai, Y., Lyu, Y., Wang, Q., An, S. C., Chen, J. C., Chen, Y. S., Liu, L., Li, J. Bin, & Gao, Z. C. (2014). Identification of klebsiella pneumoniae strains harboring inactive extended-spectrum beta-lactamase antibiotic-resistance genes. *Chinese Medical Journal*, *127*(17), 3051–3057. https://doi.org/10.3760/CMA.J.ISSN.0366-6999.20140628

Xu, Y., Bai, X., Jin, Y., Hu, B., Wang, H., Sun, H., Fan, R., Fu, S., & Xiong, Y. (2017). High prevalence of virulence genes in specific genotypes of atypical enteropathogenic *Escherichia coli*. *Frontiers in Cellular and Infection Microbiology*, *7*(APR), 255910. https://doi.org/10.3389/FCIMB.2017.00109/BIBTEX

Yang, C., Rehman, M. A., Yin, X., Carrillo, C. D., Wang, Q., Yang, C., Gong, J., & Diarra, M. S. (2021). Antimicrobial Resistance Phenotypes and Genotypes of *Escherichia coli* Isolates from Broiler Chickens Fed Encapsulated Cinnamaldehyde and Citral. *Journal of Food Protection*, *84*(8), 1385–1399. https://doi.org/10.4315/JFP-21-033

Yarabbi, H., Mortazavi, S. A., Yavarmanesh, M., & Javadmanesh, A. (2020). In Silico Study of Different Signal Peptides to Express Recombinant Glutamate Decarboxylase in the Outer Membrane of *Escherichia coli*. *International Journal of Peptide Research and Therapeutics*, *26*(4), 1879–1891. https://doi.org/10.1007/S10989-019-09986-1/FIGURES/3

Yu, G., Smith, D. K., Zhu, H., Guan, Y., & Lam, T. T. Y. (2017). ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution*, *8*(1), 28–36. https://doi.org/10.1111/2041-210X.12628

Zakaria, D., Matsuda, S., Iida, T., Hayashi, T., & Arita, M. (2023). Genome Analysis Identifies a Novel Type III Secretion System (T3SS) Category in Vibrio Species. *Microorganisms*, *11*(2), 290. https://doi.org/10.3390/MICROORGANISMS11020290/S1

Zeng, M., Xia, J., Zong, Z., Shi, Y., Ni, Y., Hu, F., Chen, Y., Zhuo, C., Hu, B., Lv, X., Li, J., Liu, Z., Zhang, J., Yang, W., Yang, F., Yang, Q., Zhou, H., Li, X., Wang, J., … Wang, M. (2023). Guidelines for the diagnosis, treatment, prevention and control of infections caused by carbapenem-resistant gram-negative bacilli. *Journal of Microbiology, Immunology and Infection*, *56*(4), 653–671. https://doi.org/10.1016/J.JMII.2023.01.017

Zhang, P., Shen, Z., Zhang, C., Song, L., Wang, B., Shang, J., Yue, X., Qu, Z., Li, X., Wu, L., Zheng, Y., Aditya, A., Wang, Y., Xu, S., & Wu, C. (2017). Surveillance of antimicrobial resistance among *Escherichia coli* from chicken and swine, China, 2008–2015. *Veterinary Microbiology*, 203, 49-55. https://doi.org/10.1016/j.vetmic.2017.02.008

Zhang, Q., & Ye, Y. (2017). Not all predicted CRISPR-Cas systems are equal: Isolated cas genes and classes of CRISPR like elements. *BMC Bioinformatics*, *18*(1), 1–12. https://doi.org/10.1186/S12859-017-1512-4/TABLES/3

Zhang, X., Xia, S., Ye, Y., & Wang, H. (2021). Opportunistic pathogens exhibit distinct growth dynamics in rainwater and tap water storage systems. *Water Research*, *204*, 117581. https://doi.org/10.1016/j.watres.2021.117581

Zhao, S., White, D. G., Ge, B., Ayers, S., Friedman, S., English, L., Wagner, D., Gaines, S., & Meng, J. (2001). Identification and Characterization of Integron-Mediated Antibiotic Resistance among Shiga Toxin-Producing *Escherichia coli* Isolates. *Applied and Environmental Microbiology*, *67*(4), 1558–1564. https://doi.org/10.1128/AEM.67.4.1558-1564.2001

Zhao, X., Zhao, H., Zhou, Z., Miao, Y., Li, R., Yang, B., Cao, C., Xiao, S., Wang, X., Liu, H., Wang, J., & Yang, Z. (2022). Characterization of Extended-Spectrum β-Lactamase-Producing *Escherichia coli* Isolates That Cause Diarrhea in Sheep in Northwest China. *Microbiology Spectrum*, *10*(4), e01595-22. https://doi.org/10.1128/spectrum.01595-22

Zheng, T., Li, P., Li, L., & Zhang, Q. (2021). Research advances in and prospects of ornamental plant genomics. In *Horticulture Research* (Vol. 8, Issue 1). https://doi.org/10.1038/s41438-021-00499-x

Zhou, J., Yang, J., Hu, F., Gao, K., Sun, J., & Yang, J. (2020). Clinical and molecular epidemiologic characteristics of ceftazidime/ avibactam-resistant carbapenem-resistant *Klebsiella pneumoniae* in a neonatal intensive care unit in China. *Infection and Drug Resistance*, *13*, 2571–2578. https://doi.org/10.2147/IDR.S256922

Ziege, R., Tsirigoni, A. M., Large, B., Serra, D. O., Blank, K. G., Hengge, R., Fratzl, P., & Bidan, C. M. (2021). Adaptation of *Escherichia coli* Biofilm Growth, Morphology, and Mechanical Properties to Substrate Water Content. *ACS Biomaterials Science and Engineering*, *7*(11), 5315–5325. https://doi.org/10.1021/ACSBIOMATERIALS.1C00927/ASSET/IMAGES/MEDIUM/AB1C00927_M001.GIF

Zou, L., Meng, J., McDermott, P. F., Wang, F., Yang, Q., Cao, G., Hoffmann, M., & Zhao, S. (2014). Presence of disinfectant resistance genes in *Escherichia coli* isolated from retail meats in the USA. *Journal of Antimicrobial Chemotherapy*, *69*(10), 2644–2649. https://doi.org/10.1093/JAC/DKU197

APPENDICES

Appendix A: Chapter 3 Supplementary Material

Table S 1: *Escherichia coli* pathotype determination based on virulence factors

InPEC Pathotypes	**Virulence Factors (Target Genes)**
Enteroinvasive *E. coli* (EIEC)	*ipaH, ipaC*
Shiga Toxin-Producing *E. coli* (STEC)	*stx1, stx2, eae*
Enteroaggregative *E. coli* (EAEC)	*aatA, aaiC, aggR*
Enterotoxigenic *E. coli* (ETEC)	*elt, est* - ST gene, LT gene
Enteropathogenic *E. coli* (EPEC)	*bfpA, eae*
Diffusely adherent *E. coli* (DAEC)	*afa*
ExPEC Pathotypes	**Virulence Factors (Target Genes)**
Uropathogenic *E. coli* (UPEC)	*fimH, papA/papG, hlyA, usp, cnf1, iutA, fyuA, afa, ompT, sfa, chuA, vat, fycV*
Avian pathogenic *E. coli* (APEC)	*iroN, iss, ompT, tsh, hlyF, cvaC, iutA, sfa, papGII, fimH, nlpl*
Neonatal meningitis *E. coli* (NMEC)	*kpsMTII, ibeA, sfa/foc, iss, neuC, iutA, fyuA, iroN, ire, hlyF, astA, tsh*

Heat stable enterotoxin (ST), heat-labile enterotoxin (LT).

Table S 2: Sequence types, Serotypes and Phylogroups

		PHYLOGROUPS					
Sequence Type	**Serotype**	**No. OF ISOLATES (n=35)**	**A**	**B1**	**B2**	**C**	**E**
ST1858	O6:H5	4	0	0	4	0	0
ST975	O101:H37	4	4	0	0	0	0
ST88	O8:H27	3	1	0	0	2	0
ST10	O108:H10	3	3	0	0	0	0
ST95	O2:H4	3	0	0	3	0	0
ST2952	O64:H19	3	0	3	0	0	0
ST120	O8:H9	2	1	1	0	0	0
ST58	O155:H21	1	0	1	0	0	0
ST3478	O11:H10	1	1	0	0	0	0
ST5675	O5:H4	1	1	0	0	0	0
ST6381	O137:H41	1	0	0	0	0	1
ST79	O2:H4	1	0	0	1	0	0
ST4656	O65:H9	1	1	0	0	0	0
ST29	O26:H11	1	0	1	0	0	0
ST540	ONT:H30	1	1	0	0	0	0
ST398	ONT:H4	1	1	0	0	0	0
ST2598	O175:H16	1	0	0	0	0	0
ST93	O15:H17	1	1	0	0	0	0
ST3	O111:H2	1	0	1	0	0	0

Table S 3: Blastn data

Blastn Analysis						
No.	**Isolate Accession**	**Query Coverage (%)**	**Per. Ident (%)**	**e-value**	**Species level**	**NCBI Accession**
S1	SAMN41920844	98%	100%	0.0	*E. coli*	CP054214.1
S2	SAMN41920845	100%	100%	0.0	*E. coli*	CP049979.1
S3	SAMN41920846	100%	99.94%	0.0	*E. coli*	CP054379.1
S4	SAMN41920847	100%	99.93%	0.0	*E. coli*	CP024978.1
S5	SAMN41920848	100%	99.97%	0.0	*E. coli*	CP051222.1
S6	SAMN41920849	100%	100%	0.0	*E. coli*	AP023237.1
S7	SAMN41920850	100%	100%	0.0	*E. coli*	CP054214.1
S8	SAMN41920851	97%	100%	0.0	*E. coli*	CP141844.1
S9	SAMN41920852	100%	99.97%	0.0	*E. coli*	CP128590.1
S10	SAMN41920853	100%	100%	0.0	*E. coli*	CP053284.1
S11	SAMN41920854	95%	99.97%	0.0	*E. coli*	CP034784.1
S12	SAMN41920855	98%	100%	0.0	*E. coli*	CP097213.1
S13	SAMN41920856	100%	100%	0.0	*E. coli*	CP136025.1
S14	SAMN41920857	97%	100%	0.0	*E. coli*	CP053231.1
S15	SAMN41920858	100%	97.57%	0.0	*E. coli*	CP044177.1
S17	SAMN41920859	98%	100%	0.0	*E. coli*	CP054454.1
S18	SAMN41920860	100%	99.81%	0.0	*E. coli*	CP047461.1
S19	SAMN41920861	97%	100%	0.0	*E. coli*	CP114893.1
S20	SAMN41920862	86%	99.98%	0.0	*E. coli*	LR134315.1
S21	SAMN41920863	98%	99.97%	0.0	*E. coli*	CP135671.1
S22	SAMN41920864	100%	99.98%	0.0	*E. coli*	CP054345.1
S23	SAMN41920865	100%	100%	0.0	*E. coli*	CP110097.1
S24	SAMN41920866	97%	100%	0.0	*E. coli*	CP054214.1
S25	SAMN41920867	98%	99.98%	0.0	*E. coli*	LR134000.1
S26	SAMN41920868	100%	99.97%	0.0	*E. coli*	CP066032.1
S27	SAMN41920869	98%	100%	0.0	*E. coli*	LR134236.1
S28	SAMN41920870	100%	100%	0.0	*E. coli*	CP135680.1

S29	SAMN41920871	98%	99.19%	0.0	*E. coli*	CP057612.1
S30	SAMN41920872	98%	99.98%	0.0	*E. coli*	LR134315.1
S32	SAMN41920873	99%	99.97%	0.0	*E. coli*	CP024978.1
S33	SAMN41920874	98%	100%	0.0	*E. coli*	CP053231.1
S34	SAMN41920875	100%	100%	0.0	*E. coli*	CP051719.1
S35	SAMN41920876	100%	100%	0.0	*E. coli*	CP099270.1
S36	SAMN41920877	100%	100%	0.0	*E. coli*	CP051222.1
S37	SAMN41920878	94%	99.94%	0.0	*E. coli*	LN554925.1

Table S 4: Prevalence of Antimicrobial resistance genes in *E. coli* isolates

Antimicrobial Resistance	Percentage
sitABCD	42.9%
tet(B)	22.9%
qacE	14.3%
sul1	8.6%
sul2	5.7%
catA1	5.7%
blaTEM-1b	8.6%
blaTEM-104	2.9%
aadA1 and aadA2b	5.7%
aph(6)-id and aph(3'')-ib	5.7%

Table S 5: Overall results summary grouped according to species level

No.	Animal Id	Sample Type	Species	Area	Antibiotic Resistance	Resistance Gene	Virulence Genes	Serotype	Sequence Type	Phylo-Group	Pathotype
S2	1792	TURKEY drumstick	Poultry	NA	PEN, PIP, SXT	-	*fimH*, *papC*,	O155:H21	ST58	B1	APEC
S3	2003	Chicken drumstick	Poultry	MP	PEN,SXT, H_2O_2	*sitABCD*	*iutA*, *fyuA*, *iha*, *tratT*	O11:H10	ST3478	A	APEC
S4	2040	Chicken leg quarter	Poultry	NL	PEN, AMC, SXT, H_2O_2	*sitABCD*	-	O8:H9	ST120	A	UPEC
S5	2033	Chicken wings	Poultry	FS	PEN, SXT	-	*iutA*, *iss*,	O5:H4	ST5675	A	APEC
S6	2021	Chicken drumstick	Poultry	FS	PEN, SXT, CAZ	-	*fimH*, *iss*, *cvaC*, *traT*, *iutA*, *tsh*,	O104:H12	ST88	A	APEC
S7	2025	Chicken fillet	Poultry	FS	PEN, SXT, ATM, H_2O_2	*sitABCD*	*fimH*, *fyuA*,	O6:H5	ST1858	B2	UPEC
S8	2026	Chicken leg quarter	Poultry	FS	PEN, CAZ, SMT, AMC, TET, SXT, CHL	*qacE*, *aadA1* *sul1*, *catA1*, *blaTEM-1B*, *tet(B)*	*fimH*, *papC*, *iutA*, *fyuA*, *traT*	O11:H19	ST88	C	APEC
S9	2023	Chicken drumstick	Poultry	FS	PEN, SXT, H_2O_2	*sitABCD*	*fimH*, *iss*, *traT*, *iutA*,	O137:H41	ST6381	E	APEC
S10	2035	Chicken wings	Poultry	FS	PEN, SXT	-	*fimH*, *iss*, *traT*, *papC*,	O101:H37	ST975	A	ETEC
S11	2022	Chicken drumstick	Poultry	FS	PEN, SXT	-	*fimH*	O108:H10	ST10	A	APEC
S17	1557	Chicken drumstick and wings	Poultry	NW	AMP, AMC, SXT, H_2O_2	*sitABCD*	*fimH*, *iss*, *papC*, *fyuA*, *ireA*,	O2:H4	ST95	B2	UPEC
S18	2041	Chicken thigh	Poultry	FS	PEN, SXT	-	*fimH*, *hlyA*, *traT*, *stx2*	O65:H9	ST4656	A	EHEC

S28	1789	Chicken leg quarter	Poultry	EU	PEN, SXT	-	*fimH, iss, traT, papC,*	O101:H37	ST975	A	ETEC
S12	2037	Pork wors	Porcine	FS	PEN, CHL, SXT, TET, AMC, SMT	*tet(B), blaTEM-1B, qacE, aadA1, sul1, catA1*	*fimH, iss, traT,*	O101:H37	ST975	A	ETEC
S14	2029	Pork shoulder	Porcine	FS	PEN, SXT, SMT, AMC, TET	*aadA4, sul1, qacE, tet(B)*	*fimH, iss, iutA, cvaC, traT, fyuA, papC,*	O2:H4	ST79	B2	UPEC
S15	2034	Pork chops	Porcine	FS	PEN, SXT	-	*fimH, iss, traT,*	O8:H27	ST88	C	ETEC
S19	2331	Animal faeces	Porcine	GP	PEN, SXT	-	-	O107:H27	ST10	A	APEC
S22	2334	Animal faeces	Porcine	GP	PEN, ATM, SXT, H_2O_2	*sitABCD*	*fimH, ibeA, fyuA,*	O6:H5	ST1858	B2	UPEC
S26	1568	Animal faeces	Porcine	GP	PEN, AMC, SXT	-	*fimH, iss, traT, eae, bfpA,*	O111:H2	ST3	B1	EPEC
S29	9001	Animal faeces	Porcine	GP	PEN, SXT	*O15:H17*	*fimH, fyuA, traT, iha,*	O123:H16	ST10	A	APEC
S32	9002	Animal faeces	Porcine	GP	PEN, TET, SMT, CHL, AMC, SXT	*tet(B), qacE, sul1, aadA2b, catA1*	*fimH, iss, traT,*	O9:H20	ST120	B1	ETEC
S34	9015	Animal faeces	Porcine	GP	PEN, SXT, AMC, H_2O_2	*sitABCD*	*fimH, iss, traT, papC, fyuA, ireA,*	O2:H4	ST95	B2	UPEC
S1	1820	Beef and Pork Wors	Bovine and Porcine	NW	PEN, ATM, SXT, H_2O_2	*sitABCD*	*fimH, ibeA*	O6:H5	ST1858	B2	UPEC
S20	1344	Beef mince	Bovine	GP	PEN, TET, SXT, CAZ, SMT	*sitABCD, aph(6)-Id, sul2,*	*fimH, fasA, traT,*	O64:H19	ST2952	B1	ETEC

						aph(3'')-Ib, *tet(B)*					
S24	540	Beef wors	Bovine	FS	PEN, AZT, SXT, H_2O_2	*sitABCD*	*fimH, fyuA, ibeA,*	O6:H5	ST1858	B2	UPEC
S36	1431	Beef biltong	Bovine	LP	PEN, SXT, H_2O_2	*sitABCD*	*cvaC, ireA, iss, fyuA, iutA,*	O15:H17	ST93	A	UPEC
S13	2028	Lamb leash	Ovine	FS	PEN, SXT	-	*fimH, iss, traT, papC,*	O101:H37	ST975	A	ETEC
S21	7461	Animal faeces	Ovine	GP	PEN, SXT, H_2O_2	*sitABCD*	*fimH, iss, iutA, papC, traT,*	O117:H4	ST10	A	UPEC
S23	2027	Animal faeces	Unknown	GP	PEN, AMC, SXT	-	*fimH, iss, fyuA, traT, eae, hlyA*	O26:H11	ST29	B1	EPEC
S25	2089	Animal faeces	Unknown	GP	PEN, TET, SXT, SMT, AMP, CHL,	*tet(B), qacE, sul1, aadA2, blaTEM-104, blaTEM-1B*	*fimH, iss, traT,*	ONT:H30	ST540	A	APEC
S27	2024	Unknown	Unknown	GP	PEN, SXT	-	*fimH, fyuA,*	ONT:H4	ST398	A	APEC
S30	2039	Animal faeces	Unknown	GP	PEN, TET, CAZ, SXT, SMT	*sitABCD, aph(6)-Id, sul2, aph(3'')-Ib, tet(B)*	*fimH, fasA, traT,*	O64:H19	ST2952	B1	ETEC
S33	2031	Animal faeces	Unknown	GP	PEN, SXT, AMC, H_2O_2	*sitABCD*	*fimH, iss, traT, cvaC, ireA, papC,*	O2:H4	ST95	B2	UPEC
S37	9020	Animal faeces	Unknown	GP	PEN, SXT, CAZ, H_2O_2	*sitABCD*	*fimH*	O20:H9	ST2952	B1	APEC
S35	2035	Water	N/A	GP	PEN, SXT, PIP	-	*fimH,*	O175:H16	ST2598	Unknown	APEC

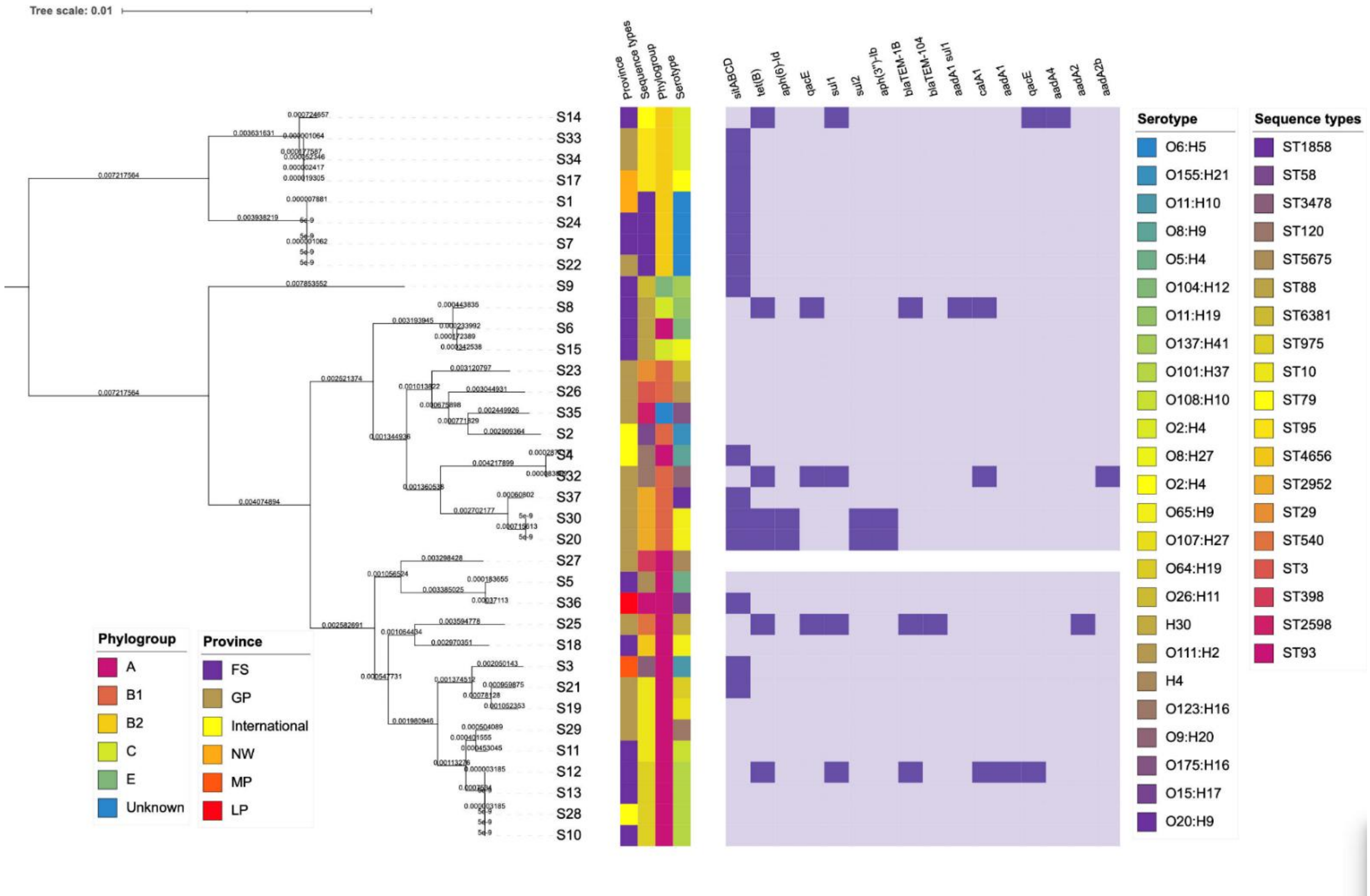


Figure S 1: Heat map of *E. coli* resistance genes, blue colour indicates presence of genes and light blue indicates absence of genes. Left panel of the map indicates the location isolates are sampled from, their phylogroup, and the phylogenetic relationship of isolates. The right panel indicates the serotypes and sequence of these isolates. Provinces represented include: Free State (FS), Gauteng (GP), North West (NW), Mpumalanga (MP), Limpopo (LP) as well as International isolates from ports of entry.

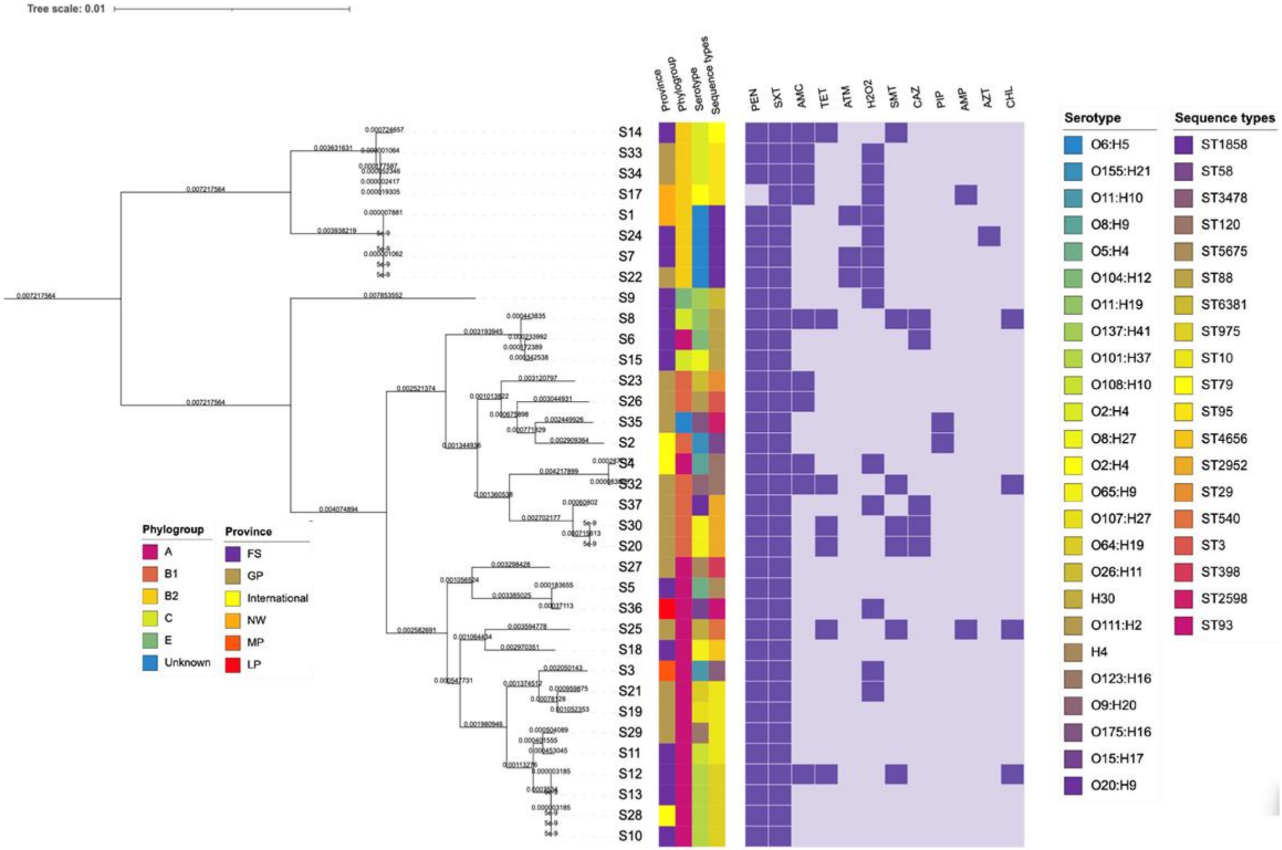


Figure S 2: A heatmap showing antibiotic phenotype of *E. coli* isolates in this study. Blue colour and light blue indicate the presence and absence of the antibiotic phenotype. The top panel of the x-axis indicates the antibiotics detected. Legend of antimicrobials: PEN penicillin, SXT cotrimoxazole, H202 hydrogen peroxide, AMC amoxycillin, SMT sulfamethazine, TET tetracycline, CHL chrloramphenicol, CAZ ceftazidime, ATM aztreonam, PIP piperacillin, AMP ampicillin, AZT Azithromycin. Phylogroups detected: A, B1, B2, C, E and one isolate was unknown. Provinces represented include Free State (FS), Gauteng (GP), North West (NW), Mpumalanga (MP), Limpopo (LP) as well as international isolates from ports of entry

Appendix B: Chapter 4 Supplementary material

Table S 6: Detected LEE proteins and their grouping

LEE encoded proteins
espL2, nleB1, nleE, nleA/espI, espM2, espW, cif, nleH1, nleA/espI, nleH2, nleF, espM1, espR1, espX4, espG, escE, cesAB, escL, escR, escS, escT, escU, etgA, cesD, escC, sepD, escJ, escI, cesL, escV, escN, escO, escP, sepQ/escQ, espH, cesF, map, cesT, eae, escD, sepL, espA, espD, cesD2, escF, escG, espF, nleG7', espX5, espX1
LEE protein groups
LEE1 espG, escE, cesAB, escL, escR, escS, escT, escU LEE 2 cesD, escC, sepD, escJ, escI LEE 3 escV, escN, escO, escP, sepQ/escQ, espH, cesF, map LEE 4 escD, sepL, espA, espD, cesD2, escF, escG, espF, LEE 5 cesT, eae

Table S 7: Physiochemical properties of LEE encoded proteins

Physiochemical properties	**LEE encoded proteins**							
	escU	***escN***	***eae***	***espA***	***espD***	***espG***	*escV*	*escD*
Amino acids	250	446	939	192	380	398	675	406
Molecular weight (Da)	28604.50	48858.45	101684.77	20534.02	39618.76	43905.64	75150.08	45316.26
Aliphatic index	108.32	105.40	76.99	88.49	91.18	88.24	119.81	105.69
Instability index	32.47	35.55	32.60	37.67	37.26	41.70	39.61	39.83
Theoretical pI	9.29	5.99	8.95	4.80	6.65	5.37	5.66	6.96
Total number of negatively charged residues (Asp+Glu)	27	54	78	20	31	46	69	41
Total number of positively charged residues (Arg+Lys)	35	50	89	16	31	35	62	41
Extinction coefficients (all pairs of Cys residues form cystine)	32890	25120	123665	-	-	24785	51465	49070
Extinction coefficients (all pairs of Cys residues are reduced)	32890	24870	123540	7450	25440	24410	51340	48820
Grand average of hydropathicity (GRAVY)	0.046	0.020	-0.358	-0.195	-0.025		0.355	-0.087

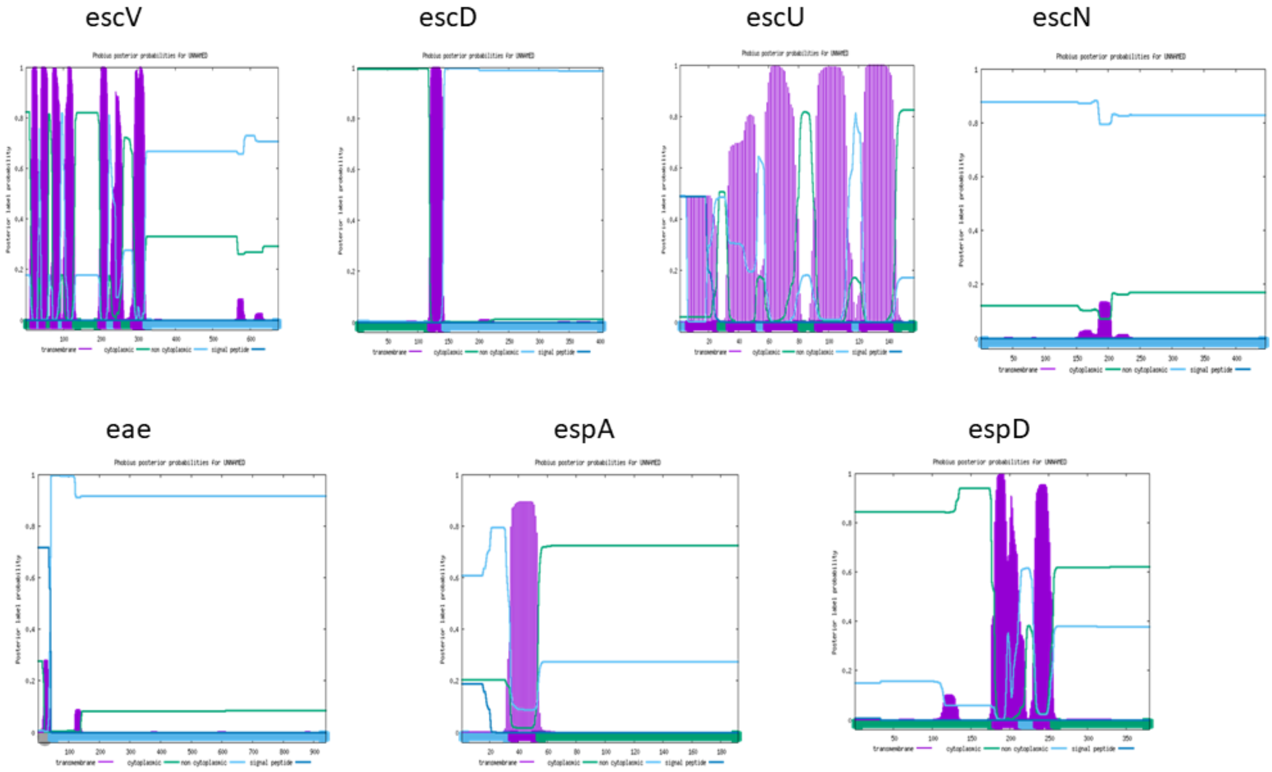
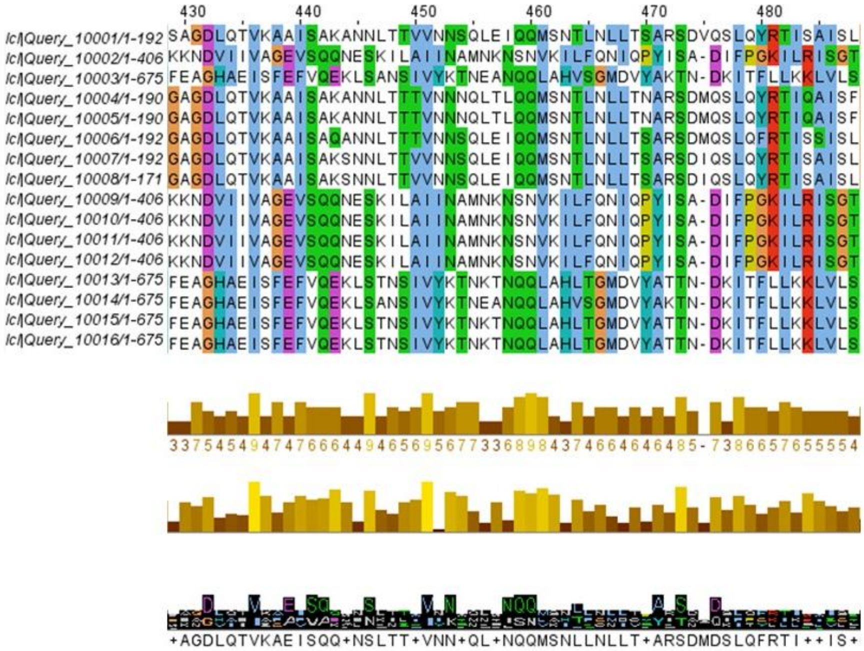


Figure S 3: Predicted transmembrane protein domains of LEE genes using Phobius.



```
                     430       440       450       460       470       480
lcl|Query_10001/1-192 SAGDLQTVKAAISAKANNLTTVVNNSQLEIQQMSNTLNLLTSARSDVQSLQYRTISAISL
lcl|Query_10002/1-406 KKNDVIIVAGEVSQQNESKILAIINAMNKNSNVKILFQNIQPYISA-DIFPGKILRISGT
lcl|Query_10003/1-675 FEAGHAEISFEFVQEKLSANSIVYKTNEANQQLAHVSGMDVYAKTN-DKITFLLKKLVLS
lcl|Query_10004/1-190 GAGDLQTVKAAISAKANNLTTTVNNNQLTLQQMSNTLNLLTNARSDMQSLQYRTIQAISF
lcl|Query_10005/1-190 GAGDLQTVKAAISAKANNLTTTVNNNQLTLQQMSNTLNLLTNARSDMQSLQYRTIQAISF
lcl|Query_10006/1-192 GAGDLQTVKAAISAQANNLTTTVNNSQLEIQQMSNTLNLLTSARSDMQSLQFRTISSISL
lcl|Query_10007/1-192 GAGDLQTVKAAISAKSNNLTTVVNNSQLEIQQMSNTLNLLTSARSDIQSLQYRTISAISL
lcl|Query_10008/1-171 GAGDLQTVKAAISAKSNNLTTVVNNSQLEIQQMSNTLNLLTSARSDIQSLQYRTISAISL
lcl|Query_10009/1-406 KKNDVIIVAGEVSQQNESKILAIINAMNKNSNVKILFQNIQPYISA-DIFPGKILRISGT
lcl|Query_10010/1-406 KKNDVIIVAGEVSQQNESKILAIINAMNKNSNVKILFQNIQPYISA-DIFPGKILRISGT
lcl|Query_10011/1-406 KKNDVIIVAGEVSQQNESKILAIINAMNKNSNVKILFQNIQPYISA-DIFPGKILRISGT
lcl|Query_10012/1-406 KKNDVIIVAGEVSQQNESKILAIINAMNKNSNVKILFQNIQPYISA-DIFPGKILRISGT
lcl|Query_10013/1-675 FEAGHAEISFEFVQEKLSTNSIVYKTNKTNQQLAHLTGMDVYATTN-DKITFLLKKLVLS
lcl|Query_10014/1-675 FEAGHAEISFEFVQEKLSANSIVYKTNEANQQLAHVSGMDVYAKTN-DKITFLLKKLVLS
lcl|Query_10015/1-675 FEAGHAEISFEFVQEKLSTNSIVYKTNKTNQQLAHLTGMDVYATTN-DKITFLLKKLVLS
lcl|Query_10016/1-675 FEAGHAEISFEFVQEKLSTNSIVYKTNKTNQQLAHLTGMDVYATTN-DKITFLLKKLVLS

                      337545494747666449465695677336898437466464648 5-738665765555 4
                      +AGDLQTVKAEISQQ+NSLTT+VNN+QL+NQQMSNLLNLLT+ARSDMDSLQFRTI++IS+
```

Figure S 4: Multiple sequence alignment of T3SS proteins (espA, escD, escV). ClustaW scheme was used for coloring residues in this alignment

Appendix C: Chapter 5 Supplementary material

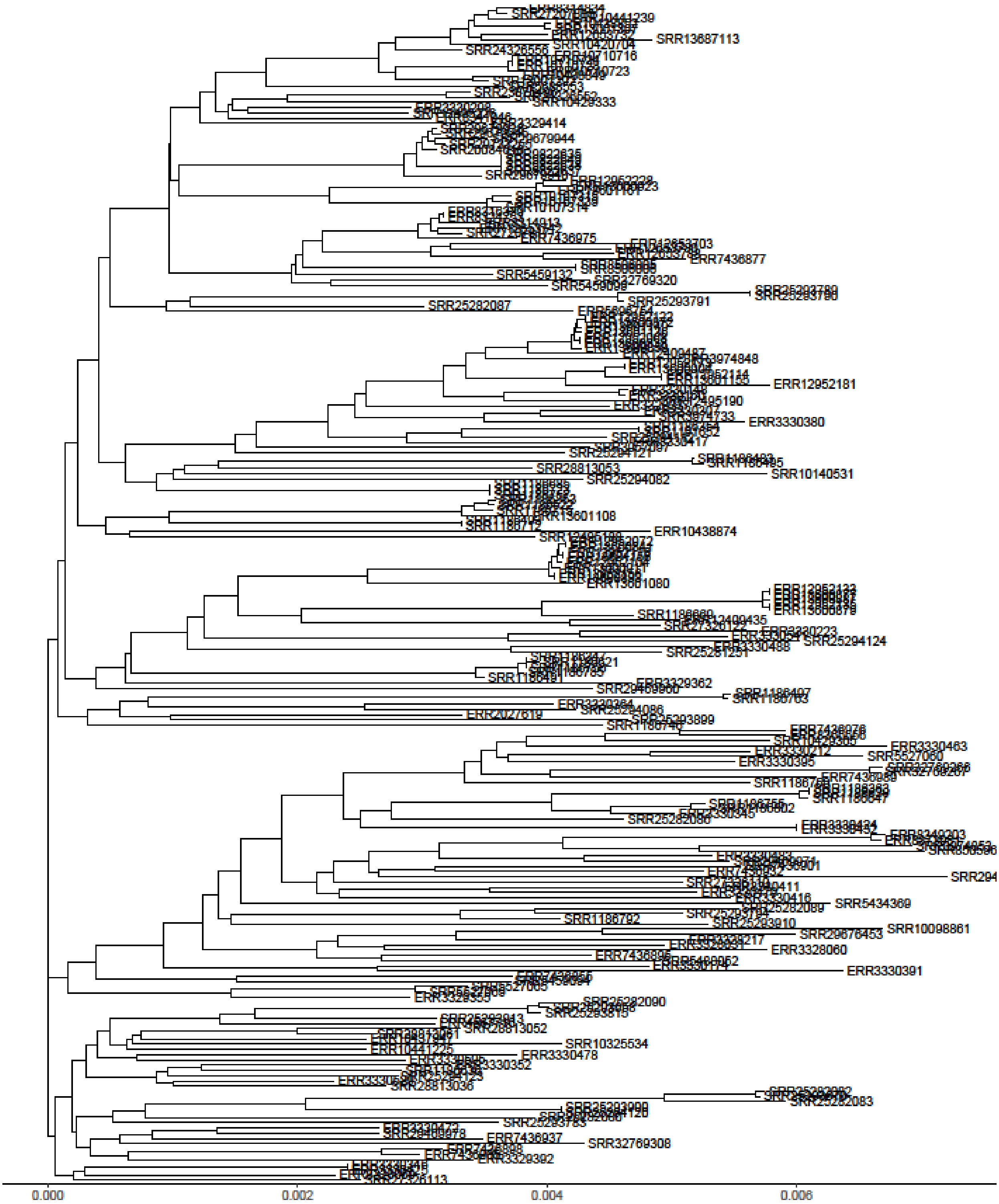


Figure S 5: A mash-distance phylogenetic tree of 331 *E. coli* isolates retrieved from Eneterobase to identify outliers. No outliers were detected in these isolates.

Table S 8: ST10 *E. coli* isolates included in this study (n=159)

Isolate ID	Source Type	Source Details	Collection Year	Country
ERR10437947	Human	Human urine	2020	Ghana
ERR10438849	Human	Human urine	2020	Nigeria
ERR10438852	Human	Human urine	2020	Nigeria
ERR10438874	Human	Human urine	2020	Nigeria
ERR10441225	Human	Human urine	2020	Burkina Faso
ERR10441239	Human	Human urine	2020	Burkina Faso
ERR10710716	Fish	Local fish market	2021	Djibouti
ERR10710723	Human	Pus	2021	Djibouti
ERR10710724	Ovine	Sheep faeces	2021	Djibouti
ERR10710728	Fish	Local fish market	2021	Djibouti
ERR12409435	Human	Chicken faeces	2021	Tanzania
ERR12653703	Water/River	Surface water	2021	Ghana
ERR12653732	Water/River	Surface water	2021	Ghana
ERR12653742	Water/River	Surface water	2021	Ghana
ERR12653780	Water/River	Surface water	2021	Ghana
ERR12653788	Water/River	Surface water	2021	Ghana
ERR12952072	Human	Human faeces	2021	Ethiopia
ERR12952104	Human	Human faeces	2022	Ethiopia
ERR12952116	Human	Human faeces	2022	Ethiopia
ERR12952133	Human	Human faeces	2022	Ethiopia
ERR12952135	Human	Human faeces	2022	Ethiopia
ERR12952156	Human	Human faeces	2023	Ethiopia
ERR13600841	Human	Human faeces	2021	Ethiopia
ERR13600877	Human	Human faeces	2022	Ethiopia
ERR13600879	Human	Human faeces	2022	Ethiopia
ERR13600887	Human	Human faeces	2022	Ethiopia
ERR13600893	Human	Human faeces	2023	Ethiopia
ERR13600911	Human	Human faeces	2022	Ethiopia
ERR13601080	Human	Human faeces	2021	Tanzania
ERR13601108	Human	Human faeces	2022	Tanzania
ERR13601156	Human	Human faeces	2022	Ethiopia
ERR2027619	Human	Human faeces	2014	DRC
ERR3328059	Bovine	Beef meat	2015	Kenya

ERR3329337	Human	Human faeces	2015	Kenya
ERR3329355	Rodent	Rodent faeces	2015	Kenya
ERR3329362	Swine	Pig faeces	2015	Kenya
ERR3329392	Human	Human faeces	2015	Kenya
ERR3330174	Swine	Pig faeces	2015	Kenya
ERR3330223	Human	Human faeces	2015	Kenya
ERR3330298	Avian	Geese meat	2016	Kenya
ERR3330307	Human	Human faeces	2016	Kenya
ERR3330346	Avian	Chicken meat	2016	Kenya
ERR3330352	Rodent	Rodent faeces	2016	Kenya
ERR3330380	Environment	Household environment	2016	Kenya
ERR3330391	Avian	Chicken faeces	2016	Kenya
ERR3330417	Rodent	Rodent faeces	2016	Kenya
ERR3330425	Avian	Chicken faeces	2016	Kenya
ERR3330472	Avian	Chicken faeces	2016	Kenya
ERR3330478	Avian	Chicken faeces	2016	Kenya
ERR3330488	Human	Human faeces	2016	Kenya
ERR3330529	Avian	Chicken faeces	2016	Kenya
ERR3330541	Environment	Household environment	2016	Kenya
ERR3330595	Environment	Household environment	2016	Kenya
ERR4567716	Water/River	Village drinking water	2019	Kenya
ERR5696754	Human	Human faeces	2014	Rwanda
ERR7436855	Human	Human faeces	2019	Madagascar
ERR7436877	Human	Human faeces	2019	Madagascar
ERR7436898	Human	Human faeces	2019	Madagascar
ERR7436905	Human	Human faeces	2018	Madagascar
ERR7436937	Human	Human faeces	2019	Madagascar
ERR7436975	Human	Human faeces	2018	Madagascar
ERR8314834	Human	Human faeces	2017	Niger
ERR8314913	Human	Human faeces	2017	Niger
ERR8316339	Human	Human faeces	2017	Niger
ERR8324753	Human	Human faeces	2017	Niger
ERR8341046	Human	Human faeces	2017	Niger
SRR10140531	Human	Human faeces	2019	Nigeria
SRR10325534	Avian	Poultry litter	2019	Nigeria
SRR10420704	Avian	Poultry litter	2019	Nigeria

SRR10429333	Human	Human faeces	2019	Nigeria
SRR1186347	Human	Human faeces	2009	Tanzania
SRR1186353	Human	Human faeces	2009	Tanzania
SRR1186354	Human	Human faeces	2009	Tanzania
SRR1186483	Human	Human faeces	2009	Tanzania
SRR1186491	Human	Human faeces	2009	Tanzania
SRR1186493	Human	Human faeces	2009	Tanzania
SRR1186495	Human	Human faeces	2009	Tanzania
SRR1186497	Human	Human faeces	2009	Tanzania
SRR1186522	Human	Human faeces	2009	Tanzania
SRR1186621	Human	Human faeces	2009	Tanzania
SRR1186636	Human	Human faeces	2009	Tanzania
SRR1186669	Human	Human faeces	2009	Tanzania
SRR1186685	Human	Human faeces	2009	Tanzania
SRR1186712	Human	Human faeces	2009	Tanzania
SRR1186715	Human	Human faeces	2009	Tanzania
SRR1186723	Human	Human faeces	2009	Tanzania
SRR1186746	Human	Human faeces	2009	Tanzania
SRR1186756	Human	Human faeces	2009	Tanzania
SRR1186757	Human	Human faeces	2009	Tanzania
SRR1186763	Human	Human faeces	2009	Tanzania
SRR1186785	Human	Human faeces	2009	Tanzania
SRR1186792	Human	Human faeces	2009	Tanzania
SRR1191652	Human	Human faeces	2009	Tanzania
SRR12495188	Human	Human faeces	2018	Gambia
SRR12495226	Human	Human faeces	2018	Gambia
SRR13001292	Human	Human urine	2020	Nigeria
SRR13001397	Human	Human urine	2020	Nigeria
SRR13687113	Human	Human faeces	2019	Benin
SRR20084616	Shellfish	Seafood	2015	Tunisia
SRR23075496	Poultry	Raw meat	2017	Nigeria
SRR24326552	Poultry	Raw meat	2017	Nigeria
SRR24326556	Poultry	Raw meat	2017	Nigeria
SRR25281251	Human	Human faeces	2019	South Africa
SRR25282080	Swine	Pig faeces	2019	South Africa
SRR25282082	Swine	Pig faeces	2019	South Africa

SRR25282083	Swine	Pig faeces	2019	South Africa
SRR25282087	Swine	Pig faeces	2019	South Africa
SRR25282090	Swine	Pig faeces	2019	South Africa
SRR25293658	Swine	Pig faeces	2019	South Africa
SRR25293670	Swine	Pig faeces	2019	South Africa
SRR25293783	Swine	Pig faeces	2019	South Africa
SRR25293789	Swine	Pig faeces	2019	South Africa
SRR25293790	Swine	Pig faeces	2019	South Africa
SRR25293791	Swine	Pig faeces	2019	South Africa
SRR25293794	Swine	Pig faeces	2019	South Africa
SRR25293815	Swine	Pig faeces	2019	South Africa
SRR25293899	Swine	Pig faeces	2019	South Africa
SRR25293900	Swine	Pig faeces	2019	South Africa
SRR25293913	Swine	Pig faeces	2019	South Africa
SRR25294082	Human	Human rectal swab	2019	South Africa
SRR25294112	Human	Human rectal swab	2019	South Africa
SRR25294120	Human	Human rectal swab	2019	South Africa
SRR25294121	Human	Human rectal swab	2019	South Africa
SRR25294123	Human	Human rectal swab	2019	South Africa
SRR25294124	Human	Human rectal swab	2019	South Africa
SRR27207815	Poultry	Chicken faeces	2015	Nigeria
SRR27207817	Poultry	Chicken faeces	2015	Nigeria
SRR27326113	Human	Human faeces	2022	Ethiopia
SRR27326122	Human	Human faeces	2021	Ethiopia
SRR28813036	Human	Human faeces	2022	Ethiopia
SRR28813052	Human	Human faeces	2022	Ethiopia
SRR28813053	Human	Human faeces	2022	Ethiopia
SRR28813061	Human	Human faeces	2022	Ethiopia
SRR29469960	Animal faeces	Pig faeces	2018	South Africa
SRR29469978	Meat	Meat	2016	South Africa
SRR29679938	Poultry	Chicken faeces	2015	Nigeria
SRR29679940	Poultry	Chicken faeces	2015	Nigeria
SRR29679944	Poultry	Chicken faeces	2015	Nigeria
SRR29679945	Poultry	Chicken faeces	2015	Nigeria
SRR29733255	Poultry	Chicken faeces	2015	Nigeria
SRR32769308	Reptile	Cloacal swab	2022	Zimbabwe

SRR32769320	Reptile	Cloacal swab	2022	Zimbabwe
SRR3957097	Human	Human faeces	1994	Zambia
SRR3974733	Human	Human faeces	1994	Zambia
SRR5459094	Avian	Chicken faeces	2016	Uganda
SRR5459099	Avian	Chicken faeces	2016	Uganda
SRR5459132	Avian	Chicken faeces	2016	Uganda
SRR5527065	Avian	Chicken faeces	2016	Uganda
SRR5527069	Avian	Chicken faeces	2016	Uganda
SRR5688553	Dairy	Milk cheese	2016	Egypt
SRR8506005	Human	Human faeces	2012	Niger
SRR8506006	Human	Human faeces	2012	Niger
SRR9822635	Poultry	Chicken faeces	2015	Nigeria
SRR9822637	Poultry	Chicken faeces	2015	Nigeria
SRR9822638	Poultry	Chicken faeces	2015	Nigeria
SRR9822649	Poultry	Chicken faeces	2015	Nigeria